



LUND UNIVERSITY

På AI-teknikens axlar

Om kunskapssociologin och stark artificiell intelligens

Kåhre, Peter

2009

[Link to publication](#)

Citation for published version (APA):

Kåhre, P. (2009). *På AI-teknikens axlar: Om kunskapssociologin och stark artificiell intelligens*. [Doktorsavhandling (monografi), Sociologi]. Department of Sociology, Lund University.

Total number of authors:

1

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Avhandlingen diskuterar stark artificiell intelligens. Sådan AI som kan skapa kunskap på egen hand. Det påpekas att det finns en skillnad mellan två huvudformer av AI-design: robotar som uppträder intelligent och intelligenta programvaror som letar efter kopplingar mellan kunskap som finns lagrad på olika håll. Avhandlingen hävdar att intelligenta robotar knappast kan skapa kunskap, men att AI-program som använder internet för att nå distribuerad kunskap kan sägas skapa ny kunskap genom att de kan observera nya relationer mellan data, information och etablerad kunskap.

Sociologen Niklas Luhmann menar att det är sociala system som kommunicerar och inte människor. Att stark distribuerad AI kan fungera har att göra med att de kan bidra till den process som enligt Luhmann behövs för att de sociala systemen kan interagera med sin omgivning. Den processen är viktig för att systemen skall observera nya kunskapsrelationer.



LUNDS
UNIVERSITET

Lund Dissertations in Sociology
ISBN 91-7267-289-7 ISSN 1403-6061
printed by Media-Tryck, sociologen 2009
Department of Sociology, Lund University
P.O. Box 114 SE-221 00 Lund

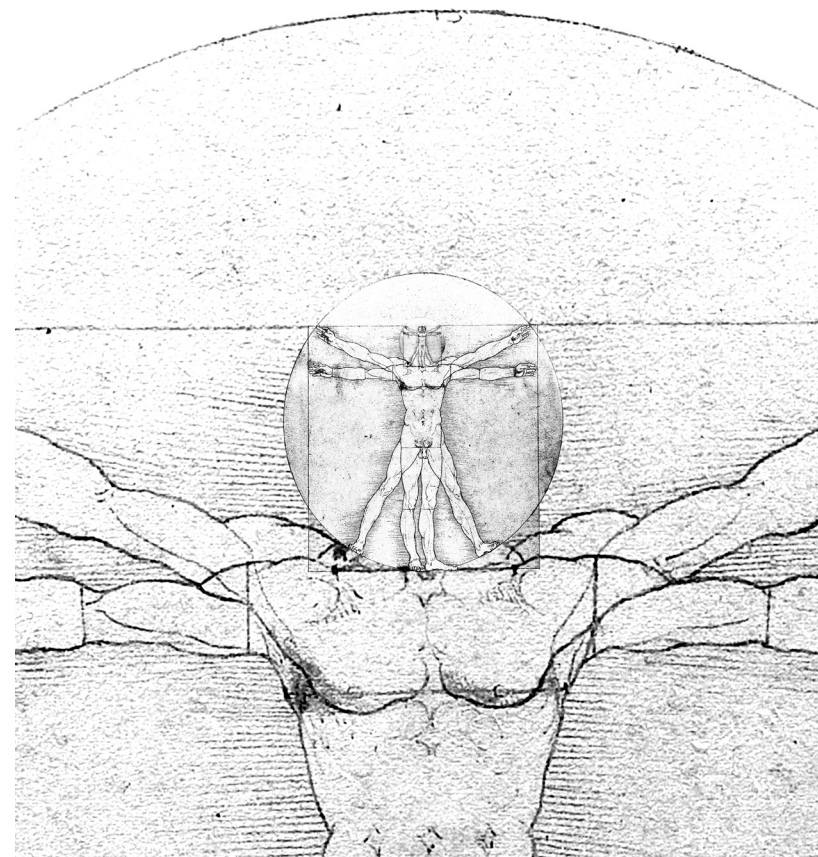
På AI-teknikens axlar

Peter Kåhre

Peter Kåhre

På AI-teknikens axlar

Om kunskapssociologin och stark artificiell intelligens



Lund Dissertations in Sociology 87

PÅ AI-TEKNIKENS AXLAR

Peter Kåhre

På AI-teknikens axlar

Om kunskapssociologin
och stark artificiell intelligens



Lund Dissertations in Sociology 87

En komplett förteckning över
Sociologiska institutionens utgivning
finns sist i boken, och på
www.soc.lu.se/info/publ

COPYRIGHT ©
Peter Kåhre 2009

GRAFISK FORM
Media-Tryck

OMSLAGCOLLAGE
Markus Kåhre

TRYCK
Media-Tryck, Lund University, Lund, Sweden 2009

ISBN 91-7267-289-7
ISSN 1102-4712

Innehållsförteckning

1. Sociologin, artificiell intelligens och social kommunikation.	11
1.1 Inledning.....	11
1.2 Syftesformulering.....	12
1.3 Problemställningens anknytning till sociologisk teori	13
1.3.1 Intelligens som en förmåga att se längre.....	15
1.3.2 Den röda tråden i ett svårt problemkomplex.....	16
1.4 AI som sociologiskt studieobjekt	19
1.5 Två strategier för stark AI.....	22
1.6 Hur syftet skall utredas	25
1.6.2 Diskussion av studieobjektet	26
1.6.3 Debatten om AI	26
1.6.4 Det sociala i AI-diskussionen.....	27
1.6.5 Luhmannsk socialitet och DAI	27
1.6.6 Sociala processer som artefaktisk aktivitet.....	27
1.6.7 Social kommunikation som personlig interaktion	28
1.6.8 Det sociala som holism – Systemteorin och det Starka Programmet.....	28
1.6.9 Ett försök till syntes – det artificiella som funktionalitet i komplexa sociala processer.....	28
1.6.10 Uppsummering av argumentationen	29
1.7 Vad som inte skall göras.....	29
1.8 Metodologiska överväganden	32
1.9 Centrala begrepp i avhandlingen	35
2. Diskussion av studieobjektet.....	37
2.1 Vad lägger vi in i begreppet social, intelligens och kunskap	37
2.2 Kunskapsprocesser som problemlösning eller som expanderings utöver det som vi redan känner till.....	46

2.3 AI som robotar med mänsklig kapacitet eller som program för holistiska sociala processer	47
2.4 Vad innebär social kommunikation ur ett sociologiskt perspektiv	51
2.5 Hur kan komplexa sociala processer modelleras i en sociologisk teoretisk modell?	52
3. Debatten om AI.....	57
3.1 Turings test och jämförelsen med människan.....	57
3.2 Den konnektionistiska modellens vetenskapsteoretiska antaganden.....	61
3.3 Searles kinesiska rum	63
3.4 Kritiken mot Searle	66
3.5 Searles argumentation och diskussionen i denna avhandling	67
3.6 För konnektionism och mot symbolisk AI	69
3.7 Den situationsberoende handlingen, rationalismen och AI.....	70
4. Det sociala i AI-diskussionen	75
4.1 Metodologisk individualism eller Holism och AI	75
4.2 AI och praktisk kunskapsförmedling	76
4.2.1 Randall Collins experiment	76
4.2.2 Relevansproblematiken, människan och möjligheterna för observation ...	78
4.3 Den sociologiska AI-diskussionen	79
4.3.1 Den tidigaste diskussionen	79
4.3.2 Senare diskussion.....	80
4.3.3 Den senaste diskussionen och forskningsprogrammet <i>Socionics</i>.....	80
4.4 Dreyfus, konnektionismen och den sociologiska kritiken	83
4.5 Människans ursprung enligt utvecklingsbiologin.....	88
5. Luhmanns socialitet och Distribuerad AI.....	95
5.1 Second Order Cybernetics, relativismen och det posthumana	95
5.2 Luhmanns systemteori och kunskapsprocesserna.....	100
5.2.1 Distinktionen.....	100
5.2.2 Information som skillnad	101
5.2.3 Observatören	103
5.2.4 Kontingens och dubbel kontingens	104
5.2.5 Strukturell koppling.....	105

5.3 Meningsproblematiken	106
5.4 Luhmanns funktionalism som en radikal samhällsteori	107
6. Sociala processer som artefaktisk aktivitet	109
6.1 Medvetandets behov av redskap	109
6.2 Expanding och redskap	113
6.3 Semiotiken som det artificiella hos Vygotsky.....	115
6.4 Reflexion, social kommunikation och omvärlden.....	117
6.5 Det sociala som text/artefakt.....	122
6.6 Den sociala kommunikationen och handlingsbegreppet.....	125
6.7 Redskapsanvändning och konnektionism.....	128
7. Social kommunikation som personlig interaktion.....	133
7.1 Personlig interaktion och AI	133
7.2 Sociala system som statiska eller föränderliga.....	139
7.3 Biblioteken som ett standardiserat system	141
7.4 Tre perspektiv på AI	145
7.5 Människors kunskapsförmedling i praktiken.....	147
8. Det sociala som holism – Systemteorin och det Starka Programmet.....	151
8.1 Det Starka Programmet	151
8.2 Debatten kring Bloors Starka Program.....	153
8.2.1 Debatten mellan datakognitivister och David Bloor	154
8.2.2 Debatten mellan David Bloor och vetenskapsteoretiker	159
8.3 Luhmanns funktionella relativism och ontologisk realism	162
8.4 Luhmann om kodning och programmering	168
9. Ett försök till syntes – det artificiella som funktionalitet i komplexa sociala processer.....	171
9.1 Kunskapssociologins relativistiska grundsyn	171
9.2 Det sociala som situationism eller som ”environmentalism”	173
9.2.1 Kontextens användning inom sociologisk diskussion	173

9.2.2 Från embodiement till environmentalism	178
9.3 Luhmanns systemteori eller talhandlingsteori	188
10. Uppsummering av argumentationen.....	201
11. English Summary	205
Referenslista.....	211

Tack till:

Denna avhandling har en lång tillkomsthistoria. Jag började arbetet redan i mitten på 90-talet med något som från början var ett försök att diskutera hur man kommunicerar med hjälp av en förmedlande institution som ett bibliotek. Avhandlingen kom så småningom att koncentreras på frågan om Artificiell Intelligens allt eftersom bibliotekens litteraturförmedling blev mer beroende av litteratur lagrad i elektroniska medier och vi såg att sökmotorer som Internet verkligen betydde en skillnad för biblioteksinstitutionens förutsättningar. Det som en gång började som en ambition att placera in biblioteksinstitutionen i Habermas teori om kommunikativt handlande övergick till att bli en avhandling om artefaktens möjlighet av vara en del av den systemkommunikation som Luhmanns teori bygger på.

Under årens lopp har många närstående försvunnit från min krets av samtalspartner. Bland dem som har försvunnit från jordelivet vill jag först och främst tacka min far som länge gjorde det möjligt för mig att våga satsa på något så osäkert som ett avhandlingsarbete.

Jag vill också tacka alla kamraterna på Stockholms Universitetsbibliotek som tagit del av mina tankar under många år. Speciellt John Maule som både har läst min text och kommit med många goda råd. John är en mycket trofast och omdömesgill vän som man alltid kan lita på. Han är dessutom en äkta skåning som gett mig mycket kunskap om det skånska livet på landsbygd och i stad. Många gånger har jag dessutom sovit över i Växjö där han och fru Katarina numera bor. Ett speciellt tack också till Marian Dabrowski som är en god vän och den som ständigt har muntrat upp mig. De år jag bott på Åland har Marian och hans fru Joanna alltid erbjudit sitt gästrum när jag passerat Stockholm.

Ett mycket stort tack vill jag också rikta till Anette Mjöberg som oräkneligt antal gånger erbjudit mig övernattnings i Eslöv. Avhandlingsarbetet tog så lång tid att dottern Mette har vuxit upp från en flicka som jag ibland satt barnvakt åt till att bli en ung kvinna. Under senare år tillkom Anettes make Conny som både bjöd på goda middagar och fin whisky. Tack också till Anettes pappa Jöran som också gick bort innan denna text blev klar.

Jag vill också tacka den glesa kretsen av luhmannianer i Sverige som ibland bidragit med synpunkter på min text. Tack också till prefekten Gunnar Andersson som engagerade sig så att det blev möjligt för mig att till sist lägga fram avhandlingen.

Avhandlingsarbete brukar beskrivas som en ensam process. I mitt fall har det varit ett mycket ensamt arbete. Jag såg en gång en TV intervju där Kristina Lugn berättade att hon levtt ett så ensamt liv att litteraturen blivit hennes bästa vän. I mitt arbete har mina närmaste vänner bestått av de texter som jag refererar till i denna avhandling. Jag vill därför tacka alla som finns med i litteraturlistan till denna avhandling. Ett speciellt tack till Niklas Luhmann som fick mig att observera ett alternativ till en världsåskådning. Jag vill också tacka alla som bidragit till att så

mycket av den vetenskapliga publiceringen numera distribueras i elektroniskt format. Utan den utvecklingen hade det varit svårt att arbeta med denna avhandling från ett så avlägset örike som Åland. Ett tack också till Markus Kåhre som gjort collaget till omslaget och till Jack Hilling som läst korrektur.

1. Sociologin, artificiell intelligens och social kommunikation

1.1 Inledning

I denna avhandling skall jag diskutera sociologins behandling av den datateknologi som går under namnet artificiell intelligens, i fortsättningen använder jag förkortningen AI. Sociologin är ofta tämligen kritisk till AI. Min egen hållning till AI var också från början kritisk. Läsningen av sociologiska texter tyder på att sociologin ofta har en kritisk hållning till teknik som leder till att ett redskap som datoriserad AI bedöms med skepsis. Paradoxalt nog är det fördjupningen i sociologisk teori, speciellt den sociologiska diskussionen av ontologiska utgångspunkter, som lett mig fram till en positiv hållning till möjligheterna av en genuin artificiell intelligens. Den läsningen gjorde mig skeptisk till att sociologin kan bygga på en så kallad metodologisk individualism (i fortsättningen förkortad till MI) som utmärker sig genom att försöka grunda alla sociala beteenden på individuella handlingar. När jag istället kom fram till uppfattningen att sociologin måste grunda sig på en holistisk uppfattning om vad sociala fenomen handlar om, kändes det istället naturligt att sociologin bör förhålla sig positivt till AI. Jag kommer att utveckla denna position i avhandlingen. Här vill jag bara hävda att mitt positiva intresse för AI handlar om att jag tror att sociologin måste inrikta sig på att studera sådana sociala fenomen som handlar om något annat än det som människorna gör med sina medvetna handlingar.

Avhandlingen kommer att fokusera på så kallad **stark AI**, sådan AI som kan skapa kunskap på egen hand. Den aspekten av AI kan uppfattas som att den står väldigt långt från sociologiska frågeställningar eftersom den diskuterar hur AI fungerar utan inblandning av människor. Det är den aspekten på AI som gör att den har så stark koppling till frågan om huruvida sociala processer skall ses som resultat av individuella handlingar eller som resultat av från människorna fristående sociala faktorer. Diskussionen om stark AI har således stora likheter med den

grundläggande sociologiska diskussionen. Det är den senare uppfattningen, den holistiska sociologins position, som blir intressant om vi vill diskutera stark AI ur ett sociologiskt perspektiv. Dessutom kan argument för stark AI användas för att stödja den holistiska sociologins positioner.

Den MI som betonar individuella handlingar är som sagts mycket kritisk till stark AI. Föreläsare för så kallad holistisk sociologi brukar också vara kritiska mot AI därför att de oftast anser att tekniken är styrd av sociala intressen. I avhandlingen kommer vi dock att peka på att stark AI och holistisk sociologi har ett gemensamt problem. Det problemet handlar om att kunna förklara emergens, dvs. hur en holistisk social nivå kan förändra sig själv utan att det handlar om en process som är ett resultat av människors handlingar. En tes som kommer att drivas i avhandlingen är att holistiska sociala processer alltid har varit beroende av en funktionalitet som kommer från de artefakter som finns i en social och humanistisk kulturell struktur. Stark AI kommer då att uppfattas som en ny egenskap i denna struktur, som i sig inte representerar ett nytt fenomen, men som effektiverar funktionaliteten på ett sådant sätt att vi upplever det som ett helt nytt fenomen.

En sådan uppfattning leder också till att vi måste diskutera AI som ett fenomen som kan hjälpa oss att skapa samhällsprocesser som har en bättre emanciperande funktionalitet. Stark AI kan naturligtvis innehålla en sådan funktionalitet att samhället utvecklas i inskränkande riktning. Men det finns också många andra egenskaper i samhället som kan utvecklas i den riktningen. Trots det så uppfattar de flesta sociologer med rätta att emanciperande villkor för individerna bara kan skapas genom att samhällsstrukturerna utvecklas. Mitt intresse för att ge en fördjupning av diskussionen kring stark AI har mycket att göra med att jag uppfattar att stark AI ger en möjlighet för samhället att utvecklas i emanciperande riktning därför att samhällsutvecklingen blir mindre beroende av människors intressen. Det primära syftet med avhandlingen är att ge ett bidrag till den diskussion om stark AI som redan existerar.

1.2 Syftesformulering

I denna avhandling skall jag analysera vilken roll artificiell intelligens har i de sociala processer som utvecklar kunskapen. Analysen görs utifrån den moderna sociologiska systemteorins antaganden om karaktären av social kommunikation och om de sociala systemens förhållande till sin omvärld. Framför allt såsom denna systemteori har utvecklats av den tyske sociologen Niklas Luhmann (1927-1998). Den analysen kommer att fördjupas genom att jag också lyfter fram konkurrerande sätt att se på sociala processer inom sociologin, samt genom att presentera den kritik mot stark AI som kommer från representanter för denna sociologi.

Diskussionen förs utifrån ett antagande om att dagens materiella förutsättningar för kunskapsspridning genom elektronisk publicering och elektronisk kommunikation, ger nya förutsättningar för uppfattningen om hur AI kan vara en del av sociala processer kring kunskapsskapande.

Läsaren bör observera att min diskussion handlar om de argument för och emot AI som har lagts fram. För många kan det vara svårsmält att artificiella program kan vara lika kapabla som människor. Utifrån ett holistiskt sociologiskt perspektiv finns det många möjligheter att argumentera för att AI kan vara stark. Men det finns faktiskt inget i denna avhandling som kan bevisa detta och det hör inte till syftet. Till syftet hör däremot att analysera den existerande diskussionen kring stark AI. Utifrån den analysen kan man hitta rätt så mycket som talar för att det finns rimlighet i argumenten för stark AI. Dessutom understryker jag att väldigt mycket av det som den sociologiska disciplinen lär oss om sociala faktorer, faktiskt talar för att sociologer bör bejaka AI.

1.3 Problemställningens anknytning till sociologisk teori

Frågan om stark AI debatteras i stor utsträckning utifrån antaganden om att det är sociala processer som medger att ny kunskap kan skapas. Utifrån Luhmanns systemteori betyder det att sociala processer tillför en funktionalitet som medger att de sociala systemen kan gå utöver vad som finns inom ett system. Med Luhmanns terminologi innebär det att de sociala systemen kan bli påverkade av vad som finns i systemens omvärld. Den AI-kritik som levereras av framför allt symboliskt interaktionistiskt orienterade sociologer, tar oftast fasta på att sociala processer som skapar ny kunskap fungerar genom att människor interagerar med varandra. Därför menar man att stark AI inte kan skapa ny kunskap utan inblandning av människor.

Det finns en del i Luhmanns teori som talar för att AI-system inte kan skapa ny kunskap. En aspekt av teorin handlar om att systemen är operativt slutna. I så fall kan inte något nytt uppstå inom systemen. Men man kan också läsa Luhmann som att han framför allt talar om de funktioner som får systemen att förändras, även om den förändringen formellt måste handla om att ett gammalt system ersätts av ett nytt system. Dessa förändringar skapas genom att omvärlden irriterar systemen så mycket att de till slut inte kan upprätthålla sina funktioner. Det intressanta med den form av AI-design som vi diskuterar i denna avhandling, är att den handlar om att utnyttja omvärldens komplexitet för att förändra systemen. Det sker genom att systemen samtidigt som de är operativt slutna också är kognitivt öppna. Att systemen är operativt slutna bör däremot vara ett argument för att en form av strategi för stark AI som kallas klassisk AI inte kan fungera.

I en artikel om Luhmanns teori i tidskriften *Organization* framhåller författarna att Luhmanns teori består av två delar. Den första delen handlar om autopoiesis, d.v.s. hur sociala system upprätthåller sina egna funktioner. Den andra delen handlar om hur distinktioner uppstår (Seidl&Becker 2006:11-12). Det är den senare aspekten på Luhmanns teori vi måste ta fasta på när vi diskuterar hur de sociala processer som utvecklar ny kunskap kan se ut. Det finns möjligheter att tolka Luhmann som att han menar att kommunikation inom sociala system inte kan utveckla ny kunskap eftersom den typen av kommunikation bara syftar till att upprätthålla systemets funktioner. Den handlar bara om att kommunicera den etablerade kunskap som ges av rådande systemreferens. Vi kommer att se att det finns många som menar att det är kommunikation som leder till att ny kunskap skapas. Det finns anledning att framhålla att kommunikationen måste ha en speciell karaktär om den skall leda till något nytt. I denna avhandling kommer vi att visa på att kommunikation som bara handlar om att människor samtalar med varandra sällan leder till något nytt. Den kommunikationen måste få en kapacitet som artefakter tillför om den skall leda framåt. Det är först då som den kan leda till att nya distinktioner skapas. Den kommunikation som skapar distinktioner (som leder till ny kunskap) måste skapa en relation mellan systemen och deras omvärld. Här är problemet att Luhmann allt som oftast hävdar att kommunikation bara är något som sker inom det sociala systemet. Det finns delar i Luhmanns texter som kan tolkas som att kommunikation också handlar om relationen mellan systemen och deras omvärld. Det viktiga är dock att det är de processer som skapar distinktion som är relevanta för hur ny kunskap skapas. Och vi kommer senare in på att det finns moderna luhmannianer som pekar på att Luhmanns kommunikationsbegrepp behöver utvecklas så att det också inbegriper kommunikation mellan systemen och deras omvärld. Hur som helst handlar det om processer där kommunikation inte handlar om att människor kommunicerar med varandra. Det är det perspektivet som ger en möjlighet för att stark AI kan vara del av den sociala kommunikationen. Kommunikationsbegreppet i sig är ett problem eftersom Luhmanns användning av begreppet skiljer sig från de flesta andras. De kritiker av AI som stödjer sig på argument som handlar om kommunikation menar att det handlar om att personer samtalar med varandra. Enligt Luhmann kan människor inte kommunicera, det är bara sociala system som kommunicerar.

Om de sociologer som är kritiska till AI alls uppskattar AI, så anser man att AI-system bara kan vara redskap för individerna. Den uppfattningen kallas för **svag AI** och är inte alls kontroversiell. Det finns möjlighet att hävda att det framför allt är denna form av AI som gör att människorna blir påverkade av intressen nedlagda i tekniken. Eftersom stark AI skall fungera utan påverkan från människor bör den vara mer neutral.

Ur ett holistiskt perspektiv handlar frågan om stark AI inte bara om teknikens kapacitet. Den handlar också om att det blir ifrågasatt om enskilda individer kan

skapa kunskap på egen hand. Frågan om stark AI handlar då mer om vad det är för funktioner i de kunskapsskapande processerna som har den mest avgörande betydelsen. Är det den funktionalitet som människorna ger processerna, eller den funktionalitet som kommer genom att artefaktiska system som språk, litteratur och konst, också förmedlar mellan system och omvärld. Om det finns tillräckligt många avgörande funktioner i de artefaktiska systemen, finns det fog för att tala om att AI verkligen kan vara stark. Frågan om stark AI, liksom frågan om människornas kapacitet, blir då i själva verket en graderingsfråga.

1.3.1 Intelligens som en förmåga att se längre

Titeln på denna avhandling är inspirerad av boken *On the Shoulders of Giants* av den amerikanske kunskapssociologen Robert K. Merton (1910-2003). Mertons bok kan säkert tolkas på många sätt. Jag vill framhålla att det finns mycket i texten som handlar om att människorna inte tänker sig fram till sanningar genom att använda sina medvetanden. Framför allt måste förstas Mertons bok ses som en argumentering för att tänkandet alltid måste bygga på det som andra människor har tänkt. Det är emellertid ett tämligen trivialt påstående som bara blir intressant om vi funderar mera på vad det egentligen innebär att vi använder det som andra redan tänkt ut.

Mertons bok blir intressant just därför att han angriper uppfattningen om att det tänkande som andra redan tidigare presterat, faktiskt har samma innehåll som de som senare använder detta tänkande, uppfattar innehållet i de texter i vilka detta tidigare tänkande finns bevarat. Dessutom skriver Merton i efterordet till en senare bok, *The Travels and Adventures of Serendipity*, att den kunskap som upptäckts inte är den samma som sedan levereras i vetenskapliga texter (Merton, 2004:272-273). Han hänvisar till Francis Bacon som menade att kunskap aldrig presenteras i samma ordning som den har upptäckts (Bacon, [1605]1863:70).

Titeln på Mertons bok anger också att kunskapsprocesser bygger på att klassiska kunskapsgiganter skapar de perspektiv som nya kunskapsproducenter behöver ha för att kunna skapa något nytt. På jättarnas axlar innebär att dvärgar kan se längre än dessa jättar genom att de får bättre förutsättningar att observera vad som sker i världen. Det handlar inte om att jättarnas tänkande påverkar dvärgarnas tänkande genom samtal eftersom vi tänker oss att sådan påverkan sker genom att de personligen möter varandra i ett tankeutbyte. Då borde titeln vara *Genom jättarnas huvud*. Istället går denna påverkan genom texter där det kan gå långt tid mellan att ett budskap formulerats och tagits emot. Mertons perspektiv anger snarare att människor, som har mindre intelligens än kunskapsgiganterna, ändå kan bli mer kapabla än dessa giganter genom att stå på deras axlar.

Uttrycket på jättars axlar är en mycket gammal metafor för att kunskapsutveckling inte bara handlar om att människor skapar ny kunskap genom att tänka. Vi måste också ha redskap som tillåter oss att komma längre än vad vi

kan nå genom att bara använda vår hjärna. Ett sådant redskap bör vara den datateknik som används för att skapa artificiellt intelligenta datasystem. Merton har också sociologiserat begreppet *serendipity* i boken *The Travels and Adventures of Serendipity* (2004).¹ Den var länge outgiven, men Merton refererar i *On the Shoulders of Giants* till det då outgivna manuskriptet (Merton1993:156). Serendipity beskriver det fenomen som man upplever som att man snubblar över relevant kunskap utan att man aktivt har letat efter just den kunskapen. En intressant detalj med Mertons (och Barbers) behandling av begreppet *serendipity* är att de anser att det åtminstone inte bara handlar om en slump när vi exponeras för sådant som vi inte sökte efter. Istället handlar det om att det skapats en struktur som är gynnsam för att serendipitet skall uppstå. Frågan är om en sådan struktur måste skapas av människor eller om den kan uppstå genom autonoma processer.

För att utreda syftet kommer jag att fokusera analysen på en strategi för att bygga system för stark AI som kallas *Distribuerad Artificiell Intelligens* (DAI). Till skillnad från traditionella AI-strategier som försöker skapa AI-system som skall motsvara den förståelseförmåga som människans hjärna har, handlar det här om att AI-system kan förbättra människornas möjligheter att observera över större områden. På så sätt får människorna tillgång till artificiella axlar som de kan sitta på. Det är ett perspektiv som ger anledning att fråga vad som genererar funktionaliteten – människan eller redskapen? Merton har också lanserat begreppet latent funktioner. Alltså funktioner som inte är avsiktliga men som uppstår eller framkommer efterhand som någon form av bieffekt. Mitt intresse för AI-tekniken handlar om att jag uppfattar att den tekniken innehåller många latent funktioner som kan revolutionera kunskapsprocesserna. Viktigare är att latent funktioner anger att funktionalitet inte bara leder i en riktning. Att funktioner kan vara latent innebär att begreppet får en annan innebörd än den vanliga som anger någon form av ändamålsenlighet. Hos Luhmann betyder också funktion något som står tämligen neutralt gentemot ändamålsenlighet.

Det finns också en annan likhet mellan Luhmanns teori, försöken med artificiell intelligens och Mertons bok *On the Shoulders of Giants*. Det handlar om att samma skepsis som visas mot Luhmanns teorier och mot försöken att utveckla AI, också till en början visades mot Mertons bok (Simon&Gagnon,1967:424-425). Orsaken är antagligen att alla tre förhållningssätten anger att människans medvetna tänkande och analyserande bara har en relativ betydelse. Numera lär denna bok vara den som är mest uppskattad av Mertons produktion.

1.3.2 Den röda tråden i ett svårt problemkomplex

AI-frågan ger många problem med att formulera den klara problemställning som en avhandling bör ha. Det beror på flera omständigheter. En omständighet har att

¹ Som han har skrivit tillsammans med Elinor Barber.

göra med att diskussionen kring AI på intet sätt är homogen. En orsak till det har att göra med att diskussionen ofta förs som en debatt mellan föreläsare för olika kunskapsområden. Föreläsare för de olika formerna av stark AI för fram tämligen annorlunda argument för sina olika strategier fastän de så att säga borde befinna sig inom samma lag. Argumenten för och emot stark AI representerar radikalt olika ståndpunkter i de flesta dimensioner; vetenskapsteoretiska utgångspunkter, synen på vad människan och den mänskliga kulturen är, programmets kapacitet och vad kunskap och intelligens egentligen innebär. Det är inte enkelt att enhetligt presentera den diskursen. Johan Asplund, som jag tar upp i ett flertal sammanhang som en kritiker av AI, har skrivit att han inte kunde få någon känsla av fullbordan när han lämnade in manuskriptet till den bok om AI som han skrivit – *Genom Huvudet*. I det avseendet är vi helt överens. Det finns också en stor skillnad i synen på vad AI-program egentligen skall klara av. Den vanligaste uppfattningen handlar om att programmen skall kunna tänka på samma sätt som människor gör när de förmodas skapa ny kunskap. Den litteratur som används för att argumentera för möjligheterna av stark DAI, bygger istället på att skapandet av ny kunskap handlar om att kunna gå utöver det som vi kan göra med tänkandet. Bägge dessa positioner måste presenteras, vilket inte låter sig göras på ett enkelt sätt.

En annan omständighet som skapar oklarhet har att göra med ett fenomen som är starkt kopplat till en av de röda trådarna i denna avhandling. Den tråd som handlar om att kunskapsskapande kräver att människan kan gå utöver det som hon redan kan hantera med sitt medvetande. Stark AI är ett fenomen som är så främmande för en människa som har lärt sig att nå all kunskap genom sitt eget medvetande, att hon har väldigt svårt att rent känslomässigt ta till sig allt man behöver ställa på huvudet för att kunna diskutera stark AI. Och även om man kan komma över denna känslomässiga fallgrop har vi inte de analysredskap som behövs för att rationellt kunna formulera problemet. Det ligger helt enkelt på sidan om medvetandets förmåga och därmed också utanför det som enkelt kan formuleras med analyserbara frågeställningar.

Jag har många gånger försökt att formulera problemställningen på ett konkret sätt men det har aldrig funnits något praktiskt perspektiv som förmår belysa hela problematiken. Jag vill påstå att problemformuleringen nog är den del av avhandlingen som jag har ägnat mest tid åt av allt i hela texten. Att det inte går att formulera en enkel och preciserad problemställning har jag efterhand kommit underfund med har att göra med en av de röda trådarna i avhandlingen. Att ny kunskap är beroende av att människan ges nya möjligheter att observera genom att situationerna förändras. Den problemformulering jag har gett försöker beskriva avhandlingens syfte utifrån det sociologiska teoretiserande som jag har kommit fram till att är det mest relevanta – Luhmanns systemteori. Det krävs en viss insikt i Luhmanns teori för att till fullo förstå vad som ingår i den problemformuleringen.

Det tog mig mycket länge att alls kunna greppa hur problemen kan diskuteras i ett sociologiskt sammanhang. En av orsakerna till det är att den

kunskapssociologiska litteratur som har skrivits om teknikanvändning på många sätt leder oss fel. Den leder till att fokusera på faror med teknik. Och hindrar oss att se på teknikens emanciperande möjligheter. Detta förhållningssätt hos sociologer kan innebära att det inte alls går att skriva en sociologisk avhandling om stark AI. Kanske är tiden inte mogen för en sociolog att skriva om stark AI. Å den andra sidan bör det vara så att tiden bara kan bli mogen om någon först leder observationerna i rätt riktning. Och jag refererar i min text till en ganska stor existerande sociologisk diskussion om AI. Den positiva sociologiska diskussionen om AI bygger på ett ifrågasättande av så mycket som anses givet inom sociologin att den som vill förstå den röda tråden i min text nog måste ge sig tid att sätta sig in i detta annorlunda sociologiska perspektiv.

Många sociologer verkar sålunda vara tveksamma till att AI frågan är ett intressant sociologiskt studieobjekt. Därför innehåller detta kapitel till stor del en diskussion om varför AI är ett intressant sociologiskt ämne. Dessutom har så gott som hela kapitel 2 tillkommit som en beställning på ett klargörande hur den sociologiska vinklingen vi behöver för att diskutera AI-frågan skiljer sig från mer traditionella sätt att analysera teknik utifrån en sociologisk utgångspunkt. Denna framställning behövs möjligen för att ge läsaren förutsättningar att förstå hur diskussionen om AI skiljer sig från traditionell sociologi. Framställningen visar på vilka oklarheter som finns kring sociala kunskapsprocesser utan att det är möjligt att där visa på vilka lösningar vi kan erbjuda med det analytiska grepp som används i avhandlingen. På så sätt kan denna förklaring också skymma avhandlingens syfte.

Jag har försökt hitta det analytiska redskapet hos den luhmannska teorin. Problemet med den teorin är, som redan antytts, att den inte är entydig. Den analys av AI som har gjorts med hjälp av systemteorin har dessutom byggt på en utveckling av Luhmanns teori. Denna utveckling av teorin har gjorts av elever till Luhmann. I det arbetet har de använt mer komplicerade delar av teorin. Denna teoriutveckling kan upplevas som mycket mer komplicerad än den presentation som ges i många översiktsverk. För min del vill jag påstå att mycket av det motstånd som finns mot Luhmanns teori handlar om att man nöjt sig med att läsa dessa översiktliga beskrivningar av teorin. Uppsalasociologen Vessela Misheva menar att Luhmanns teori bara kan förstås inom sitt eget teoretiska universum (Misheva,2005:26). Viktigast ur mitt perspektiv är att det inte går att ge ett bidrag till AI-diskussionen bara genom att redovisa de enklare beståndsdelarna i teorin. För att kunna använda teorin för att analysera AI behöver den dessutom utvecklas med hjälp av andra teorier som har använts i AI-diskussionen.

Avslutningsvis vill jag peka på att avhandlingen innehåller tre olika temata. Det första handlar om en presentation och diskussion av AI-frågan. Det andra handlar om den Luhmannska systemteorin och dess förhållande till AI. Det tredje handlar om grundläggande sociologiska utgångspunkter vilka är nödvändiga att diskutera därför att sociologins förhållande till AI fördunklas av disciplinens oklarhet kring karaktären av ett så centralt begrepp som det sociala. Dessa tre

temata bildar var för sig en tråd som skall bidra till att belysa möjligheterna av att stark AI kan vara en del av de sociala processer som är inblandade i skapandet av ny kunskap. Den genomgående tråden handlar om att diskutera vilken roll redskap och artefakter spelar i sociala kunskapsprocesser och om det går att placera in stark AI i dessa processer. I likhet med Randall Collins menar jag att skapandet av AI-program framför allt är något som kräver en insats från sociologer (Collins, 1992:vii).

1.4 AI som sociologiskt studieobjekt

Jag har lärt mig att det finns tre inriktningar på sociologers invändningar mot att se AI som ett sociologiskt studieobjekt.

Den **första** handlar om att AI inte är något verkligt. Man menar att det inte finns några egentliga AI-system som man kan diskutera vare sig med empirisk metod eller med analytisk diskussion. Det bör dock finnas många sociologiska frågeställningar som inte manifesterar sig som konkreta studieobjekt. T.ex. diskrimineringen av invandrare på arbetsmarknaden kan naturligtvis bara ses som mer allmänna tendenser. Det är inte en entydig mekanism som leder till att invandrare alltid diskrimineras. För vi vet så klart att inte alla invandrare diskrimineras. I denna avhandling kommer vi att utgå från att intelligens handlar om en kontextfaktor och att artificiell är ett begrepp som också handlar om sådant som böcker och konstföremål. Det leder till att artificiell intelligens har förekommit så länge som människan har använt olika former av redskap för att utveckla sig som en social varelse. AI är på så sätt något som i allra högsta grad har varit verkligt under en lång tid, men som vi sällan kan urskilja som ett avgränsat fenomen. AI handlar därför om något som står för en del av den funktionalitet som finns inblandad i sociala processer. AI representerar något verkligt, men AI har samtidigt en sådan karaktär att det blir svårt att exemplifiera fenomenet med hänvisning till konkreta företeelser. Därmed liknar begreppet AI begreppet interaktion, som ofta kommer upp i diskussionen om hur kunskapsprocesser fungerar. Vi menar ofta att de processer som skapar kommunikation mellan människor handlar om interaktion. Men för den skull kan vi inte diskutera interaktioner som något konkret. Det finns tekniska redskap som underlättar både intelligenta processer och processer som handlar om interaktion. Vi kan inte ta några sådana system som ett exempel på hur intelligens eller interaktion fungerar. Det vi måste göra är istället att försöka sätta in de respektive begreppen i en teoretisk diskussion och försöka diskutera vilken roll de kan ha i sociala processer. Att vi kan uppleva att skolan idag inte är maximalt anpassad för att underlätta elevens interaktioner med lärare och undervisningssystem, är naturligtvis inte något

som vi kan ta till intäkt för att skolundervisning aldrig kan leda till att individerna kan lära sig matematik, språk o.s.v. Genom att diskutera vad som bör ingå i interaktionsbegreppet kan vi istället diskutera vilka möjligheter det ändå finns att utveckla skolan så att den bättre hjälper eleverna att skaffa sig den kunskap de behöver. Samma förhållande gäller för hur vi skall värdera den artificiellt intelligenta datateknik som vi idag har. Det är ett praktiskt redskap som har en långt kortare utvecklingsperiod bakom sig än den moderna skolan. Vi kan inte värdera skolans möjligheter att hjälpa eleverna att skaffa sig kunskap genom att ta fasta på dess nuvarande tillkortakommanden. På samma sätt är det ännu mindre rimligt att värdera AI-systemens kapacitet genom att ta fasta på bristerna i de AI-system som idag finns. Däremot kan vi naturligtvis lära oss något av hur existerande system fungerar. Av dagens skolsystem kan vi sannolikt lära oss att det skulle vara ännu sämre om vi övergav skolan och istället satsade på att barnen skall skaffa sig utbildning via de personer som finns i deras egen sociala miljö. Att de AI-system som idag finns faktiskt tycks tillföra kvaliteter till kunskapsprocesserna, bör göra oss mer intresserade av att analysera vilken kapacitet dessa system kan ha om vi utvecklar dem.

Den **andra** invändningen handlar om att ett tekniskt system aldrig kan vara ett sociologiskt studieobjekt. Det är också ett påstående som gör att väldigt mycket av det som sociologer studerar skulle diskvalificeras som objekt för den sociologiska analysen. Ett utbildningssystem är naturligtvis till stora delar ett tekniskt system. Vi kan istället hävda att både AI och utbildningssystem är typiska uttryck för sådant som skapar det sociala systemet. Om vi inte studerar sådana system blir vi tvungna att studera människan som en individ utanför det sociala systemet. Således skulle sociologin på sin höjd reduceras till psykologi. Det finns därför goda grunder för att hävda att studiet av AI-teknik är ett grundläggande sociologisk studieområde.

En **tredje** invändning handlar om att man menar att även om ett tekniskt system kan studeras som ett resultat av mänskliga handlingar, kan det aldrig studeras som något som ersätter den mänskliga kapaciteten. Denna invändning bör vara den som är allra mest relevant för en studie av stark AI. Stark AI innebär nämligen att vi måste föra diskussionen på ett sådant sätt att det blir möjligt att fånga upp att sociala processer handlar om något annat än att människor interagerar. Det angreppssättet gör att frågan om stark AI måste diskuteras som ett teoretiskt objekt som måste diskuteras utifrån ontologiska frågeställningar. Den AI-kritiska diskussionen förs sällan utifrån sådana utgångspunkter. Detta är ett faktum som gör att den sociologiska diskussionen sällan förs utifrån ett observationsperspektiv som tillåter diskussionen att se hela problematikens vidd. Vi kommer att se att AI är en bra metafor för att diskutera många grundläggande frågeställningar kring holistisk sociologi. Det handlar om sådant som kausalitet, mekanismer och funktionalitet. Och om hur en social nivå kan fungera utan att hela tiden behöva knytas till människors handlingar. Vi kan därför på goda grunder

hävda att studiet av stark AI bör utgöra ett av den moderna sociologins verkligt viktiga analysområden.

Mycket av det intresse som sociologin visar för att diskutera AI, handlar som sagts om att se hur den funktionalitet som AI-systemen innehåller blir styrd av sociala intressen. Man menar att det är det enda sätt man kan förhålla sig till AI. Den kritiken riktas mot en form av AI-system som kallas symbolisk AI. Det är en form av AI-teknik som bygger på att den programmeras med givna regler. Denna avhandling hävdar att det är en korrekt uppfattning att den formen av AI aldrig kan skapa ny kunskap på egen hand, eftersom den alltid är beroende av de intressen som människorna lägger i deras funktionalitet.

Skulle det gälla för alla AI-system skulle det samtidigt innebära att sociologin måste vara skeptisk till att det går att använda den nya tekniken för att utveckla samhällets strukturer så att medborgarna ges bättre möjligheter att ta till sig kunskap. Det skulle nog betyda att sociologin ställer sig utanför ett av de mest tydliga avtrycken i samtiden.

Den analys av AI som vi för i denna avhandling diskuterar huruvida AI-system kan skapa kunskap på egen hand genom att de får förmågan att programmera om sig själva. I teknisk mening handlar det om huruvida så kallad konnektionistisk programmeringsmetod kan skapa något som kallas genetiska program. Vi bör uppmärksamma att den frågeställningen har stora likheter med frågor vi måste ställa för att diskutera om sociala processer kan förändra samhället utan att förklara dessa processer med mänskliga handlingar.

Vi bör framhålla att en diskussion av DAI innebär att strukturernas eller kontexternas betydelse ökar. Det innebär i sig att omfånget av det som blir relevant ökar exponentiellt. Det som är relevant för intelligens handlar då inte bara om funktioner som kan jämföras med vad som sker i det mänskliga medvetandet. Eftersom komplex funktionalitet i ett DAI-perspektiv handlar om kontexternas förmåga att exponera komplexitet, blir bland annat frågor som handlar om konstnärliga värden som design och annan estetisk kapacitet relevant. Vi blir dock tvungna att begränsa diskussionen till sådana frågor som hör till den vetenskapliga kunskapsförmedlingen. Av de konstnärliga frågeställningarna blir då endast kvar sådana som kan kopplas till litteraturens förmåga att ge komplexitet.

Den debatt om AI som förs är ofta hätsk och känslofylld. Debattörerna diskuterar sällan varandras argument. Därför försöker man oftast angripa motståndaren genom att misstolka eller förlöjliga dennes argument och genom att lägga fram fler argument för sin egen sak. Detta innebär att min redovisning av debatten måste välja bort det som bara handlar om markering av positioner. Istället måste jag konstruera debatten så att den blir mer konsekvent än hos dem som debatterar.

Den konstruktion jag gör av debatten kring AI bygger på att följande delfrågeställningar är centrala för debatten om stark AI:

1. Vad ingår i begreppen social, intelligens och kunskap
2. Skall kunskapsskapande diskuteras som problemlösning eller som expanderings utöver det som vi redan känner till.
3. Skall AI diskuteras som ett försök att skapa robotar som skall jämföras med mänsklig kapacitet eller som ett försök att efterlikna holistiska sociala processer.
4. Vad innebär social kommunikation ur ett allmänt sociologiskt perspektiv och speciellt ur den luhmannska teorins perspektiv.
5. Hur kan komplexa sociala processer modelleras i en sociologisk teoretisk modell.

Vi skall i nästa kapitel utveckla den diskussion om studieobjektet som vi har startat i detta kapitel genom att närmare redogöra för vilka problem som finns inom ovanstående fem frågeställningar. I detta kapitel skall vi fortsätta med en närmare beskrivning av stark AI, hur syftet skall utredas och vad som inte ingår i den analysen samt ta upp några metodproblem.

1.5 Två strategier för stark AI

Det finns två huvudsakliga strategier för att skapa AI-system. Den tidigaste AI-strategin byggde på att AI-systemen skulle behandla kunskapen som om den utgjorde symboler för lagbundna storheter i världen. Denna AI-strategi, som brukar kallas **Klassisk AI** eller **Symbolisk AI**, var dessutom knuten till tanken på att AI-forskningen handlade om att skapa artificiella hjärnor som skulle motsvara den mänskliga hjärnans kapacitet. De AI-strategier vi här diskuterar bygger på något som oftast kallas **Konnektionistisk AI**. Den systemorienterade DAI-strategi som vi diskuterar i denna avhandling använder konnektionistisk programmering för att skapa den funktionalitet som behövs. Det finns flera sätt att definiera *konnektionism* beroende på ur vilken synvinkel vi ser på begreppet. Det är knappast möjligt att ge begreppet en entydig definition. Vi kan emellertid ta del av några beskrivningar för att se i vilka sammanhang begreppet används.

I *The Oxford Companion to Philosophy* (1995:151) ges en beskrivning av Konnektionism eller med det ord som också används här Parallell Distributed Processing. Den tar fasta på att det handlar om ett försök att göra en realistisk modell av hjärnan och mentala processer som inte handlar om att manipulera symboler. Istället handlar det om att se på mentala processer som en komplex väv

av mångdimensionerade nätverk. I en sådan modell sker informationsbearbetning parallellt och distribuerat.

An approach in artificial intelligence and cognitive science aimed at producing biologically realistic models of the brain and of mental processing; sometimes called PDP (parallel distributed processing). In 'old-fashioned' (or cognitivist) accounts, the brain is viewed as a symbol manipulator. In PDP (or connectionist) accounts, the brain is viewed as a complex weave of multilayered networks. The units of a network (which may be compared to the brain's neurons) are simple processors, and the connections between them, of which there are massively many, have different strengths. Information-processing is parallel, i.e. much is carried on simultaneously; and it is distributed, i.e. any individual connection participates in the storage of many different items of information. There is controversy about the exact significance of this new approach, and about its repercussions for debates in the philosophy of mind.

En beskrivning av neurala nätverk, som konnektionism också brukar kallas, i *Stanford Encyclopedia of Philosophy* (2007) understryker också att deras styrka har att göra med att de är tillräckligt flexibla för att kunna efterlikna förhållanden i den verkliga världen.

Neural networks exhibit robust flexibility in the face of the challenges posed by the real world. Noisy input or destruction of units causes graceful degradation of function. The net's response is still appropriate, though somewhat less accurate. In contrast, noise and loss of circuitry in classical computers typically result in catastrophic failure. Neural networks are also particularly well adapted for problems that require the resolution of many conflicting constraints in parallel. There is ample evidence from research in artificial intelligence that cognitive tasks such as object recognition, planning, and even coordinated motion present problems of this kind. Although classical systems are capable of multiple constraint satisfaction, connectionists argue that neural network models provide much more natural mechanisms for dealing with such problems.

Artikeln *Artificial Social Intelligence* (Bainbridge, 1994) i *Annual Review of Sociology* är ett av de tidigaste och bästa exemplen på en sociologi som försöker diskutera hur AI-teknik kan hjälpa till att utveckla sociologins teoribildning. Artikeln går igenom en mängd olika AI-strategier och visar på AI-teknik som använts för att modellera sociala processer.

Författarna börjar med att redovisa att den tidigaste AI-forskningen försökte modellera symboliska processer på en så hög nivå att de kunde motsvara människornas komplexa tankeprocesser. De konnektionistiska AI-strategierna har däremot försökt modellera de basala funktionerna i de biologiska nervernas nätverk i förhoppning att på så sätt arbeta sig fram till det mänskliga medvetandets nivåer.

Författarna framhåller att även om de neurala modellerna har en rot i ett [cybernetiskt] tänkande från 50-talet, så dominerade de symboliska modellerna i cirka 30 år fram till 1986. Sedan dess har de neurala modellerna vuxit i popularitet och vetenskaplig användning (Bainbridge,1994:409).

En viktig del i de symboliska strategierna är att kunskapen skall kunna representeras som ett nätverk av begrepp. Expertsystem är en del av den symboliska AI:n. Ett expertsystem är i formellt hänseende en symbolisk processor som är designad för att lösa specifika problem på samma sätt som en mänsklig expert inom området gör (Bainbridge,1994:413). Artikeln nämner att expertsystem bland annat kan underlätta kodandet av texter. Det kan förefalla underligt att de talar om nätverk av begrepp eftersom det kan likna tanken med neurala nätverk. Skillnaden är att begreppen här är fasta symboler som inte förändras av aktiviteten i nätverken, medan de symboler som också kan finnas i neurala nätverk hela tiden får nya innebörder. Vi kommer senare att se att detta är en fråga som återkommer då vi diskuterar vad som är grunden i de sociologiska teoriernas uppfattning om sociala processer. Handlar det om en funktionalitet som skall bygga på att moderna fungerar genom att deras relativistiska karaktär skapar möjligheter för dem att skapa länkar till nya relationer beroende på vilken situation de befinner sig i? Eller skall vi istället se på moderna som uttryck för fasta symboler som antingen uttrycker att begreppen avspeglar naturlagar eller avspeglar lagbundna sociala relationer?

I neurala nätverk lagras informationen i relationer mellan nervliknande strukturer. Detta innebär att den lagras på ett distribuerat sätt. Därför sker lärandet i neurala nätverk genom träning, ungefär så som människorna lär. Relationernas styrka bestäms genom en vikt som förändras när nätverken lär sig att svara på data som kommer utifrån. Det är ett förhållande som antyder att kontexterna är styrande vare sig det gäller människor eller maskiner. Nätet tränas genom att man presenterar det med en serie fall som både innehåller input och output (Bainbridge,1994:416).

En viktig del av den programmeringsteknik som skapar neurala nätverk är genetiska algoritmer. Det handlar om matematiska formler som är skrivna så att nätverken hela tiden kan lära sig att fungera bättre. Det skapar en likhet med mänskliga samhällen, som har mer information och som är bättre på att processa data än någon av de individer som ingår i samhället. Författarna menar att den genetiska algoritmen, trots den ickesociologiska grunden, understryker att verklig intelligens till sin natur är social (Bainbridge,1994:417).

Kathleen Carley som är en av författarna till artikeln ovan skriver i en annan artikel *Artificial Intelligence within Sociology* (Carley,1996) att neurala nätverk är ett

bra sätt för att förstå social struktur. Dessutom hävdar Carley att social intelligens behövs för att individens kognitiva kapacitet är begränsad. Den sociala intelligensen är därför nödvändig för att individen skall kunna skaffa sig olika typer av kunskaper (Carley,1996:13). Carley hänvisar till artikeln *The Nature of the Social Agent* (Carley&Newell,1994) som hon själv skrivit tillsammans med en av AI-forskningens pionjärer, Allen Newell. De hävdar där att en av grunderna för att tala om agentens behov av socialitet är att agenten bara har en begränsad kognitiv kapacitet. Individerna är så inbäddade i en sociokulturell och historisk miljö att de bara har begränsad kunskap om olika aspekter på denna miljö. Om de skall kunna nå kunskap som döljer sig bakom nya perspektiv måste det finnas en social funktionalitet som tillåter dem att övervinna sitt eget medvetandes begränsningar.

Carley håller fast vid de symboliska AI-strategierna, i avsikt att skapa artificiellt intelligenta aktörer som kan ersätta människan, trots att hon visar intresse för konnektionistisk programmering. DAI innebär däremot att de konnektionistiska programmen inte behöver jämföras med egenskaper hos människan. DAI skiftar istället fokus mot kontexterna och bygger på ett antagande om att intelligens, både hos mänskliga aktörer och i AI-system, handlar om distribuerade egenskaper.

Vi bör framhålla att det centrala i denna diskussion inte handlar om hur konnektionistiska program fungerar. Det centrala är istället hur social kommunikation fungerar och att utifrån den analysen bedöma rimligheten av huruvida de konnektionistiska programmen i en DAI-struktur kan skapa stark AI. Den kritiska frågan handlar då om huruvida det är viktigt att anta att den funktionalitet som finns i social kommunikation uppstår genom det som Luhmann kallar systemdifferens, eller om det krävs en komplex funktionalitet som endast kan skapas då människor tänker. Om vi översätter dessa olika uppfattningar om hur sociala processer fungerar, till debatten om huruvida kunskap kan skapas med dataprogram, så innebär uppfattningen att komplexitet skapas med människors tänkande, att dessa processer inte alls kan modelleras med AI-program. Om frågan däremot handlar om den kommunikation som uppkommer genom systemdifferens eller genom kausala sociala processer, så blir frågan snarast om processerna behöver modelleras med hjälp av konnektionistiska program eller om det räcker med en vanlig symbolisk programmeringsteknik som bygger på procedurer där ett givet värde leder till ett annat – så kallade then-if procedurer.

1.6 Hur syftet skall utredas

Ett problem med att utreda syftet är att AI-frågan ger så många anledningar att se på gamla frågeställningar på nya sätt. Det är speciellt ett problem för en sociologisk

analys av AI, eftersom AI-frågan i sig ger anledning att se på den sociologiska analysen från ett nytt perspektiv. Det är därför nödvändigt att vi först börjar med att utveckla varför AI är ett intressant sociologiskt studieobjekt. Därefter skall vi gå in på den debatt om AI som redan har förts. En debatt som både har förts av filosofer och sociologer. Vi börjar med den filosofiska debatten i ett kapitel för att sedan koncentrera oss på den sociologiska debatten i nästa kapitel. Därefter kommer vi att presentera Luhmanns sociologiska systemteori utförligare och visa på hur den kan användas för att diskutera AI. Efter det skall vi se på andra samhällsvetenskapliga skolor som diskuterar artefakternas roll i kunskapsprocesserna. Vi skall sedan pröva dessa synpunkter genom att ställa den mot den kritik av AI som har kommit från sociologer som representerar handlingsorienterad sociologi respektive holistisk kunskaps sociologi.

Efter den diskussionen skall vi se hur vi kan använda idéer om så kallad situerad kunskap för att utveckla och pröva de ståndpunkter om förhållandet mellan AI och sociologi som vi har kommit fram till med hjälp av det systemteoretiska perspektivet. Vi kommer att argumentera utifrån en uppfattning om att den naturliga sociologiska positionen bör vara att kunskapen är situerad.

Diskussionen kommer därför att föras efter nedanstående redovisning av innehållet i de kommande kapitlen:

1.6.2 Diskussion av studieobjektet

Ett stort problem med att anlägga ett sociologiskt perspektiv på frågan om stark AI är att sociologin faktiskt är mycket otydlig när det gäller ett så centralt begrepp som social. Vi kommer därför att relativt ingående redovisa varför analysen behöver göras med hjälp av Luhmanns sociologiska teori och varför det skulle vara mindre lyckat att använda sig av andra sociologiska teorier som också har använts för att diskutera redskaps och andra artefakters påverkan på kunskapsprocesser.

1.6.3 Debatten om AI

Vi skall se på några klassiska inlägg i diskussionen kring AI. Det handlar om annan litteratur än den litteratur som uttalat diskuterar AIs förhållande till sociologisk teori. Denna litteratur har dessutom karaktären av klassiker inom området. Denna klassiska litteratur handlar delvis om en filosofisk diskussion om datorers kapacitet att förstå, men också om argument för och emot AIs förmåga att hantera sådan kunskap som inte är formaliserad. Diskussionen kommer att fokusera på vad det är för typ av funktionalitet som är inblandad när ny kunskap skapas.

1.6.4 Det sociala i AI-diskussionen

I detta kapitel skall vi presentera vilka angreppssätt som sociologer använder för att knyta diskussionen om AI till sociologisk teori. Avsikten är här att presentera diskussionen så att vi ser vilka angreppssätt som används. Vi kommer att presentera både sådan sociologi som är positiv till AI och sådan som är negativ.

1.6.5 Luhmannsk socialitet och DAI

I detta kapitel skall vi utveckla diskussionen om hur den sociologiska systemteorin kan ge ett bättre svar på hur AI kan vara en del av den sociologiska teorin. Det viktigaste är att visa vilka möjligheter det finns att utifrån sociologisk systemteori tala om att sociala kunskapsprocesser inte behöver vara direkt kopplade till mänsklig handling. Därför skall vi i detta kapitel ta upp vilken betydelse systemteorin har för synen på människans kapacitet för intelligens, och på vilket sätt denna systemteori ser att social kommunikation ger de villkor det begränsade mänskliga medvetandet behöver för att bete sig intelligent.

1.6.6 Sociala processer som artefaktisk aktivitet

I detta kapitel skall vi närmare utreda frågan om hur redskap fungerar i kunskapsprocesserna samt diskutera på vilket sätt de kan knytas till en social funktionalitet i kunskapsprocesserna.

Vi skall börja med att ta upp en pedagogisk diskussion om hur vi använder artefakter i lärandet. Det handlar om den aktivitetsteori som inspirerats av den sovjetiske pedagogen Lev Seménovic Vygotsky (1896-1934)² och om hur den ser på redskapsanvändandet i kunskapsprocesserna. Motivet för att ta upp den diskussionen är att den fokuserar på hur sociala processer skapar redskapen. Dessutom har aktivitetsteorin utvecklat diskussionen om hur medvetandet använder redskapen för att expandera kunskapen bortom sådana gränser som människans begränsade medvetande inte medger. Avsikten är framför allt att hitta en början på den syntes mellan Luhmanns kommunikationsperspektiv och det perspektiv på redskapen som förlängning av de mänskliga kapaciteterna, som vi behöver för att ta ställning för eller emot stark AI. En central fråga för möjligheterna att placera intelligensen utanför människan är i vilken mån människan måste internalisera redskapen innan hon kan använda dem.

² Svensk stavning: Vygotskij, min anm.

1.6.7 Social kommunikation som personlig interaktion

I detta kapitel skall vi diskutera de argument mot AI som framläggs av företrädare för symbolisk interaktionism eftersom den teorin kan användas för att kritisera den luhmannska synen på kommunikation eller användas som ett stöd för att strukturella egenskaper verkligen betyder mycket för möjligheterna av att kunna se något nytt.

1.6.8 Det sociala som holism – Systemteorin och det Starka Programmet

Här skall vi ta upp den holistiska sociologins argumentation för att sociala processer inte behöver vara beroende av direkt interaktion mellan människor. Efter att ha gått igenom kritiken mot den tydligaste kunskapssociologiska argumenteringen för holistiska sociala kunskapsprocesser, David Bloor's Starka Program, skall vi diskutera hur vi med hjälp av Luhmanns uppfattning om sociala processers karaktär kan utveckla detta program. Avsikten är att visa på att sociala kunskapsprocesser inte bara behöver handla om hur kunskapen blir styrd av kunskapsintressen. Avsikten är istället att visa på att social funktionalitet också handlar om att utveckla kunskapen.

1.6.9 Ett försök till syntes – det artificiella som funktionalitet i komplexa sociala processer

I detta kapitel skall vi förankra tanken på att funktionaliteten i social kommunikation och kunskapsexpandering kommer från kontexterna. Det handlar om att försöka klargöra den syntes som redan har antytts av tidigare diskussion. För att kunna göra det behöver vi föra in några nya element i diskussionen.

Först skall vi dock återknyta till en sociologisk och socialantropologisk diskussion om situationernas betydelse för de handlingar individerna kan utföra. Därefter skall vi också utreda vad embodiment som en egenskap i kontexterna handlar om – det vi kallar environmentalism.

För att kunna slutföra syntesen skall vi återvända till frågan om vilka krav på kapacitet för komplex funktionalitet och reflexivitet vi behöver ställa på stark AI, och vilka möjligheter det finns att AI-systemen kan motsvara de kraven. Vi gör det bl.a. genom att ta upp en diskussion mellan Luhmann och Habermas om grunden för kommunikativa processer. En central fråga är om språket måste bygga på överenskommelser mellan individer om en gemensam innebörd, eller om språket kan fungera som artefakter som uttrycker tillfälliga positioner i systemens kommunikation.

1.6.10 Uppsummering av argumentationen

I detta avslutande kapitel skall vi göra en uppsummering av argumenten i form av ett index i 10 punkter.

1.7 Vad som inte skall göras

Diskussionen om stark AI kräver ett annorlunda perspektiv på vad som är sociala aspekter på teknikens påverkan på kunskapsprocesserna. Som vi redan har påvisat finns det mycket kunskapssociologisk forskning som betonar att tekniken innehåller intressen eller förmedlar intressen som styr kunskapsprocesserna. Det är av yttersta vikt att inse att diskussionen om stark AI inte handlar om dessa kunskapsintressen. Det som är intressant med stark AI är att den tekniken, om den fungerar, faktiskt innebär att människornas intressen blir neutraliserade av tekniken i sig. Det är en vanlig uppfattning att tekniken i sig aldrig är neutral. Det är en uppfattning som hör mer samman med den svaga AI-tekniken. Den tekniken blir alltid bunden av att vara ett redskap för olika människor. Och eftersom den är skapad för att människan skall kunna utveckla en viss kapacitet som har att göra med en viss syn på en människas kapacitet, är den naturligtvis kopplad till en värdering av vad människan bör kunna skapa. Stark AI kräver dock inte en människa för att kunna fungera. Den distribuerade aspekt på stark AI som vi diskuterar i denna avhandling innebär visserligen att dessa AI-system blir beroende av det som finns i kontexterna. På så sätt kan man hävda att den inte blir neutral. Vi bör dock framhålla att det inte är AI-tekniken i sig som i så fall är styrd av intressen. Det är i stället den mänskliga kulturen i form av texter och andra artefakter som är intressestyrd. Vi bör också framhålla att AI-tekniken medger att vi kan nå en mycket större kulturell kontext än den som bara är möjlig att nå för den människa som tack vare sitt eget medvetandes begränsning är hänvisad till det som finns i den kontext som hon själv tillhör. Människan kan under de senare omständigheterna, inte balansera upp de intressen som skapar den kontext hon befinner sig inom med de kontexter som har skapats av andra intressen.

Det Starka Programmet inom kunskapssociologin, som formulerats av David Bloor, har oftast använts för att visa på vilka kausala processer som bidragit till att en viss kunskap kunnat skapas. Att vi i denna avhandling tar upp detta program kan tolkas som att diskussionen om stark AI enbart skall föras som en diskussion om hur sociala processer skapar villkor för kunskap. En sådan diskussion brukar dock leda till att vi bara ser faror med AI-teknik. Diskussionen om stark AI bör istället handla om hur sociala processer påverkar själva den funktionalitet som skall jämföras med den funktionalitet som finns då människorna med sina medvetanden

skapar kunskap. Det Starka Programmet försöker å sin sida diskutera hur sociala faktorer i kausala processer skapar de villkor som gör att medvetandena blir benägna att komma fram till en viss kunskapsposition. Luhmann är mycket noga med att skilja funktionalitet från kausalitet. Vi ser här essensen i en sådan skillnad. Funktionalitet är något som kan leda till att människan ser nya möjligheter därför att processerna inte är determinerande. Kausalitet styrs däremot av lagar. De processerna kan bara leda i en riktning. Kanske beror tekniksociologins aversion mot AI på att den tror att alla processer som innehåller tekniska redskap måste vara kausalt determinerade!

Fokuseringen på funktionalitet ger ett perspektiv på diskussionen om AI, som anger att vi skall jämföra AI-tekniken med verkliga människor och inte med en idealistisk uppfattning av vad människan gör. Den diskussionen bör utgå från att människor verkligen kan ta sig ur beroendet av intressen om de verkar under ideala förhållanden eller verkligen anstränger sig för att inte påverkas av intressen. I det perspektivet bör vi framhålla att frågan om stark AI handlar om huruvida programmen kan skapa kunskap på samma sätt som människor när de inte styrs av olika intressen. Frågan om stark AI måste då uttryckligen handla om huruvida tekniken kan fungera oberoende av intressen. Vi bör också säga att om tekniken kan fungera på det sättet, så finns det anledning att påstå att AI-systemen faktiskt har förutsättningar att fungera mer fristående från intressen än människor. Det som gör att människor ändå blir lockade av att låta sig styras av andra människor är naturligtvis att de har känslor som styr dem åt olika håll. Det är viktigt att vi skiljer mellan intressestyrning som handlar om att de praktiska förutsättningarna är sådana att människorna inte kan tänka sig fram till något annat än vad situationerna medger. Och sådan intressestyrning som handlar om att människornas tidigare erfarenheter lockar dem att tänka i samma banor som förut. Stark AI blir naturligtvis påverkad av den första typen av intressestyrning eftersom den kulturella strukturen inom ett socialt system är färgat av den systemreferens som finns inom detta system.

Det bör vara i enlighet med känd erfarenhet att människor kör fast i gamla spår fastän de inte är tvingade. Det kan ha att göra med att människors medvetanden inte är "programmerade" för att alltid försöka tänka i nya banor. Vi kommer senare att se att en fördel med AI-system är att de faktiskt kan kommunicera över ett område som är enormt mycket större än vad människor med deras begränsade medvetanden faktiskt kan. Det kan därför vara lättare att programmera datorer att hitta nya spår. Vi kommer att se att det är en aspekt på stark AI som faktiskt talar för att AI-systemen ur vissa aspekter kan fungera "starkare" än vad människor kan just därför att de har lättare att nå utöver systemets gränser än vad människor kan. Fastän ett socialt system tenderar att styra även AI-redskapen har dessa redskap större möjligheter än människor att se över gränserna.

I sammanhanget kan vi nämna att till exempel Mertons kunskapssociologi brukar kritiseras för att den bara intresserar sig för vetenskapssamhällets formella processer. Vi behöver inte ta ställning till om detta är en riktig kritik mot Merton. Vi kan hävda att kunskapssociologin som disciplin definitivt bör intressera sig för intressestyrning av kunskapsprocesserna. Vi bör framhålla att diskussionen om stark AI måste fokusera på funktionaliteten i kunskapsprocesserna om den skall kunna ta ställning till om AIs kapacitet kan jämföras med mänsklig intelligens. På så sätt måste diskussionen ligga närmare den analys av formella processer som Merton sysslade med än den som letar efter dolda intressen. Däremot bör den sociologi som letar efter dolda intressen vara intresserad av en diskussion som kan visa på att det finns sätt att göra de formella processerna mindre styrda av intressen. Om vi kan visa på att AI-program kan fungera fristående från människor bör vi också ha presenterat en möjlig strategi för att göra kunskapsprocesserna mindre beroende av mänskliga intressen.

Luhmanns systemteori brukar stämpas som konservativ. Men det finns väldigt få skäl till att benämna den luhmannska teorin som konservativ. Teorin har hämtat inspiration från många andra teorier av vilka ett flertal kan sägas vara konservativa. Men det är påfallande att teorin, som den ser ut idag, till sina viktigaste beståndsdelar har hämtat influenser från tänkare som definitivt inte kan sägas tillhöra någon konservativ skola. Där ingår till exempel *den andra ordningens cybernetik* såsom den formulerades framför allt av Gregory Bateson. Den bygger dessutom på en uttalat ontologisk relativism som på intet sätt rymms inom en konservativ teoritradition.

Diskussionen om stark AI borde enbart föras utifrån dess egna premisser. Men eftersom den sociologiska kritiken av AI, i hög grad har dominerat den sociologiska diskussionen om AI, måste vi diskutera den kritik mot AI som kommer från de bägge sociologierna; den MI-orienterade sociologin och den kunskapssociologi som fokuserar på intressen. Dessutom måste vi ta upp den generella kritiken av AI. Vi måste dock vara noga med att vi gör det därför att vi bör förhålla oss till kritiken och inte för att diskutera de alternativa program som framförs. När vi diskuterar dessa andra program kan vi därför inte redovisa allt som de teorierna innehåller. Vi kan bara redovisa det innehåll i dessa teorier som leder till en kritik mot AI. Däremot bör det vara till fördel för diskussionens klarhet om vi också kan visa på hur dessa kritiska sociologier kan utveckla sina positioner om de tar till sig vissa av de aspekter på funktionella sociala processer som AI-diskussionen visar på.

Vi bör nämna att en funktionalistisk analys av AI-teknikens kapacitet att skapa kunskap på egen hand blir mycket mera komplicerad än en diskussion av vilka intressen AI-tekniken eventuellt kan innehålla. Det handlar i praktiken om två helt olika discipliner. Den funktionella analysen av AIs kapacitet kräver att vi diskuterar frågor som har att göra med medvetandets kapacitet, med innehållet i kunskapsprocesser, och med frågan om kunskapen ontologiska karaktär. Frågan

om AIs eventuella intressestyrning kräver på det hela taget bara att vi empiriskt studerar vilka intressen som kan finnas inblandade.

1.8 Metodologiska överväganden

Analysen bygger på att teoretiskt diskutera de sociala processer som föregår skapandet av ny kunskap. Framför allt genom att diskutera den sociala kommunikation som är inblandad i dessa processer. Vi skall diskutera i vilken mån det finns möjlighet att dessa processer kan återskapas med den modell för stark AI som bygger på konnektionistisk programmering och DAI struktur. Det blir då viktigt att diskutera hur kunskap alltid har skapats och vilken roll redskapen har haft, samt att förankra denna diskussion i den sociologiska teorin. Diskussionen kommer ofta att föras som ett abstrakt resonemang. Att så är fallet beror på att endast det abstrakta resonemanget erbjuder ett språk som inte binder analysen av hur det verkligen är till uppfattningar om hur vi tror att förhållandena ser ut. Det finns anledning att hävda att vi har ideologiserat våra fördomar om hur världen ser ut och befäst dem med ett vardagligt språkbruk. Luhmann kallar detta för att systemreferensen är given. Sociologins systemreferens gentemot AI tycks ofta leda till att AI-frågan diskuteras utifrån ett fördomsfullt perspektiv.

Det finns stora svårigheter att klart redovisa olika uppfattningar om AI därför att olika ontologiska utgångspunkter är korsvis inblandade i debattörernas ståndpunkter. Som sagts finns det en huvudskillnad mellan MI och holism i synen på AI. Den skillnaden kompliceras genom att både den luhmannska holismen och många MI-orienterade uppfattningar om kunskapsproduktion har gemensamt att de utgår från relativism. Dessutom har den kognitiva vetenskapsteori, som är tilltalad av stark symbolisk AI, en uppfattning om att kunskap skapas i kausala processer. Den synen har den gemensamt med det Starka Programmet. Vidare finns det väldigt olika utgångspunkter avseende synen på vad både kunskap och intelligens innebär. I synen på vad kunskap är tycks den luhmannska sociologin och den vygotsky-orienterade pedagogiken ligga varandra nära, medan det Starka Programmet och olika MI-orienterade utgångspunkter är allierade. I synen på vad det sociala är finns en släktskap mellan den luhmannska sociologin och den pedagogik som pekar på betydelsen av sociala processer kring begreppsutveckling o.dyl., i och med att bägge visar på att artefakter skapas genom andra processer än sådana som är knutna till omedelbara mänskliga handlingar. Dessutom förekommer det väldigt olika uppfattningar om vad funktionalitet och kausalitet handlar om. Och slutligen bör vi framhålla att det finns väldigt olika uppfattningar om hur artefakter fungerar i kunskapsprocesser. Hos både Luhmann och hos teorier som bygger på s.k. situationism härleds artefakternas betydelse till att de är

kontextfaktorer som påverkar genom att relationer i kontexterna förändras genom olika kombinationer. Hos de flesta konstruktivister är däremot artefakter något som handlar om en förlängning av mänsklig handling men inte för alla. Den franske vetenskapssociologen Bruno Latour menar snarast att artefakter är något som är helt fritt från det mänskliga (Latour,1995). På något sätt framstår det som att han tar ett steg längre än bägge varianterna av stark AI, eftersom artefakter samtidigt är något mer än resultat av naturlagar och något som handlar om mer än kulturellt skapade kontexter. Latour har dock på senare tid argumenterat för det han kallar *sociology of associations*. Då närmar han sig den syn på artefakter som vi har tillskrivit både situationismen och systemteorin (Latour,2005).

Vi har redan nämnt Johan Asplunds diskussion om AI. Asplund har i *Avhandlingens språkdräkt* diskuterat behovet av att skriva om en text flera gånger för att få den bra (Asplund,2002a). Asplunds texter är mycket tankeväckande och perspektivrika. Ofta ger han många möjligheter till att tolka vad han skriver. Kontentan av resonemanget i denna bok förefaller ändå vara att han hävdar att man redan när man börjar skriva skall ha klart för sig vad man vill med en text. Vet man inte det kan man heller inte skriva klart och begripligt. Den erfarenhet jag har av att skriva om stark AI som kräver att man ställer om sitt eget tänkande, är snarast att det etablerade språket frestar en att stanna kvar vid traditionellt tänkande. Även om man vet vad man vill skriva stöter man på svårigheter att uttrycka det med en stilistiskt god text eftersom stilistisk värdering ofta hänger samman med det förväntade. Med Luhmann kan vi säga att stilistik också värderas utifrån en given systemreferens.

Min erfarenhet leder mig till en rekommendation om förhållandet mellan tanke och språk som handlar om att vi bör använda det redskap som språket erbjuder till att tänja på begreppen så mycket att de leder tankeverksamheten i nya banor. Det är ett sätt att använda våra kulturella artefakter som ett redskap för en kommunikativ funktionalitet. Först efter att vi har använt språket till att rucka på relationer så att våra uppfattningar förändras, finns det en möjlighet att medvetet använda språket till att kommunicera nya möjligheter. Utan den processen går det inte att upptäcka nya distinktioner.

När man kommit så långt inser man att det är en svårighet att få de som har vant sig vid ett gängse språkbruk att förstå att begreppen i den nya kontexten betyder något annat. Detta kommunikationsproblem blir inte mindre av att vi ibland i denna text måste referera till ett alldeles egenartat sätt hos Luhmann att uppfatta fenomen och problematik.

De formuleringsproblem som vidhänger denna text kan vi sätta i relation till några råd som framhålls i en av de mest rekommenderade böckerna om sociologisk framställning, Howard Beckers bok *Writing for social scientists*. Becker talar mycket om att vi bör undvika överflödiga uttryck. Men att en del av de långa och överflödiga uttrycken inte kan tas bort eftersom de inte har någon underlydande mening som kan ersättas (Becker,1986:7). Nu menar nog Becker att sådant visar

att hela resonemanget i så fall är irrelevant. Han menar att sådana överflödiga uttryck uppkommer genom meningslösa markeringar. Han ger ett par huvudförklaringar till varför dessa uttryck uppkommer. Ett problem har att göra med aktörskap (agency): vem gjorde det som dina påståenden menar att har gjorts (Becker,1986:8). Det andra har att göra med att sociologer begår ett liknande teoretiskt fel då de säger att samhället gör si eller så, eller att kulturen tvingar människorna att göra saker. Sociologers oförmåga eller ovillighet att göra kausala påståenden leder på samma sätt till dålig stilistik (Becker,1986:9). Becker och Asplund är sociologer som jag har största respekt för men jag vill ändå hävda att jag är övertygad om att en fokusering på aktörskap och kausalitet skulle göra att vi inte kan diskutera AI-frågan på ett sätt som gör en skillnad. Skall vi kunna diskutera om dataprogram kan göra något som liknar det människor kan göra måste vi relativisera aktörskapet. Det är ett behov som i synnerhet blir viktigt när vi diskuterar distribuerad AI som handlar om något annat än att vi jämför robotars kapacitet med människors kapacitet. Dessutom tar den luhmannska sociologiska analysen avstånd från kausalitetsbegreppet. Utifrån Beckers råd är det naturligtvis denna hållning både till aktören och till kausalitet som skapar situationer som kan leda till en känsla av överflödiga resonemang. Jag menar att det inte går att analysera de komplexa problemen kring stark AI på något annat sätt än genom att arbeta med ett otydligt aktörskap och genom att avstå från kausala spekulationer. Det är det analytiska grepp som vi måste använda trots att vi får betala med svårigheter att göra diskussionen klar. För mig är det uppenbart att Asplunds diskussion av AI (Asplund,2002b) är färgad av att han så gärna vill försvara medvetandets betydelse

I en av de bättre introduktionsböckerna till Luhmanns teori framhåller författarna att teorins karaktär skapar ett problem som gör att den inte är så lätt att förstå. De lyfter fram att Luhmann själv menar att den som skriver något skall sträva efter att göra sig förståelig. De menar att Luhmann framhåller att sociologin inte är en lära som handlar om det man associerar till efter vad som kommer upp efter det första ögonkastet. Det handlar istället om sådant som hör till de betänkligheter och frågeställningar som kommer upp vid det andra ögonkastet (Nassehi&Kneer,1997:19).

I en annan bok framhålls att teorin är komplex därför att Luhmann menar att en social teori måste kunna observera den komplexitet som finns i det moderna samhället (King&Thornhill,2003). Det understryks att Luhmann framför allt försöker undvika att förklara sanningen eller essensen av det moderna samhället i en endaste teoretisk redogörelse. Författarna menar att man kan påstå att den första principen hos Luhmann är att hävda att det alltid är möjligt att se på saker på olika sätt och att möjligheterna av att samhället kan vara annorlunda alltid finns där. Man kan aldrig undvika reduktionism men man skall aldrig hävda att en sådan teori är den slutgiltiga (King&Thornhill,2003:1). Författarna Michael King och Chris Thornhill framhåller att teorin inte rymmer inom en ideologisk ram eller är en

teori som bygger på en speciell uppfattning om människans natur. På så sätt kan man inte använda teorin för att visa på klara kausala samband eller testa teorin genom empirisk undersökning (King&Thornhill,2003:1-2).

Dessa författare framhåller också att Luhmanns teori inte handlar om människor utan om system. Och systemen består inte av människor utan av kommunikationer. Det innebär att den läsare som är van vid att uppfatta sociologin som något som handlar om att observera människor blir förvirrad (King&Thornhill,2003:2).

1.9 Centrala begrepp i avhandlingen

Vi kommer att använda en mängd begrepp i denna avhandling som är ovanliga eller som används på ett annorlunda sätt.

De ovanliga begreppen är oftast sådana som används inom AI-litteraturen. Hit hör begrepp som stark och svag AI, konnektionism, symbolisk eller klassisk AI samt distribuerad artificiell intelligens och Multi Agent Systems. I AI-litteraturen används också en mängd mer allmänna begrepp som har en annorlunda användning inom denna diskurs. Hit hör själva intelligensbegreppet och begrepp som artefakt, aktör, agent, förmedling, kognition och funktionalitet.

Det finns också en mängd begrepp som hör till den luhmannska sociologiska systemteorin. Begrepp som hör hit är autopoiesis, strukturell koppling, kontingens, distinktion, och observatör. Vi bör framför allt uppmärksamma att denna systemteori också använder en del mycket centrala begrepp på ett annorlunda sätt. Framför allt handlar det om begreppen kunskap och information, kommunikation, social, förståelse, självreferens, reflexivitet och reflexion, mening, omvärld, miljö och system, funktionalitet och kausalitet samt relativism.

I texten förekommer också begrepp som är gemensamma för sociologin men som inte har någon större tillämplighet i vardagsterminologin. Dessa begrepp är inte helt entydigt definierade inom sociologin, så även sociologer kan vara oense om hur de skall användas. Begrepp som hör hit är emergens, metodologisk individualism (MI), metodologisk holism, behaviorism, biologism, talhandlingsteori, social struktur, nätverk, handling, interaktion, mikro- och makro-sociologi samt motiv och intention.

Några begrepp förekommer både inom sociologi och inom AI-diskussion. Hit hör begrepp som kunskapsexpansion, situerad kunskap, posthumanism, environmentalism, redskap, multiagent samt cybernetik.

Slutligen förekommer begrepp som hör till en filosofisk diskussion om kunskapens natur. Ett sådant begrepp är tyst kunskap (tacit knowledge). Andra sådana begrepp är relevans, semiotik, epistemologi och ontologi.

Det är inte möjligt att redan här ge uttömmande definitioner av dessa begrepp. Användningen av de flesta begrepp är sådan att de bara kan förstås i det sammanhang de hör till. Min avsikt med ovanstående redovisning är att göra läsaren uppmärksam på att det är skäl att lägga en gängse begreppsanvändning åt sidan, för att istället vara uppmärksam på att begreppen styrs av den kontext de används i.

Tyvärr finns det ytterligare ett problem med diskussionen kring AI. Det har att göra med att de som är för eller emot vissa aspekter på AI, ofta bygger sina positioner genom att lägga in en annan innebörd i begreppen än den som motståndaren använder. Jag försöker i min text göra läsaren uppmärksam när så sker. Det innebär kanske att avhandlingens språk blir lidande eftersom läsaren inte kan luta sig tillbaks mot en homogen begreppsanvändning. Detta därför att det inte alltid går att presentera argument utan att använda det språk som den som framför argumenten använder.

2. Diskussion av studieobjektet

2.1 Vad lägger vi in i begreppet social, intelligens och kunskap

Vi har redan sett att vad som kan höra till det **sociala** är centralt i diskussionen kring möjligheterna av stark AI. Problemet är att det sociala måste preciseras till något mer än ett begrepp för interaktioner mellan personer inom ett samhälle. Vi måste vara klara över om begreppet pekar på att människor interagerar med varandra, eller om det innefattar den samhällsliga struktur inom vilka människorna lever. Är det rimligt att alla kulturella artefakter som människorna använder också hör till det sociala? I Vygotskys teorier är det tämligen klart att redskap hör till den sociala sfären. I Luhmanns teori är det sociala knutet till de sociala systemen som består av något annat än att människor interagerar med varandra. Men också Luhmann har ambivalent hållning till vad social kommunikation innebär. Det är förhoppningen att den diskussion om stark AI som vi för i denna avhandling kommer att hjälpa till att utveckla vad som ingår i sociala förutsättningar för kommunikation (Luhmann,2002).³

³ Jag har letat på många ställen efter entydiga definitioner av begreppet "social". Begreppet förekommer oftast tillsammans med andra ord som social handling, social interaktion etc. Tesaurusen (förteckningen över indexeringstermer som används) i databasen Sociological abstracts som förtecknar sociologisk litteratur innehåller en uppsjö av begrepp där social ingår som första led tillsammans med ett annat begrepp. Allt från "Social Acceptance" till "Social Values". Jag har inte räknat alla variationer, men det finns nog närmare tusentalet där social ingår. I Thomas Brantes Sociologiskt lexikon finns heller ingen diskussion av enbart begreppet social. Den svenska Nationalencyklopedins nätupplaga anger: (fr., av lat. *sociālis* 'kamratlig', av *so'cius* 'kamrat'; 'deltagare'; 'bundsförvant'), samhällelig, samhälls-, samhällsbildande, relaterad till andra; i utvidgad betydelse omvårdnads-, omsorgs- (t.ex. i termer som *socialförsäkring*, *socialtjänst*); ibland också sällskaplig (<http://www.ne.se.ludwig.lub.lu.se/artikel/310814>). Inte heller Svenska Akademiens ordbok ger oss särskilt bra svar. I nätupplagan hittar jag den etymologiska betydelsen: "av lat. *socialis*, sällskaplig, till *socius*, medlem, deltagare, bundsförvant, kamrat, Om de betydelser ordet har säger ordboken: 1. om person: som (gärna) lever i. ingår i ett samhälle; 2. om djur: som gärna bildar i. lever i. ingår i ett samhälle; 3. om växt i. växtart: som (ensam i. samman med andra) ingår i slutna bestånd o. därvid anpassar sig till övriga växter; 4. om ngt sakligt, stundom äv. med personligt huvudord: som tillhör i. avser i. rör i. hör samman med i. utgår från (hela) samhället; 5. ss. adv., motsv. 4. En reflexion är att

Vi behöver också diskutera hur begreppet uppfattas av andra sociologiska teorier. Vi har redan visat på att Bruno Latours aktör-nätverksteori behandlar artefakternas betydelse i kunskapsprocesser. Kopplingen mellan artefakter och socialitet har också diskuterats av andra inom den så kallade konstruktivistiska skola som Latour hör till. Karin Knorr Cetina talar om det sociala som ett objekt i kunskapsprocesserna (Knorr Cetina, 1997). En kollega till Knorr Cetina vid sociologiska institutionen i Bielefeld, Alex Preda, hävdar att Knorr Cetinas angreppssätt är bättre än Luhmanns för att hantera tingens betydelse i sociala processer. Preda menar att Luhmann tvingas hänskjuta alla ting till faktorer i omgivningen (Preda, 1999:348). Preda hävdar att Knorr Cetinas och konstruktivisternas uppfattning om att tingen blir laddade med social kunskap genom de människor som har skapat dem, skapar ett bättre redskap för att analysera artificiella inslag i kunskapsprocesser. Detta är dock inte den enda tolkningen man kan göra av Luhmanns position. Enligt den bör det till exempel vara en skillnad på om vi talar om ting eller artefakter. De förra finns naturligtvis utanför det sociala systemet och handlar om system/omvärlds relationer, medan artefakter mycket väl kan ha en funktion inom de sociala systemen. Vi kommer att se att det finns anledning att problematisera Luhmanns uppfattning.

Intelligens är ett begrepp som inte alls är så entydigt definierat som vi är vana att tro. Begreppet betyder enligt *Svenska akademiens ordlista* förstånd, skarpsinne, begåvning; och kan användas om intelligent person men också om intelligent väsen (Svenska akademiens ordlista, 1998:34).

Diskussionen kommer att utgå ifrån att vi bör stoppa in mer i begreppet än att bara handla om en kvalitet i en människas hjärna. Luhmann har i en artikel, *Gibt es ein System der Intelligenz?*, tydligt angett att intelligens handlar om en systemegenskap. Den kan därför inte begränsas till något som finns inom en människa. Trots att Luhmann är medveten om att system kan fördumma, menar han att ett system ändå är det som anger vad som skall identifieras som intelligent. Därför ansluter han sig till den definition som Talcott Parsons gett, och som innebär att begreppet skall ses som ett symboliskt generaliserat medium. Det innebär att begreppet hör samman med det som ingår i ett system (Luhmann, 1992a:58).

I pedagogiskt orienterad AI-forskning kopplas intelligens till Vygotskys teorier om redskapsanvändande. En finsk forskare inom det område som kallas *Kunskapsmanagement*, Ilkka Tuomi, har gjort en längre genomgång av intelligensbegreppet hos Vygotsky i sin bok *Corporate knowledge*. Tuomi framhåller att Vygotsky genom att analysera samspelet mellan redskap och tänkande, kom fram till att människorna är kapabla att skapa kreativa kognitiva redskap som kan bli integrerade delar av deras tankeverksamhet (Tuomi, 1999a:114).

social mer syftar på det som ingår i begreppet samhälle än att det syftar på egenskaper hos individerna. I så fall har "social" faktiskt en stor koppling till artificiella skapelser och inte bara till människan.

Vygotsky frågade hur ett nyfött barn blir en intelligent vuxen och hur intelligensen hos vuxna transformeras av den socio-kulturella process som har skapat det begreppsliga systemet, språket, den sociala praktiken och de kognitiva artefakterna. Han menar då att intelligens först och främst har att göra med praktisk handling och inte med abstrakt tänkande. Det är något som gör att en intelligent handling inte behöver styras av mänsklig intentionalitet. Snarare bör vi se intelligens som en form av social process.

Tuomi framhåller att för Vygotsky är intelligens något som har att göra med beteendeutvecklingens **tredje stadium**. Nedärvda reaktioner och naturliga beteendesätt står för det **första stadiet**. Dessa kallar Vygotsky instinkter och de har biologiska funktioner för artens självbevarande. De behöver heller inte tränas in. Det **andra stadiet** representeras av sådana reaktioner som inte är nedärvda men som uppstår från djurets egen erfarenhet. Det består av reaktioner på de förutsättningar som erbjuds. Det handlar om villkorade reflexer som uppstår gradvis och sakta. Skillnaden mot det **tredje stadiet** är att beteendena i detta tredje stadium karaktäriseras av att de åstadkommer en abrupt förändring i beteendena. Ett intelligent beteende uppstår när en pågående handling avbryts av ett hinder. Vygotsky menar att lägre stående djur, som höns, råttor och hundar försöker lösa ett problem genom att öka den aktivitet som de tidigare gång efter gång provat. Människor och, enligt Vygotsky också schimpanser, reagerar däremot på misslyckandet och försöker komma på nya sätt att lösa problemet.

Tuomi menar att schimpanser uppvisar ett beteende som är en mellanform mellan djurs och människors beteenden. Poängen är att schimpanser, liksom människor, reagerar på ett misslyckande genom att försöka hitta sätt att lämna den givna kontexten (Tuomi,1999a:136-137). Schimpansens ändrade beteende bör ha att göra med att det var själva aktiviteten som till slut fick den att tänka i nya banor. Alternativet kan förstås vara att schimpansen visar sig ha mänsklig tankeförmåga. Men det skulle samtidigt betyda att vi inte längre kan koppla intelligens till något som är unikt mänskligt. Så länge som vi hävdar att intelligens måste kopplas till egenskaper i människans hjärna blir det svårt att förklara varför människan och inte djuren är intelligenta. Genom att koppla intelligens till kulturella faktorer gör vi faktiskt intelligensen mer mänsklig trots att det är en intelligensbeskrivning som är mer till fördel för tanken på stark AI. Om vi skriver om detta i luhmannska termer, kan vi påstå att djuren oftast är för mycket inriktade på att upprätthålla ett invariant socialt system för att de skall kunna ta tillvara de möjligheter som människorna genom social intelligens kan se i differensen mellan system och omvärld. Genom att människorna kan använda de kulturella artefakterna för att underlätta förutsättningarna för kommunikation får de större chans att genom kontexten se nya möjligheter. De blir därför inte på samma sätt som djuren bundna till den biologiska evolutionen. Djurens självreferens riktar sig bara inåt mot det egna systemet medan människorna kan använda artefakter som redskap för att reflektera över sina egna gränser.

Luhmanns uppfattning leder till att vi skall se på intelligens som en strukturell egenskap. Det innebär att en människas möjligheter att bete sig intelligent har att göra med vilken miljö hon befinner sig i, eller tidigare har befunnit sig i. En fördel med ett sådant intelligensbegrepp är att det kopplar loss analysen från biologistiska, rasmässiga och nedärvda egenskaper. Naturligtvis kan intelligens också handla om att egenskaper i människornas hjärnor gör att de kan bete sig mer intelligent i en del situationer än vad andra människor kan. Ur ett strukturellt perspektiv handlar det om att också hjärnan hör till den struktur där intelligens kan finnas.

Detta strukturella perspektiv på intelligens innebär också att den funktionalitet som krävs för intelligens inte handlar om en särskilt komplicerad funktionalitet. Den intelligenta funktionalitet som till exempel hjärnan utför, handlar då om den enklare funktionalitet den behöver för att kunna reflektera. Och inte om den mer komplicerade funktionalitet den behöver för att kunna förstå och analysera.

Termen **artificiell** innebär i sig ett problem i en diskussion som centreras kring vad människan kan som inte kan åstadkommas på artificiellt sätt. Problemet hör samman med att begreppet artificiell är knutet till vad människan kan göra eftersom det enligt *Svenska Akademiens Ordlista* betyder konstgjord eller konstlad (1998:34). Alltså som något som människan åstadkommit till skillnad från det som finns naturligt. Samtidigt innebär konstgjord att det handlar om något som är annorlunda än det mänskliga. Termen ligger också nära en annan term som förekommer ofta i denna avhandling – **artefakt**. Den termen betyder enligt SAOL konstprodukt, konstgjort föremål (1998:34). Diskussionen om stark AI handlar om huruvida detta konstgjorda föremål också kan få sådana egenskaper att det själv kan skapa produkter - i.e. kunskap.

I en ny svensk avhandling tar Tarja Susi upp hur svårt det är att definiera begreppet artefakt. Hon redovisar fyra olika definitioner (Susi,2006:28): 1. En produkt av mänsklig konst eller hantverk, eller i arkeologisk mening en produkt av förhistoriskt eller av ursprungsinnevärdare producerat hantverk, till skillnad från ett liknande föremål i naturligt format. (Concise Oxford dictionary, 1995); 2. En produkt av mänsklig tillverkning. Speciellt ett redskap eller ett ornament. (New English dictionary and Thesaurus, 1994); 3. Något som har skapats av människor för ett praktiskt ändamål. Speciellt ett föremål som hör samman med en särskild period. (Merriam-Webster OnLine); 4. Ett objekt som är tillverkat av mänskligt hantverk. (The Oxford Illustrated Dictionary, 1962/1975)⁴

Susi menar att dessa definitioner utesluter några allmänt vedertagna uppfattningar om vad en artefakt är, bland annat att språket är en artefakt. Dessutom utesluter de ovanstående definitionerna att artefakter också kan bestå av interna kognitiva artefakter såsom minnestekniker, begrepp och heuristik⁵. Hon

⁴ Referenserna här finns hos (Susi,2006)

⁵ Heuristik betyder enl. *Nationalencyklopedin*: **heuristi'k**, dels metod för att upptäcka eller bilda ny relevant kunskap, dels läran om sådana metoder (<http://www.ne.se/ludwig.lub.lu.se/artikel/202570>). I

redovisar istället en mängd andra definitioner. Hon nämner att bl.a. Steven Pinker hör till dem som ser på artefakter som en blandning av mekanik och psykologi (Susi,2006:29).

Hur vi ser på **kunskapsbegreppet** är naturligtvis centralt. Debatten kring stark AI handlar ju om huruvida AI-system faktiskt kan skapa något som vi kallar kunskap. Påståendet att endast svag AI är möjlig, handlar om att AI-systemen bara kan processa information som sedan människan måste omvandla till kunskap. I själva verket visar det sig att det är mycket svårt att hitta en entydig definition av kunskap. Speciellt svårt är det att knyta kunskapsbegreppet till några egenskaper som bara människor kan åstadkomma. Den moderna systemteoretiska sociologin utgår från en mycket relativistisk uppfattning om vad som är sanning eller kunskap. Olika kunskapsstorheter beror av i vilken position de står i förhållande till andra kunskapsstorheter. Kunskap blir då mer av en strukturell egenskap.

Det förekommer ofta att man i debatten talar om DIKW, vilket står för data, information, knowledge and wisdom (Helfer,1998:44ff). Vanligen ser man data som den minst bearbetade formen, medan kunskap och visdom ses som förädlade former. Dessa försök att definiera begreppen grundas på att det finns data och information som representerar verkliga värden i världen. Och att kunskap och vishet är något som skapas av människan. Enligt denna uppfattning ses information som någon form av stabila och objektiva fakta, medan kunskap är något som bara människor kan uppfatta. Därför är det bara information som kan förmedlas medan kunskap är något som individerna själva måste ta till sig.

Det finns många skäl att ta avstånd från en sådan uppfattning. Inte minst vardagliga erfarenheter av hur kunskapsförmedling fungerar anger att vi måste se på relationen mellan kunskap och information på ett mer komplicerat sätt. Vi brukar tala om att en person överför kunskap till en annan person. Vilket borde vara omöjligt om det är så att bara individen själv kan ta till sig kunskap. Ett annat skäl handlar om att vi brukar tala om att en bok innehåller kunskap. Det bör inte vara möjligt om kunskap bara kan finnas inom en människa.

Det är också en vanlig uppfattning att tekniken i sig aldrig kan förmedla någon kunskap. Tekniken kan bara vara ett hjälpmedel för en människa som förmedlar kunskap. Därför menar man att vi skall koncentrera oss på hur informationsförmedling underlättas med ny teknik. Här bör vi observera att den traditionella användningen av begreppet information skapar en intellektuell kullerbytta. Eftersom den tror att information kan objektiveras tror den också att

Expertsystem; heuristik i form av regler för goda gissningar och plausibla slutsatser på expertnivå. Hänsyn bör därvid kunna tas till att kunskaper kan vara ofullständiga och osäkra, varvid systemet skall kunna ange graden av tilltro till gjorda slutledningar (<http://www.ne.se/ludwig.lub.lu.se/artikel/166129>). **Heuristisk metod,** sokratisk metod, undervisningsmetod där eleven stimuleras att genom av läraren ställda frågor självständigt nå fram till kunskap och insikt (<http://www.ne.se/ludwig.lub.lu.se/artikel/202571>).

individerna med deras förstånd och med en klar språkanvändning kan använda informationen till att överföra den subjektivt skapade kunskapen. Den intellektuella kullerbyttan består då av att man tror att medvetandet suveränt kan förmedla kunskap genom att använda information som redskap för förmedlingen, trots att man hävdar att kunskapen är subjektiv.

Detta innebär att en relativistisk hållning till kunskap medför att medvetandet har begränsningar i sin förmåga att uppfatta värdet i ny kunskap eller i ny information. Det finns inte några givna regler för hur människan skall tolka världen. Det innebär också att det finns mindre anledning att bygga en kritik av stark AI på argument som handlar om att vi aldrig kan ersätta medvetandets kapacitet för förståelse.

Om vi ifrågasätter de gängse definitionerna av information och kunskap kan vi också vända på de logiska relationer som vi såg att dessa gängse definitioner tar för givna. Ilkka Tuomi har till exempel hävdad att data är mer än kunskap (Tuomi,1999b). Han föreslår att vi vänder upp och ner på hierarkin mellan data, information och kunskap. Enligt honom uppkommer data först när kunskap och information redan existerar. Det har att göra med att det inte finns några isolerade enkla fakta innan någon har skapat dem genom att använda sin kunskap. Data kan bara uppstå om en meningsstruktur, eller semantik, först slås fast och används för att presentera information. Därför uppstår data endast som en lösning på ett praktiskt problem (Tuomi,1999a:237).

Redan här bör vi ta upp **Luhmanns kunskapsbegrepp**. Hos Luhmanns handlar det om en form av struktur som tillåter meningsfull kommunikation. Problemet är att han samtidigt menar att det inte finns någon icke-mening. Det gör att Luhmanns kunskapsbegrepp framför allt beskriver en struktur som är bra på att lyfta fram den komplexitet som finns i omvärlden. Kunskap blir då något som kan vara en del av funktionaliteten i de sociala systemens kommunikation. Och samtidigt något som ligger utanför människan som en slags rationell möjlighet för kommunikation. Luhmann säger själv att kunskap ger villkor för en lärande process genom att kunskapen bygger in möjligheter för lärande i existerande förväntansstrukturer. Han utvecklar påståendet genom att hävda att om möjligheter för lärande skall kunna utvecklas måste kunskapsbärande situationer under tiden förberedas. Kunskapen måste uttalat eller outtalat ta hand om sin egen möjlighet att förändras. Det betyder att den inte längre kan söka efter att skydda sina egna förväntningar eller värdet av sin egen struktur. Istället måste den ägna sig åt att noggrant fastslå de villkor som kan orsaka en förändring (Luhmann,1995:328-329). Även om vi här inte får någon klar definition av vad kunskap är enligt Luhmann, ger resonemanget oss anledning att hävda att kunskap inte behöver vara något som kräver en så komplicerad funktionalitet att inte AI-system kan skapa den.

Rudi Laermans, en holländsk kultursociolog, har tagit upp det problematiska med att Luhmann egentligen är emot kulturbegreppet. Enligt Laermans handlar

det mest om att Luhmann vill distansera sig från Parsons funktionalism eftersom mycket av Parsons konservatism knyts till behovet av att reproducera kulturella värden. Denna uppfattning hos Luhmann skapar också vissa problem med kunskapsbegreppet. Däremot framhåller Laermans att Luhmann i *Die Wissenschaft der Gesellschaft* faktiskt definierar begreppet kunskap:

Knowledge is (the) condensation of observations, Luhmann states. He stresses that not only in science but also in daily life, knowledge is produced again and again via repeated observations. As said, an observation consists of one or more distinctions or forms that are one-sidedly used in order to indicate something as *this* (and not *that*). Repeated observations by means of the same forms result in the knowledge that *something is this* (Laermans2007:72).

En vanlig uppfattning, som brukar hänföras till Wittgenstein, är att kunskap är något som inte bara finns i den mänskliga hjärnan utan också något som är inbäddat in den mänskliga kroppen. I artikeln *From embodied to socially embedded agents* (Dautenhahn et.al.,2002) argumenterar författarna för att embodiment måste vara något som är socialt inbäddat.

Dessa författare menar dock att denna definition oftast är alldeles för bunden till kroppen. Vi kommer senare att se att vi ofta kopplar AI till robotar. Det kan vara en följd av att vi är så vana med att intelligens har att göra med mänsklig tankeförmåga. Författarna till artikeln menar istället att diskussionen om att ha en kropp ofta artikuleras genom hänvisningar till dynamiken i strukturen hos fysiska system. De kommer därför själva med en definition av *embodiment* som innebär en relation mellan ett system och dess omgivning.

A system **S** is embodied in an environment **E** if perturbatory channels exist between the two. That is, **S** is embodied in **E** if for every time **t** at which both **S** and **E** exist, some subset of **E**'s possible states with respect to **S** have the capacity to perturb **S**'s state, and some subset of **S**'s possible states with respect to **E** have the capacity to perturb **E**'s state (Dautenhahn et.al.,2002:400).

I boken *Natural-Born Cyborgs* (2003) av Andy Clark är huvudpoängen att den mänskliga naturen är skapad av vårt umgänge med den teknologi och den kultur med vilka vi uttrycker oss. Clark menar att det innebär att människan är teknikberoende och att denna varelse, som i den moderna litteraturen har kallats Cyborgen (av Cybernetic Organism), inte enbart hör till den moderna datoriserade tidsåldern. På så sätt är det svårt att egentligen se var skillnaden går mellan sådana intelligenta handlingar som människorna utför i direkt kommunikation med varandra och sådana där människorna är beroende av att kommunicera via

artefakter. Bara det att vi måste kommunicera med varandra via ett språk talar för att artefakternas roll i den sociala kommunikationen har en större betydelse än vi kan tro. Om social kommunikation inte bygger på att människor kommunicerar blir det svårt att förklara funktionaliteten om det inte handlar om en kommunikation som går genom kulturella artefakter. Diskussionen om Cyborgen bör därför vara väl lämpad att klargöra hur social kommunikation fungerar. Det måste emellertid klargöras om redskapen kan fungera självständigt eller om de måste kombineras med mänsklig intentionalitet. Det är en fråga om hur den komplexitet ser ut som måste finnas i kommunikationen.

I detta sammanhang vill jag också säga att mitt eget intresse för AI-frågan hör samman med att jag är bibliotekarie och länge har arbetat på ett stort forskningsbibliotek, Stockholms Universitetsbibliotek. Ett bibliotek och framför allt det som brukar kallas det bibliografiska systemet (bland annat bokkataloger och citeringsdatabaser) bildar ett stort och komplicerat redskap som redan innan datoriseringen, databasernas och Internets existens, innehöll många tekniska och materiella komponenter. Biblioteksinstitutionen kanske är ett av de allra äldsta exemplen på att åtminstone människornas kunskapsskapande alltid har varit teknikberoende. Den har alltid levt på att vara den institution som innan datoriseringens tidsålder kunde erbjuda något av den artefaktiska funktionalitet som behövs för kunskapsutveckling.

Teorier om att kunskap finns lagrad i kontexter ger också vid handen att den kunskapen kan ha en egen funktionalitet. Annars måste alltid information bearbetas av en människa innan den blir till något vi kan kalla kunskap. Det innebär att kunskap inte bara blir till ett objekt som måste tolkas genom en aktör. Vi kan jämföra en sådan uppfattning med den som den konstruktivistiska skolan har. Karin Knorr Cetina menar i artikeln *Sociality with Objects* (Knorr Cetina, 1997) att vi behöver en teori om objektifiering. Vi behöver beskriva hur uppfattningen om att den sociala miljön alltid består av människor, förändrats mot att också uppfatta att sådana relationer också kan etableras genom objekt. Knorr Cetina föreslår att en social vetenskaplig diskurs producerar kunskapsobjekt som uppfattas som ständigt föränderliga och oförutsedda tecken som aldrig blir kompletta och därför hela tiden behöver fortsatt utveckling. Hon menar att en objektorienterad socialitet skall ses som en form av subjektivitet som produceras genom att subjektet tillskansar sig objektets vilja. Knorr Cetina framhåller att det innebär något annat än sociologins vanliga uppfattning av teknologi. Den vanliga uppfattning som handlar om att teknologin alierar människorna (Knorr Cetina, 1997:23).

Det är en viktig utgångspunkt för denna avhandling att stark AI måste kunna uttrycka något annat än maktintressen. Vi har redan varit inne på att mycket av den kunskapssociologiska litteratur som är inriktad på att diskutera hur makt påverkar vetenskapen och vetenskapens objekt, blir tämligen ointressant för att utreda frågan om stark AI. Knorr Cetinas utgångspunkt ligger däremot mycket

närmare det som här behöver diskuteras. Hon utvecklar sin uppfattning ovan genom att säga:

I have tried to supplement the above approaches by breaking open notions such as knowledge and expertise to bring to the fore the objectual relationships which I think define knowledge processes. The idea of an object that is relevant to understanding these relationships contrasts sharply with our received notions of an instrument, a commodity or an everyday thing. I maintain that the libidinal, reciprocal and in other ways binding components of experts' object ties make it plausible to construe these relationships as forms of sociality rather than simply as 'work' or 'instrumental action' (Knorr Cetina,1997:23).

Vi kan använda Knorr Cetinas argument som stöd för att objekten, eller artefakterna, kan innehålla socialitet. Frågan om stark AI aktualiserar om dessa AI-system kan skapa kunskap utan någon inblandning av människor. Ty som epistemiska objekt blir de bundna till den kraft som människor lägger ner i dem.

Också Bruno Latour talar om objektens betydelse för att skapa kunskap. Han diskuterar hur människan är beroende av de teknologiska nätverk forskningen bedrivs i. Det har tidigare framstått som att han alltid ser redskapen som direkt kopplade till människors intressen. På så sätt har han varit mest intresserad av de negativa effekterna, eller den alienerande funktion, som Knorr Cetina kritiserade. Men i *Reassembling the social* (Latour,2005), anger han att han delvis anser att den tidigare inriktningen varit felaktig. Han menar att begreppet social används för allt som har anknytning det sociala livet. Problemet börjar när man knyter det till något materiellt (Latour,2005:1).

Han kommer så fram till att vi behöver omformulera vad som menas med det sociala. Han föreslår att vi skiljer standardsociologins uppfattning om det sociala från en som kommer från en underavdelning som han väljer att kalla *critical sociology*. Standardsociologins uppfattning kallar han "sociology of the social" medan han benämner den senare varianten som "sociology of associations" (Latour,2005:8-9).

Latour exemplifierar skillnaden mellan de bägge sociologierna och deras olika syn genom att visa på att den känns igen i de olika uppfattningar som finns mellan de klassiska sociologerna Gabriel Tarde och Emile Durkheim:

Tarde always complained that Durkheim had abandoned the task of explaining society by confusing cause and effect, replacing the understanding of the social link with a political project aimed at social engineering. Against his younger challenger, he vigorously maintained that the social was not a special domain of reality, but a *principle of connections* [min kursivering]; that there was no reason to separate 'the social' from other associations like biological organisms or even atoms; that no break with philosophy, and especially metaphysics, was necessary in order to become a social science; that sociology was in effect a kind of inter-psychology (Latour,2005:13).

2.2 Kunskapsprocesser som problemlösning eller som expanderings utöver det som vi redan känner till

Diskussionen om AI har traditionellt handlat om en kapacitet som skall motsvara den kalkylerande rationalitet som människan använder då hon försöker lösa ett problem. Redskapsanvändandet aktualiserar en viktig dimension inom diskussionen om stark AI. Den dimensionen frågar huruvida skapandet av ny kunskap skall ses som problemlösning eller som en process som försöker leta efter nya möjligheter. Problemlösning hör samman med att kompetenta aktörer analyserar problem, medan den senare processen kräver att vi skapar möjligheter att se något nytt.

Diskussionen i denna avhandling ansluter sig till en uppfattning om att problemlösning inte är det perspektiv vi skall ha när vi diskuterar intelligens. Intelligens uppfattas som en förmåga att **gå ut över det som redan är känd kunskap**. Problemlösning handlar ur det perspektivet bara om en förmåga att hitta lösningar på redan kända problem. Skillnaden kan ibland vara hårfin eftersom problemlösning ofta kan handla om att hitta på nya sätt att lösa ett givet problem. Ofta handlar det också om att man tvingas omdefiniera problemet för att hitta en lösning. Det viktiga med denna omdefiniering är att den kräver att man kan gå utanför den kontext som man trodde att var relevant för problemet. Det är en situation som kräver att medvetandet kan gå utöver det som en person redan kände till. Det kallar jag i denna avhandling för en *expanderande social process* med ett begrepp som jag har lånat från aktivitetsteorin (Engeström,1987).

Problemlösning sker ofta inom områden som är formaliserade. Den typ av AI-system som kallas expertsystem är inriktade på denna typ av problemlösning. Mycket talar för att de löser sådana problem på ett effektivt sätt. Det som AI-kritikerna antagligen menar att dessa system inte kan hantera är just sådan omdefiniering som handlar om att gå utöver redan kända problem.

Argumentationen i denna avhandling bygger däremot på att människorna, liksom expertsystemen, är alldeles för bundna av det som redan är känt, för att vi skall kunna förvänta oss att någondana kan lösa sådana problem på ett effektivt sätt utan att låta processen påverkas av kontextfaktorer. Visserligen kan en människa få hjälp av andra människor som är verksamma inom andra områden. Men det handlar då om en kunskap som redan är någorlunda känd, eftersom människor har svårt att förstå andra människor som inte har samma referensramar som de själva. Den kunskap som är icke-formaliserad kan bara påverka oss i en process som så småningom leder till att olika sociala system närmar sig varandra.

Möjligen finns det personer någonstans i världen som har sådana perspektiv att de skulle kunna visa på nya vägar framåt. Men vi bör uppmärksamma att det är svårt att förstå människor som inte tillhör vårt sociala system. Vad människan i sådana lägen alltid har gjort är att ta hjälp av litteraturen. På så sätt kan en människas tankar överföras till en annan människa fastän de aldrig direkt har samtalat med varandra. Men det handlar inte om personlig interaktion utan om en kommunikation som använder sig av funktionaliteten i artefakterna. Vi bör uppmärksamma att texter, liksom andra av de artefakter som människan skapat, har en speciell egenskap. Den egenskapen handlar om att innehållet kan förändras beroende på vem som läser det, och beroende på i vilken kontext texterna framställs. I det finns en förmedlingskapacitet som visar på att artefakternas funktion handlar om något som går utöver det människorna kan skapa med sina handlingar. Det handlar då om att det artificiella utvecklar människans möjligheter till reflexion. På så sätt kan man argumentera för att det sociala inte bara är förenligt med AI, utan också att AI kan utveckla människans sociala kapacitet där mänskliga interaktioner inte räcker till.

2.3 AI som robotar med mänsklig kapacitet eller som program för holistiska sociala processer

Den kritik av AI, som utgår från medvetandets kapacitet i problemlösandet, menar också att AI-system måste ses som robotar som motsvarar mänskliga aktörer. Om vi istället tar avspark i en holistisk sociologi behöver vi ta fasta på hur den sociologin diskuterar sociala processer. Vi kommer att se att holistisk sociologi, i belysning av de frågeställningar som aktualiseras genom den distribuerade AI-designen, måste vara mycket klarare över vilken funktionalitet som kan ersätta den som härrör sig från människornas handlande. Framför allt måste den vara mycket klarare över om grunderna bygger på kausala processer eller funktionella processer. I den luhmannska sociologin tas uttryckligen avstånd från kausala sociala processer till förmån för funktionella processer. Vi skall se att den diskussionen har paralleller

med en diskussion kring AI som handlar om en jämförelse mellan den symboliska AI-tekniken och den konnektionistiska. Den kausala sociologins positioner gör att den så mycket liknar utgångspunkterna för symbolisk AI, att den har svårt att försvara varför sociala faktorer kan ersätta mänskligt resonerande. Det går lättare att försvara de sociala faktorernas roll om vi ser på dem som funktionella processer. På så sätt kan vi lyfta fram några frågor kring stark AI som är viktiga.

En berör det sätt den konnektionistiska AI:n skiljer sig från den symboliska AI:n. Den frågan är viktig att utreda eftersom den sociologiska kritiken oftast använder symbolisk AI som mall för att kritisera all stark AI. Den symboliska AI:n kritiserar oftast för att den knyter kausalitet till kognitiva faktorer. Istället hävdar den sociologiska kritiken att kausaliteten kommer från sociala faktorer. Om vi kan förklara att konnektionistisk AI inte arbetar med kausalitet utan med funktionalitet, kan vi förklara att kritiken mot symbolisk AI inte kan överföras på konnektionistisk AI.

Kunskapssociologer skapar ofta en konstlad motsättning mellan sociala och kognitiva faktorer i kunskapsprocesserna baserat på att de olika faktorerna står för olika kausala påverkanssamband. En funktionell syn på processerna ger en bättre möjlighet att förmedla mellan kognitiva och sociala faktorer i kunskapsprocesserna. Det ger i sin tur en bättre möjlighet att beskriva hur de sociala processer som är inblandade i kunskapsskapande verkligen ser ut. Detta skapar en möjlighet att bättre frilägga möjligheterna av att stark AI kan vara förenlig med sociala aspekter på skapandet av ny kunskap, vilket också kan användas för att stärka den holistiska sociologins positioner.

Att vi kopplar bort frågan från den symboliska AI:ns kausala perspektiv och anlägger ett funktionellt perspektiv, ger också en möjlighet att bättre diskutera om AI handlar om robotar som skall likna människor eller om något som mer handlar om sociala processer. AI-frågans anknytning till medvetandets kapaciteter har gjort att AI oftast ses som en fråga som handlar om robotar. Inom AI-forskningen finns det en stor och företrädesvis teknisk forskning som handlar om att skapa robotar som kan göra samma saker som människor. Sådana robotar kan vara tämligen enkla, vilket gör att de knappast behöver innehålla någon komplicerad form av intelligens, även om de kan kräva rätt så avancerad dataprogrammering för att kunna fungera på avsett sätt. På den kommersiella marknaden finns till exempel gräsklippare och dammsugare, som själva kan klara av att klippa en gräsmatta och dammsuga en lägenhet. Den fråga som uppkommer är vad en robot skall göra för att kunna kallas intelligent. Vanligen menar man att roboten är intelligent om det roboten åstadkommer inte kan skiljas från det en människa med sitt tänkande åstadkommer. Här uppkommer frågan om hur människan lyckas bete sig intelligent. Frågan brukar oftast diskuteras som en fråga som handlar om att lösa problem. En robot som klipper en gräsmatta löser givetvis ett visst problem men de flesta problem är ju mer komplexa än så. AI-system som skapats för att lösa givna problem kallas, som sagts, vanligen för expertsystem. Den första tidens utveckling

av AI-system motiverades av en vilja att skapa expertsystem som skulle hjälpa människan att lösa problem. AI blev då något som skulle göra människan mer kapabel att utföra intelligenta handlingar eller att fatta intelligenta beslut. Typiska sådana system handlade om olika former av stöd för läkare att upptäcka symptom och att föreslå behandlingsmetoder.

Ambitionen att utveckla expertsystem handlar om att man tror att människans möjligheter att bete sig intelligent är beroende av att hon kan använda olika former av redskap som förlängning av den kapacitet hon har i sin hjärna. Forskningen om robotar har ofta utvecklats genom att man försöker konstruera redskap som gör människorna intelligentare. En engelsk professor, Kevin Warwick, blev för några år sedan känd sedan han låtit inoperera ett datachips under sin hud. På så sätt kunde datorerna i den byggnad där hans laboratorium ligger, känna igen honom utan att han själv behövde identifiera sig. Bland annat öppnades kodade dörrar automatiskt samtidigt som en högtalare önskade honom välkommen. I ett senare experiment utvecklades dessa färdigheter genom att ett mer avancerat implantat inopererades. Eftersom detta implantat också var kopplat till Warwicks nervsystem kunde professorn direkt kommunicera med datorer genom att hans eget nervsystem aktiverade implantatet till att sända signaler till datorn. På samma sätt kunde en dator sända signaler till Warwicks nervsystem. Efter att Warwicks egen fru fått ett motsvarande implantat kunde makarna kommunicera med varandra genom dessa implantat (Warwick,2002).

Warwicks exempel kan synas som trivialt. Men det visar på att en av uppfattningarna om att intelligens måste höra samman med människan, handlar om att intelligens inte bara har att göra med människans hjärna. Den har också att göra med egenskaper i hela den mänskliga kroppen och den mänskliga kulturen. Vi kommer senare att visa på att den sociologiska kritiken mot AI ofta blir grumlig därför att man stöder sig på argument som hör ihop med *embodiement* (som bör vara psykologiska argument) istället för argument som hör samman med *environmentalism* (som bör vara sociologiska argument). I detta perspektiv betyder det att kunskapen också finns i den miljö som bildas av vår kultur. Vi kan nämna att i den luhmannska teorin består omgivningen inte bara av natur. En omgivning för ett socialt system kan också bestå av ett annat socialt system. Den kunskap som finns inom ett annat socialt system blir på så sätt en faktor som tillhör omgivningen.

Warwicks experiment bör naturligtvis tolkas som ett exempel på hur långt vi kan komma med svag AI. Det kan tolkas som att redskapen ger människan större möjligheter att utveckla sin intelligens, men inte som att det är redskapen som tillför intelligens. Å andra sida kan vi hävda att vi alltid måste se varifrån den initierande kraften kommer. Kommer den från människans intentioner eller från den funktionalitet som finns i redskapen. I det sammanhanget bör vi framhålla att en holistisk sociologi måste grunda de sociala processerna på något som finns utanför människornas handlingar. Den fråga som då uppstår är hur sociala

processer kan vara intelligenta utan att dessa processer grundas på människors handlingar. Det positiva svar till den möjligheten som ges i denna avhandling, grundas på att det handlar om kommunikation. Och kommunikation kräver en annan funktionalitet än den som är iblandad i människors tänkande.

Embodiement gentemot environment aktualiserar frågan om det som gör människor intelligenta är egenskaper som hör samman med människans biologiska natur, eller om dessa egenskaper hör samman med det samhälle som människorna lever i, och som innehåller olika former av redskap, i form av kulturella artefakter som ett språk, och tekniska redskap som böcker eller datorer. Frågan om stark AI kräver att vi måste ta ställning till skillnader som ofta är hårfina.

Även om AI-frågan alltid har dominerats av en jämförelse mellan AI och människan så har teorier om environmentalism inspirerat till ett annat forskningsområde kring AI. Istället för att skapa kompetenta robotar tar den sikte på att skapa effektiva program. Det är i huvudsak denna senare inriktning som vi skall diskutera i denna avhandling. Eftersom det handlar om program handlar det också om att AI skall jämföras med funktionella sociala processer i samhället, istället för med interaktioner mellan människor. Detta är en viktig distinktion som sällan uppmärksammas.

I praktiken är det inte alldeles enkelt att se skillnaden mellan dessa forskningsområden. Det hör samman med att också robotar kräver program och att programmen i sig kan ses som aktörer. Det som mest skiljer dessa båda varianter åt är att den senare variantens betoning av processer gör att den menar att det essentiella med intelligens är att kunna ta vara på den komplexitet som finns i världen. Därmed fokuserar den på externa processer.

Andy Clark visar att också robotutvecklingen idag tar sikte på att robotens kapacitet inte behöver vara komplicerad. Han redogör för de experimentella robotar som Rodney Brooks har skapat. Istället för att ta modell av människan har Brooks försökt utveckla dem efter modell från kackerlackor. Skillnaden mellan människor och kackerlackor är att de senare inte tros ha någon större kapacitet för interna bearbetningar av information, men desto större kapacitet för att registrera vad som pågår runt omkring dem. På så sätt har Brooks sluppit problemet med att bygga upp detaljerade och komplexa inre modeller av medvetandet. Istället skall dessa robotar vara sensibla nog för att kunna välja ut några kritiska aspekter i en situation. På så sätt är det mer avgörande att de kan använda den komplexitet som finns i omgivningen än att de själva kan skapa komplexitet. Brooks menar att även den mänskliga hjärnan fungerar på det sättet. Den försöker också använda världen som modell för den komplexitet hjärnan skapar (Clark, 2003:68).

Den processinriktade formen av AI-strategi liknar den holistiska sociologin, eftersom bägge discipliner försöker förstå processer utanför människan. Om sociologin är den vetenskap som förklarar hur människan fungerar i relation till olika sociala miljöer, till skillnad från psykologin som förklarar hur människans interna processer fungerar, så bör detta processperspektiv på AI ligga nära de frågor

sociologin studerar. Det är därför inte underligt att AI-frågan har kopplats starkt till frågan om hur den så kallade mikro-makro länken inom sociologin kan se ut. Alltså hur holistiska sociala processer påverkar villkoren för individerna och hur enskilda individers handlingar påverkar de sociala processerna.

2.4 Vad innebär social kommunikation ur ett sociologiskt perspektiv

När vi har kommit närmare de frågeställningar som blir aktualiserade av den sociologiska diskussionen av stark AI, bör vi diskutera hur den sociologiska teorin och diskussionen om stark AI kan utveckla varandra. Det är här som det är väsentligt att föra diskussionen utifrån den uppfattning om social kommunikation som den luhmannska systemteorin erbjuder.

Att kunskap skapas i kommunikativa processer brukar vara ett argument som sociologiska kritiker av AI tar fasta på. En av de strängaste kritikerna av stark AI är, som nämnts, den svenske sociologen och socialpsykologen Johan Asplund. Han grundar sin kritik av AI på att kunskap skapas i en kommunikation som bygger på personliga interaktioner mellan olika människor. Asplund argumenterar i boken *Genom Huvudet* (2002b) för att själva problemlösningsprocessen är en social process som bygger på en dialog mellan människor. När vi senare diskuterar Asplunds argument skall vi föreslå att Luhmanns uppfattning om kommunikation ger perspektiv på Asplunds diskussion som anger att kommunikationsperspektivet snarast leder till argument för stark AI.

En annan sociologiprofessor som också har knutit diskussionen om AI till kommunikationen kring en problemlösningsprocess som sker i en dialog mellan människor, är Randall Collins. Men där Asplund menar att människors kapaciteter att interagera med varandra handlar om en så specifikt mänsklig förmåga, att det inte finns någon möjlighet att artificiella aktörer kan klara av dessa processer, tror Collins att den sociala processen i en framtid kan utföras av socialt kompetenta artificiella aktörer.

Vi kan framhålla att diskussionen här handlar om aktörers kapacitet. För att förtydliga ett viktigt argument, som är svårt att klargöra, så bygger diskussionen om AI i denna avhandling på att förutsättningen för kunskapsutveckling inte handlar om enskilda aktörers kapacitet. Istället handlar diskussionen om att kulturella artefakter skapar den funktionalitet som krävs för att systemen kan dra nytta av den komplexitet som finns i miljön. Och att det är ett förhållande som talar för möjligheterna av stark AI. Den artefaktiska funktionaliteten leder till processer som blir mer komplicerade än de processer som bara bygger på att två eller flera

personer interagerar med varandra, eftersom de senare blir bundna till ett och samma sociala system.

King och Tornhill visar att kommunikation för Luhmann har en mycket speciell mening. Han skiljer tydligt på kommunikation och interaktion. Informella utbyten av åsikter mellan individer handlar bara om interaktion. För att interaktioner också skall handla om kommunikation måste denna kommunikation produceras av sociala system. Kommunikation består i grunden av en syntes mellan information, yttranden och förståelse. Genom kommunikation förmedlas information i en form som medger att den kan förstås. Kommunikation handlar på så sätt om något som är både verbalt och icke-verbalt. Gester och handlingar är också information under förutsättning att de blir förstådda. På så sätt blir ett samhälle något som inte handlar om meningsfulla kommunikationer utan det består endast av kommunikationer och innefattar alla kommunikationer. Däremot kan inte medvetandet kommunicera och hör därför inte till samhället. Människa och samhälle skapar förutsättningar för varandra men någon kommunikation mellan dem är inte möjlig (King&Thornhill,2003:11-12). Vi bör emellertid återigen framhålla att det finns möjligheter att tolka Luhmanns kommunikationsbegrepp på ett något annorlunda sätt och att det pågår en diskussion om hur kommunikationsbegreppet skall revideras.

2.5 Hur kan komplexa sociala processer modelleras i en sociologisk teoretisk modell?

Vi har redan flera gånger påmint om att diskussionen om stark AI handlar om i vilken mån AI-system kan innehålla den komplexa funktionalitet som vi menar att människans hjärna är kapabel till. Vi behöver därför ta upp en grundläggande diskussion om vilken komplexitet vi kan tänka oss att innefattas i ett socialt perspektiv på kunskapsskapande.

Det innebär att funktionalitet inte bara blir ett begrepp med vilket vi beskriver processer. Vi måste också se det som ett analysobjekt i sig. Det är i och för sig inget nytt. Begreppet har alltid hört till ett av sociologins analysobjekt och det är speciellt framträdande i den funktionalistiska teoritradition som den luhmannska sociologin hör till. Vi behöver utveckla den diskussionen genom att visa på hur den komplexa funktionaliteten kan grunda sig på kontextfaktorer istället för på faktorer som har att göra med egenskaper som kommer från tänkande människor eller människor som interagerar med varandra. För att göra detta behöver vi fördjupa hur den av Vygotsky inspirerade pedagogiken ser på hur redskap kan föra kunskap vidare. Den diskussionen för oss åter in på frågan om huruvida kunskap som förkroppsligad i den mänskliga kroppen, i.e. *embodiement*,

kan omformuleras som något som finns i den humanistiska kulturella kroppen, i.e. *environmentalism*. Det betyder att vi behöver föra in element i diskussionen som handlar om att kunskapens innehåll har att göra med dess plats i en kontext och i relation till annan kunskap. Sådana element hjälper oss att se hur artefakter genom positionering kan skapa funktionalitet i kunskapsprocesserna. Diskussionen i denna avhandling förs utifrån ett perspektiv som tar fasta på att den kontext som representeras av elektroniskt lagrad litteratur förmår skapa flera nya kvaliteter genom att kontexten blir enorm till omfånget.

Betonandet av kontexternas betydelse understryker relevansen av det darwinistiska utvecklingsbiologiska perspektiv som bl.a. intresserade Vygotsky. Det är ett perspektiv som också tidigt anknöts till den funktionssociologiska traditionen. Parsons tog upp en sådan diskussion och den har förts vidare av Luhmann. Vi bör notera att utvecklingsbiologin här används som en möjlighet att förklara hur människan kan förändras.

Poängen med att anknyta till den utvecklingsbiologiska diskussionen i denna text är tvåfaldig. Den ena handlar om att den AI-kritiska sociologin, både den som utgår från MI-perspektiv och den som utgår från en kausal syn på holistiska processer, blir fast vid en uppfattning om människans kapaciteter som mer tar fasta på att se henne som en biologisk varelse än som en kulturvarelse.

Sociologins oklarhet över det sociala har sannolikt att göra med att företrädare för disciplinen idag tappar bort frågeställningen om vad som skiljer den sociala människan från den människa som bara fungerar som en biologisk organism. Det är en svårighet som återkommer i många stora frågeställningar där den ontologiska frågan ställs på sin spets. Den fråga som har mest beröring med diskussionen av AI är debatten kring sociobiologin. Likheten med AI-frågan syns i att sociobiologisk teori försöker förklara människan på annat sätt än som en skapelse av en gudomlig kraft. Den tidigaste formen av sociobiologi försökte knyta människan till mycket statiska beteenden som helt förklaras med egenskaper som har att göra med hennes biologiska utveckling. En modernare sociobiologi tonar ner biologisk utveckling och hävdar istället att människan har utvecklats genom att hon själv lyckats skapa en alltmer kvalificerad kulturell miljö. En miljö som sedan har hjälpt henne att höja sig till allt högre kulturella nivåer. Till sådana nivåer som innebär att vi har börjat tala om att människan till skillnad från djuren är en intelligent varelse. Det framstår som att vissa sociologers kritik av AI är sprungen ur en uppfattning som inte riktigt inser att man då ligger närmare att förklara människans sociala beteenden med biologiska argument istället för med kulturella argument. Ty de senare borde innebära att man inser att människans socialitet alltid varit beroende av kulturella artefakter och att det gör att frågeställningarna kring stark AI inte borde vara särdeles kontroversiella. Vi bör framhålla att detta senare perspektiv inte borde vara främmande för sociologin eftersom den sedan länge också har försökt förklara socialitet genom att knyta den till en kulturell struktur. Sådana drag finns hos både Parsons och Durkheim. Det är värt att nämna att också Marx och Engels

försökte knyta den darwinistiska utvecklingstanken till faktorer som har att göra med något som inte bara handlar om biologisk utveckling. Engels skrev till exempel 1876 en text som heter *Arbetets andel i apans förvandling till människan* (Engels, 1965). Vi skall också nämna att en diskussion om funktionalitet sedan länge också har haft en betydelse i strategier för att utveckla samhället i progressiv riktning. Till exempel i den så kallade gillesocialism eller funktionsocialism, som var mycket inflytelserik i den engelska arbetarrörelsen i början på 1900-talet.⁶

Den andra poängen med den utvecklingsbiologiska diskussionen handlar om att en strategi som bygger på att komplexitet kommer från kulturella strukturer nödvändigtvis måste grunda den uppfattningen på en relativistisk ontologisk uppfattning om kunskapens natur. Det som skiljer den luhmannska funktionalismen från till exempel den som Parsons formulerade är just att den bygger på en uttalad relativism. Vi skall se att det är en diskussion som gör att kulturell utveckling i ett kort perspektiv kan föra med sig stor förändring. Därför behöver inte en teori om människans kapacitet bara hålla sig till den långa tidsrymd som brukar antas föregå en förändring av den mänskliga kroppens genetiska kapaciteter. Det innebär bland annat att den kapacitet som människan har, inte blir beroende av att människans hjärna måste genomgå en lång biologisk utvecklingsprocess innan den kulturella kontexten kan påverka människans intelligens. Genom att den kulturella utvecklingen leder till att redskap som AI utvecklas, kan människan lyfta sig över sina biologiska begränsningar.

Sociologins funktionalitetsbegrepp grundar sig på Durkheims funktionssociologi och har utvecklats av Talcott Parsons. Det är ofta svårt att skilja mellan kausala och funktionella processer hos Parsons och Durkheim eftersom bägges teorier bygger på en realistisk ontologi. Därför kopplas funktionalitet hos dem snarare till stabilitet än komplexitet. Luhmanns funktionalism bygger däremot på en relativistisk position, samtidigt som han är mycket noga med att göra en skillnad mellan funktionella processer och kausala processer. Bägge dessa inslag gör att man kan hävda att det faktiskt är en stor skillnad mellan Luhmann och Parsons. Parsons använder funktionsbegreppet för att se hur givna strukturer fungerar. Därför benämns Parsons funktionalism *structural functionalism*. Luhmann är mindre intresserad av givna strukturers funktioner och vänder på det teoretiska perspektivet. Istället söker han efter alternativa lösningar på problem. Det grundläggande antagandet handlar om att det finns alternativa lösningar på nästan alla problem än de som redan valts. Lösningar som leder till att vi ser nya möjligheter vilka leder till att systemen hela tiden förändras. Detta leder till att Luhmann benämner sin egen funktionalism som en *functional structuralism* (International Encyclopedia of the Social & Behavioral Sciences:s.5851). Luhmann ersätter Parsons kausalfunktionalism med det som han kallar

⁶ Se till exempel G.D.H. Coles böcker, *Guild Socialism Restated* (1920a) och *Social Theory* (1920b).

ekvivalensfunktionalism. Luhmanns funktionalism handlar därför inte om kausala relationer utan om sammanhang mellan problem och lösningar på problem. Detta öppnar för alternativa lösningar på det ursprungliga problemet. På så sätt blir världens komplexitet den yttersta utgångspunkten för den funktionella analysen (Nassehi&Kneer,1997:42-44).

Det viktiga är att relativismen i Luhmanns teori är det som skapar grunden för komplexiteten i de funktionella processerna. Det är i sin tur något som just gör att vi behöver skilja mellan två typer av funktionalitet. Alltså mellan den brutala eller deterministiska funktionalitet som kausala teorier uppfattar att funktionalitet handlar om. Och något som framstår som mer finstämda funktionella processer som inte är determinerande och som fungerar just därför att samhällets differentiering skapar en ständig möjlighet till ompositionerande aktivitet. Vi kommer att se att diskussionen om stark AI handlar just om huruvida programmen måste fungera kausalt därför att programmen inte kan förändra sin kod eller om vi kan baka in en sådan finstämd funktionalitet i processerna att programmen får en chans att förändra sina koder. Endast den senare är förenlig med DAI.

Funktionalitet brukar också kopplas samman med att något skall vara användbart eller avsiktligt skapat för att tillfredställa ett visst behov. King och Tornhill tar upp en diskussion om hur Luhmanns funktionsbegrepp skiljer sig från den som finns hos antropologer som Malinowski och Radcliffe-Brown. Luhmanns funktionsbegrepp är, liksom hos Parsons, mycket knutet till en förmåga att upprätthålla vissa grundfunktioner. King och Tornhill framhåller att Luhmann menar att utvecklingen av samhällets delsystem inte har skett på ett avsiktligt eller rationellt sätt. Istället handlar det om den process där informationen utväljs och ges mening i en kommunikation. Samhällets funktionella system blir "funktionellt" i samma stund som andra kommunikativa system blir beroende av deras kommunikation. System blir funktionella bara i den mån de förmår organisera kommunikationer och dela upp dem på sätt som gör att andra kommunikativa system kan använda dem. På så sätt representerar funktionalitet inte några ideala tillstånd (King&Thornhill,2003:9-11). King och Tornhill beskriver skillnaden mellan Luhmanns och några av hans föregångare:

Although Durkheim's concept of division of labour and Talcott Parsons's action systems, Weber's theory of rationalization and Arnold Gehlen's account of technology are influences on Luhmann's ideas, Luhmann's method of functional analysis is far removed from any of these theorists. He sees, as already stated, functionally differentiated subsystems as organizing not labour divisions or social action, but meaning. Socially differentiated systems in their production of communications transform information into meaning, and without their operations meaning, and so society, could not exist. But more than that, each has its specific function in relation to the organization of meaning for society (King&Thornhill,2003:11).

I AI-diskussionen brukar anföras att AI inte kan omfatta det som inte är formaliserat. Den sociolog som kommer närmast att formulera detta som en kritik av AI är Harry Collins som (i samarbete med Martin Kusch) skrivit boken *The Shape of Actions* (Collins&Kusch,1998). Sven-Eric Liedman tar i boken *Ett oändligt äventyr* upp det som Collins och Kusch skriver.

Liedman framhåller att de där skiljer mellan *mimeomorfa* och *polymorfa* handlingar. De förra är efterliknande och kan utföras av både maskiner och människor, medan polymorfa handlingar är varierade på ett sätt som de maskinmässiga aldrig kan vara. I de polymorfa handlingarna skapas variationerna i samspelet mellan människor. De är socialt inbäddade i en mängd människors sammanlagda handlingar. Vad en människa gör är beroende av vad andra gör samtidigt. Dock betonar Collins och Kusch också att de polymorfa handlingarna samtidigt inramas och begränsas av sociala institutioner och konventioner (Liedman,2001:142).

Den fråga som uppstår är om det inte är rimligare att dessa sociala institutioner och konventioner framför allt representerar en socialitet som begränsar kunskapsprocesserna. Collins talar om att de sociala institutionerna på något sätt påverkar de polymorfa handlingarna. På så sätt anger han att det är egenskaper i institutionerna som skapar förutsättningarna för det polymorfa. Collins bör tala emot sig själv eftersom det sociala inte borde ha den förmågan om det bara avspeglade handlingar som människorna tidigare utfört på så sätt att det avspeglar sig i institutioner och konventioner. Sådant borde istället synas som formalisering. Det polymorfa borde istället vara något som får sin funktionalitet via den sociala kommunikationen.

3. Debatten om AI

Vi återkommer i detta kapitel till flera av de infallsvinklar på stark AI som vi tangerat i de första kapitlen. Det blir ibland några upprepningar av argument som vi redan bekantat oss med, men dessa argument är så centrala att de behöver framläggas utifrån flera olika aspekter för att vi skall se hela den problematik som de innehåller.

Vi skall börja med det så kallade *Turingtestet*. Efter en genomgång av grunderna för konnektionismen kommer vi till det exempel som är känt som det *kinesiska rummet* och som är framlagt av filosofen John Searle. Turingtestet och det kinesiska rummet är de två exempel som diskussionen kring stark AI oftast utgår ifrån.

3.1 Turings test och jämförelsen med människan

Turingtestet formulerades redan 1950 av den engelske matematikern Alan M. Turing i en artikel (Turing,1950). Där förklarade Turing principerna för hur en dator kunde utföra matematiska operationer. Turing var inte den som myntade begreppet artificiell intelligens men artikeln anses vara upphovet till att man började tala om AI.

Turings test bygger på att undersöka om de svar en maskin ger, går att skilja från svar en människa gett. Kan personen som ställt frågan inte skilja mellan svaren måste vi tillerkänna maskinen intelligens. Testet byggde på ett antagande om att intelligens uttrycker en form av kalkylerande rationalitet. Samt att denna kalkylerande rationalitet är så regelstyrd att det inte finns någon större anledning att vänta sig att en människas beteende skulle vara så mycket annorlunda än en maskins. Det innebär att även människans beteende enligt vissa teorier anses orsakas av kausala regler. Detta är en av de grunder som den klassiska AI:n står på. Vi skall också se att även mycket av den sociologiska teorin utgår från att människans beteenden styrs av kausala regler. Men den sociologin menar då att dessa regler handlar om något som människorna tillsammans skapar och som inte maskiner kan uppfatta.

Att intelligens uppfattas som ett beteende har också ifrågasatts. Sven-Eric Liedman framhåller att det är ifrågasatt om problemet gäller enbart ett beteende eller ett verkligt tänkande (Liedman, 2001:129ff.). De som kritiserar Turingtestet menar att ett beteende inte är samma sak som att människan verkligen har använt sin tankeförmåga. Det finns ett antal kommentarer man kan ge till detta påstående. Det mest uppenbara är hur vi skall kunna skilja ett beteende från ett tänkande. Jag uppfattar att tänkandet i funktionellt hänseende anses vara något mycket mer komplext än ett beteende. Det är väl här som uppfattningen om att hjärnan är en oöverträffad maskin för komplex funktionalitet kommer in. En annan uppfattning grundar sig på att tänkandet är något som sker internt inom medvetandet, medan ett beteende kan ske utan tankeverksamhet och på ett reflexmässigt sätt. Här uppstår en svårighet att visa vad som skiljer tänkandet från den kalkylerande rationaliteten. Den skillnad man kan se är att rationaliteten ändå utgår från regler medan tänkandet sker mer förutsättningslöst. Vi har redan varit inne på att man gärna kopplar behovet av dessa internaliserade medvetandeprocesser till att vi också behöver göra något av allt det som i vår omvärld är icke-formaliserat och därför inte kan behandlas med kalkylerande rationalitet. Det är just på denna punkt som vi behöver uppmärksamma att de som intresserar sig för hur AI kan bidra till kunskapsutveckling, menar att de tänkande medvetandeprocesserna är alldeles för begränsade för att kunna tolka det icke-formaliserade på nytt sätt. Man menar att medvetandet trots allt är bundet till den kalkylerande rationaliteten. Det är här som Luhmanns teori blir så användbar eftersom den visar att inte ens en människa kan tänka utanför de givna ramar som skapas av systemreferens. På så sätt blir "tänkandet", i processer som strävar efter att skapa ny kunskap, inte särskilt komplicerat. Den process som syftar till att utveckla systemets referensramar handlar om kommunikation och det är bara systemet självt som kan kommunicera.

Vi bör uppmärksamma att den AI-kritiska och företrädesvis konstruktivistiska sociologi som vi kommer att ta upp i denna avhandling, har ett ambivalent förhållande till rationalitetsbegreppet. Å ena sidan kritiserar den att människorna gör rationella överväganden med hjälp av sina hjärnor. Men å den andra sidan hävdar den att rationalitet styrs av regler som är skapade i och av sociala processer. Ett sätt att tolka detta är att rationalitet består av en egenskap i de sociala strukturerna och att människans medvetande inte är kapabelt att gå ut över det som reglerna tillåter. Det är ett argument som anger att människans medvetande inte kan nå det som inte är formaliserat. Och samtidigt ett stöd för Luhmanns argument att rationalitet handlar om en från människan extern egenskap. En sådan argumentation är väl förenlig med påståendet att människan måste använda redskap för att kunna nå det icke-formaliserade.

En möjlighet att tolka denna sociologiska kritik är att den hävdar att social intelligens behövs för att tillåta människan att gå utöver den kalkylerande rationalitet som ett medvetande tillåter. Det är när flera medvetanden kommunicerar med varandra som människorna kan se något nytt. Det svar som

DAI perspektivet ger, är att det blir svårt att förstå hur flera medvetanden kan fungera tillsammans om inte kommunikationen går via olika former av artefakter. Vi behöver alltid använda ett språk för att kunna kommunicera. Och oftast behöver vi kommunicera via litteratur och andra artefakter. Det går då att hävda att själva funktionaliteten kommer från artefakterna, eftersom det blir svårt att förstå att flera enskilda medvetanden, som bara är kapabla till kalkylerande rationalitet, plötsligt skulle bli kapabla att gå utöver dessa begränsningar bara därför att de träffar andra människor. Visserligen vet vi att vi ofta stimuleras av andra människor, men man bör kunna hävda att vi stimuleras av dessa människor just därför att de ur vårt eget perspektiv representerar en form av artefakt. Vi kan ju faktiskt inte direkt kommunicera med den andres medvetande och vi kan aldrig helt förstå vad en annan människa menar. På så sätt bör inte andra människor representera någon annan kvalitet än den som vi till exempel kan få genom att läsa en text.

Den dominerande sociologiska kritiken av AI bygger på att den sociala funktionaliteten hör samman med det mänskliga medvetandets egenskaper. Utifrån den luhmannska sociologin finns det större anledning att se på tänkandet som ett beteende eftersom tänkandet är styrt av de förutsättningar som observatörspositionen medger. Det som medger att människan kan gå utöver det som den kalkylerande rationaliteten tillåter, handlar då inte om att hon tänker sig fram till ny kunskap, utan om att hon befinner sig i positioner som tillåter att hon beter sig intelligent. Denna uppfattning om vad intelligens handlar om, är då förenlig med att artefakter åtminstone innehåller en del av det som vi kallar social funktionalitet. Det är svårt att bevisa att det förhåller sig på detta sätt. Den svårigheten måste värderas utifrån att det är lika svårt att verkligen bevisa att intelligens bara handlar om en intern kapacitet i människans medvetande.

Liedman framhåller också en annan fråga. Den tar fasta på om det är människan som är jämförelsepunkten eller en mer allmän rationalitet som maskinen liksom människan kan ha del i? Jag uppfattar inte att Liedman själv tar ställning till dessa frågor. Frågan bör handla om i vilken mån det finns anledning att tala om att det finns en verklig ontologisk nivå som skapar någon form av lagbundenhet som styr skapandet av ny kunskap. Existensen av en sådan lagbundenhet är något som den symboliska AI:n grundar sina positioner på. Det är en position som även den kausala holistiska sociologin bör bekänna sig till, även om lagbundenheterna kommer av att människor brukar bete sig på ett visst sätt och inte av att naturen är lagbunden.

Den senare positionen aktualiserar frågan om inte en socialitet som enbart har att göra med kapaciteter hos människorna, till slut landar i en uppfattning om att det är naturen som ger det sociala en lagbunden form. Det är på så sätt som frågeställningarna som handlar om vad det sociala är, påminner om de frågeställningar som diskuteras inom utvecklingsbiologin.

Både **klassisk AI** och **konnektionistisk AI** har använts som modeller för **stark AI**. Vi kan alltså inte säga att den ena eller den andra modellen ligger närmast stark AI. Om sociala processer är centrala för skapandet av kunskap bör det vara ett argument emot en stark klassisk AI, eftersom denna AI-form bygger på att naturen exponerar kunskap genom lagbundna processer.

Skillnaden mellan konnektionistisk och klassisk AI återkommer vi till i nästa avsnitt. Här kan vi nöja oss med att det är en kritik om att det inte är någon egentlig skillnad mellan dem, som gör det så angeläget att på djupet diskutera hur de sociala processerna fungerar. Om dessa sociala processer måste stå i direkt förbindelse med mänskliga handlingar därför att bara människan kan skapa det sociala, talar det för att det inte är någon skillnad mellan konnektionistisk AI och klassisk AI avseende deras förmåga att innehålla en social funktionalitet. Varken den ena eller den andra kan fungera därför att det som skapar förändring av moderna alltid måste skapas genom ett mänskligt medvetande.

Att ta ställning till ett sådant argument kräver att vi ger oss in i samma krångliga diskussion som den som sociologin själv för kring frågan om MI gentemot holism. Om det är så att sociala processer måste grundas på mänskliga handlingar bör det vara omöjligt att hävda en holistisk syn på samhället. Om vi däremot skall hävda holistiska argument måste vi hitta någon annan bärare av den sociala funktionaliteten än människan. Sådana argument behöver dessutom kunna visa att sociala processer handlar om annat än en naturgiven lagbundenhet. I så fall bör de argumenten likna de argument man behöver använda för att tala om att konnektionism innebär något annat än symbolisk AI. Den holistiska sociologin måste alltså visa att den funktionalitet som skapar sociala processer inte heller är beroende av att processas av det mänskliga medvetandet.

Det som har styrt försöken med klassisk AI är en uppfattning om att hjärnans komplexitet har att göra med kapaciteten att tolka den lagbundenhet som finns i världen. Den komplexa funktionalitet eller den intelligens som behövs, handlar då om en förmåga att upptäcka de lagbundenheter som finns i världen. Det är den definition av intelligens som brukar ligga bakom IQ-tester. Kritiken handlar då om att datorer inte kan ha en sådan funktionalitet att de i sig själva kan anses intelligenta. Vi har redan uppmärksammat att den sociologiska kritiken oftast utgår från att den mänskliga funktionaliteten inte handlar om att upptäcka lagbundenheter i världen utan om att skapa något nytt. Intelligens blir då något som mer handlar om kreativ kapacitet. Den handlar om att skapa något av det icke-formaliserade som finns i världen. Vilket i sin tur handlar om en förmåga att kunna observera och reflektera. Det bör därför vara fråga om en funktionalitet som handlar om andra saker än att människan internt analyserar de intryck hon får från världen.

Huruvida konnektionismen kan vara modell för en stark AI representerar den tekniska form som frågan om den artificiella intelligensens kapacitet kan ta. Vi bör hålla i minnet att det är en kapacitet som skall jämföras med den kapacitet som

människan kan ha. Det är då av yttersta vikt att vi håller i minnet att den dramaturgi som ligger bakom tanken på DAI innebär att även villkoren för människans kapacitet för intelligenta handlingar ges av egenskaper i kontexten eller i miljön.

3.2 Den konnektionistiska modellens vetenskapsteoretiska antaganden

Vi bekantade oss redan i kapitel 1 med några av grunderna för konnektionistiska modeller. I detta avsnitt skall vi se närmare på denna modell och försöka redovisa dess vetenskapsteoretiska grund. Den framställning jag har valt finns i en bok av Bechtel & Abrahamsen med titeln *Connectionism and the mind*. Innan jag redogör för hur Bechtel & Abrahamsen presenterar den konnektionistiska modellen kan vi se hur de redovisar den vetenskapsteoretiska grund som den klassiska AI:n står på.

Argumentet mot de klassiska AI-strategierna är oftast att de bygger på en rationalistisk ansats till att behandla symboler som representanter för företeelser i världen. Det är därför som den klassiska AI-strategin också kallas *Symbolisk AI*. (Begreppet *Symbolic AI* används oftast i engelska texter). Bechtel & Abrahamsen framställer vetenskapsteorins rationalistiska uppfattning som att den framhåller att det viktigaste för att nå kunskap är att resonera kring inneboende eller naturgivna idéer. Rationalisterna förnekar inte erfarenheten men de menar att den bara visar på vilka möjliga kombinationer i världen som verkligen används (Bechtel&Abrahamsen,1991:68-69).

Den symboliska AI-modellen bygger på att ändra eller lägga till värden på olika värdeenheter genom en rationell bearbetning av den information som tas in i systemet. Vi skall senare se att en av dispyterna kring de sociala faktorernas roll i kunskapsprocesserna handlar om resonerandets betydelse. I den debatten är det just rationalister som angriper företrädare för de sociala faktorernas betydelse, på den grund att de anklagar de senare för att de inte inser det rationella tänkandets betydelse. I debatten kring det rationalistiska tänkandet står konnektionismen på samma sida som företrädarna för det sociala perspektivet. I konnektionismen handlar det nämligen heller inte om att förändra värdet på själva enheterna. Istället handlar det om att ändra vikten i det samband eller de relationer som finns mellan enskilda enheter. Som vi har konstaterat är det svårt att exakt fånga upp vad sociala relationer innebär, men en strategi som vill ändra på relationer bör ligga ganska nära vad vi uppfattar ligger i ett socialt perspektiv på kunskap. Det grundläggande målet för dessa konnektionistiska nätverk är att ändra på vikterna så att det som nätverket presenterar i framtiden skall vara annorlunda. Den stora utmaningen blir då att få nätverket att göra detta själv utan att någon extern programmerare gör

detta. Bechtel & Abrahamsen skriver att det kräver att det finns lokala noder, inom den del av nätet där förändringen sker, som kontrollerar att vikterna inom nätet förändras (Bechtel&Abrahamsen,1991:47-50). Företrädarna för konnektionism menar att detta bättre motsvarar det beteende som människan använder för att hantera information. Det bör vara uppenbart att ett system som bygger på att ändra relationer, istället för att leta efter naturlagar, ligger mycket närmare en syn på hur ett socialt system fungerar. Därför kan vi tolka konnektionismen som att den omfattar att en "mänsklig" modell för informationshantering måste bygga på att människan beter sig som en social varelse.

En viktig fråga vid värderingen av de konnektionistiska modellerna är om de inte bara representerar en naiv associationism. Något som skulle innebära att kunskapsskapande bygger på ett passivt mottagande av intryck. Det skulle innebära att de bara klarar av en mekanistisk funktionalitet vilket betyder att nätverken aldrig kan förändra vikterna. Det är i sammanhanget intressant att det liknar den kritik som rationalisterna riktar mot dem som hävdar kausala sociala processers betydelse för kunskapsprocesserna. Den kritik som rationalisterna riktar mot konnektionismen sammanfattar Bechtel & Abrahamsen som att en adekvat modell av kognition måste använda sig av symboler och andra relaterade hjälpmedel för att manipulera information. Därför måste lärande innehålla ett system för hypotestestning, analogisering och manipulering av symboliska representationer (Bechtel&Abrahamsen,1991:101-102). Bechtel & Abrahamsen svarar emellertid att konnektionism visserligen är en tillämpning av associationism, men den har lärt sig av den utveckling som kognitivismen gått igenom under de senaste 20 åren. Därför innebär konnektionism mer än ren associationism fastän den härstammar från associationism (Bechtel&Abrahamsen,1991: 102-103).

Bechtel & Abrahamsen försöker förklara skillnaden mellan konnektionism och klassisk associationism med att i den senare är de elementära enheterna idéer med givna värden. Om man menar att en förutsättning för att idéer skall kunna förändras är att människan med sitt förnuft värderar dessa idéer, är det naturligt att man ser att idéerna i ett konnektionistiskt nätverk passivt tas emot av noderna. Ett konnektionistiskt nätverk bygger emellertid på att enheterna representerar något som är av mer komplext slag. Komplexiteten finns redan i det som noderna tar emot. Den komplexiteten gör att noderna kan bli påverkade på olika sätt genom att de alltid tar emot denna omvärld som relativa värden kombinerade med andra relativa värden i en kontext. Vi talar här alltså om varifrån den komplexitet kommer som ger människan/noderna mening. En tolkning av AI kritikernas positioner är att de menar att mening är något som människan tillför genom att behandla de statiska värden som världens lagbundenhet representerar. Världen är entydig medan människan skapar komplexiteten genom den kapacitet som hennes hjärna ger henne. I en sådan modell kan naturligtvis associationism inte leda till något annat än brist på mening. Resultatet blir obearbetade intryck som bara leder till torftighet.

I de konnektionistiska systemen är det världens komplexitet som skapar förutsättningarna för mening. Komplexitet har att göra med hur världen förmår exponera medvetandet för något nytt. De konnektionistiska systemens effektivitet hör samman med hur de situationer skapas som gör att vi kan observera världen på ett nytt sätt. Kan dessa situationer bara skapas av människor eller kan dessa situationer skapas på artificiellt sätt?

3.3 Searles kinesiska rum

Det mest kända argumentet mot att AI kan motsvara ett verkligt tänkande är Searles kinesiska rum. Searle formulerade argumentet i en omtalad artikel som heter *Minds, Brains and Programs* (Searle, 1980). Han framställer sitt problem på följande sätt: Kan medvetandet motsvaras av ett datorprogram? Eller mer specifikt: Om ett datorprogram simulerar eller imiterar mänskliga aktiviteter som kräver förståelseförmåga, kan vi då säga att också datorprogrammet har förmåga att förstå? Det är alltså en diskussion som ganska klart ansluter till Turingtestet och frågeställningarna kring den klassiska AI-forskningen.

Searle argumenterar också emot något som kallas systemsvaret vilket förefaller vara det som han uppfattar som det starkaste argumentet för stark AI. Detta systemsvar innehåller mycket av det som den konnektionistiska varianten bygger på. Det går ut på att förståelse inte bara är något som kräver en mänsklig hjärna. Systemsvaret menar också att förståelse är något som människan uppnår i interaktionen med den omgivning hon befinner sig i. Vi skall dock börja med Searles argument mot klassisk AI.

För att testa sin fråga om datorprogrammets förmåga att förstå gör Searle en simulering. Den handlar om att han är innesluten i ett rum och ges en massa meningar nedskrivna på kinesiska. Ett språk som han inte har någon som helst kunskap i. Inte ens så mycket att han med säkerhet kan skilja nedskriven kinesiska från japanska eller från några andra meningslösa krumelurer.

Searle ber oss även anta att han får en andra uppsättning av kinesisk text tillsammans med en regeluppsättning som talar om hur han skall kombinera den första uppsättningen texter med den andra. Dessa regler är på hans modersmål engelska. Efter detta får han ännu en uppsättning texter på kinesiska tillsammans med instruktioner som är på engelska. Genom dessa instruktioner skall han kunna korrelera delar i den tredje uppsättningen med de två första, samtidigt som dessa regler instruerar honom hur han skall svara på de kinesiska symboler som har en speciell form. Honom ovetande kallas den första uppsättningen ett *skript*. Den andra kallas en *berättelse* och den tredje *frågor*. Den uppsättning regler han får kallas ett *program*.

Searle menar att efter hand blir han mycket bra på att manipulera de kinesiska symbolerna, samtidigt som programmerarna blir bra på att skriva de program de ger honom. Därför kan någon som iakttar hans beteenden från en position utanför rummet, inte skilja hans svar från de svar en infödd kines skulle ge. Alltså skulle *Turingtestet* ange att hans beteende är ett intelligent beteende.

Poängen i exemplet, menar Searle, är att han när han manipulerar de kinesiska orden, bara producerar svar genom att manipulera otolkade formella symboler. På så sätt uppträder han bara som en dator och är bara ett uttryck för ett datorprogram. Searle menar att den processen inte kan jämföras med den där han svarar på engelska eftersom han då förstår innebörden i symbolerna (Searle, 1980:417-418).

Searle understryker att oberoende av vilka formella principer datorn programmeras med, räcker de inte för att uppnå förståelse eftersom en människa kan följa formella principer utan att förstå något.

Notice that the force of the argument is not simply that different machines can have the same input and output while operating on different formal principles - that is not the point at all. Rather, whatever purely formal principles you put into the computer, they will not be sufficient for understanding, since a human will be able to follow the formal principles without understanding anything. No reason whatever has been offered to suppose that such principles are necessary or even contributory, since no reason has been given to suppose that when I understand English I am operating with any formal program at all (Searle, 1980:418).

Han framhåller att det aldrig har framlagts några bevis för att den som förstår sitt modersmål alls har lärt sig några formella program. Searles tycks mena att frågan om intelligensens förhållande till formalisering inte är avgörande, samtidigt som han menar att det avgörande är att endast människan har den maskin, i.e. hjärnan, som kan tänka. Det är då endast hjärnan som har den komplexitet som behövs för att hantera det informella. Detta går då emot det Carley m.fl. hävdar som går ut på att social intelligens har att göra med att människornas hjärnor inte är de redskap som de kan använda för att greppa det informella. Enligt den synen är det sociala processer som exponerar människor för sådant som de inte kan uppfatta med sitt medvetande.

Så här långt bör vi uppmärksamma att det är en avgörande skillnad om intelligens kan vara distribuerad i en social struktur eller om intelligens bara finns inom människor. Om intelligens finns i en kontext som inte behöver vara omedelbart skapad av människor kan man ifrågasätta Searles argument som helt bygger på att hjärnan är överlägsen datorn. Då ifrågasätter man att frågan om komplexitet handlar om något som skall jämföras med människans hjärnkapacitet.

Vi ser åter att de argument som berör diskussionen om stark AI oftast är krångliga. Argumenten blir inte mindre krångliga av att människor när de interagerar med andra människor naturligtvis skapar kontexter för varandra. Och att detta kan ske utan att de behöver förstå varandra under interaktionen. Luhmann använder begreppet dubbel kontingens bl.a. för att visa på att människor i en interaktion sällan helt förstår den andre. Frågan kommer då att handla om huruvida situationer bara kan skapas av människor eller om också artefakter har en sådan funktionalitet. Det är ett rimligt antagande att frågan ställd på det viset är mindre kontroversiell än en fråga som handlar om huruvida maskiner kan ersätta mänskligt tänkande.

Vi bör uppmärksamma att Searle inte i första hand bygger sina argument på att människan är en social varelse. Han grundar istället sitt ställningstagande mot att maskiner kan tänka, på att de inte är lika effektiva maskiner som människornas hjärnor. Han betonar att han faktiskt menar att maskiner kan tänka. Men människornas hjärnor är de enda maskinerna som kan tänka:

"Could a machine think?" My own view is that only a machine could think, and indeed only very special kinds of machines, namely brains and machines that had the same causal powers as brains. And that is the main reason strong AI has had little to tell us about thinking, since it has nothing to tell us about machines. By its own definition, it is about programs, and programs are not machines. Whatever else intentionality is, it is a biological phenomenon, and it is as likely to be as causally dependent on the specific biochemistry of its origins as lactation, photosynthesis, or any other biological phenomena. No one would suppose that we could produce milk and sugar by running a computer simulation of the formal sequences in lactation and photosynthesis, but where the mind is concerned many people are willing to believe in such a miracle because of a deep and abiding dualism: the mind they suppose is a matter of formal processes and is independent of quite specific material causes in the way that milk and sugar are not (Searle,1980:424).

Vi bör också lyfta fram att trots att Searle argumenterar emot den klassiska AI:n, bör han ändå omfatta de antaganden om rationell bearbetning av intryck från omvärlden som den klassiska AI:n bygger på. Både Searle och klassisk AI hävdar att det komplexa skapas av en processande maskin. Det är då kapaciteten i den processande noden som är avgörande. Searle menar snarast att den klassiska AI:n inte fungerar på grund av att det inte går att bygga en dator som har hjärnans komplexa funktionalitet. Searles kritik är därför annorlunda än den sociologiska kritik som går ut på att det inte går att programmera den komplexitet som skapas i de sociala interaktionerna. Den är dock relevant för den sociologiska diskussionen

eftersom den sociologiska kritiken ändå till slut landar i en uppfattning som anger att vi aldrig kan komma förbi den mänskliga hjärnans kapacitet.

3.4 Kritiken mot Searle

En kritik mot Searle tar sikte på att han inte ser att det finns många olika grader av förståelse. Searle formulerar själv den kritiken som att den går ut på att förståelse inte bara handlar om att förstå eller att inte förstå. Den handlar istället om att förstå i mer eller mindre grad. Searle svarar att han håller med i sak, men att invändningen inte är applicerbar på hans exempel eftersom datorer inte alls kan förstå (Searle, 1980:418-419).

Argumentet innebär att datorer aldrig kan vara intelligenta eftersom artificiella aktörer inte alls kan förstå. Då går det heller inte att argumentera för att artificiella aktörer kan vara intelligenta genom att hävda att de artificiella aktörerna bara behöver ha en begränsad förmåga att förstå. De kan helt enkelt aldrig bli intelligenta eftersom de aldrig kan förstå. Searles argument framstår ändå oklart. Vad som fattas är att han inte ger någon antydning om hur människan förstår genom att hennes hjärna är en så överlägsen maskin gentemot dataprogram. Formellt handlar argumentet om att människornas handlingar styrs av intentioner. Men det ser ändå ut som att han söker hjärnans överlägsenhet för intentionella handlingar i något som handlar om att den kan hantera det informella. Searle får då ett problem med att visa på varför förståelse är så viktigt för diskussionen om AI, eftersom det blir svårt att se att vi kan hantera det som är informellt genom att lita till vår förmåga till förståelse. Lite längre fram skall vi se att Dreyfus understryker att de konnektionistiska modellerna inte behöver bygga på formella principer. Searles position kan i förhållande till Dreyfus formuleras som att hjärnan är den enda maskin som är tillräckligt kapabel att utföra de processer som konnektionistiska system bör kunna utföra. Men det innebär fortfarande att Searle tycks ta avstånd från att komplexiteten kommer från omgivningen. Det är hjärnan som skapar det komplexa. Den inställningen framkommer i det svar som Searle formulerar på en kritik mot hans experiment som går under benämningen *systemsvar*. Detta systemsvar är viktigt eftersom det ligger rätt nära en argumentation för att det sociala har att göra med faktorer som är skilda från människans intentionella kapacitet. Det är dessutom ett argument som ligger väldigt nära grunderna i en DAI-strategi. Sherry Turkle sammanfattar kritiken mot systemsvaret i boken *Leva Online* på ett sätt som också visar på kopplingen till biologiska frågeställningar. Hon tar fasta på att systemsvaret menar att intelligens ligger i ett samspel mellan olika beståndsdelar i samhället.

Även om mannen i rummet inte förstod kinesiska kunde kanske rummet i sin helhet sägas förstå kinesiska. På samma sätt finns det ingen del av hjärnan som förstår kinesiska, men hjärnan som helhet gör det. Med andra ord var intelligensen utspridd; den fanns i systemet som helhet, inte inom någon särskild agent inom systemet. Intelligensen låg inte hos något isolerat tänkande subjekt, utan i samspelet mellan en mångfald av intellektets beståndsdelar, bildligt talat, i ett samhälle av intellekt. Repliken var en indikation på vart datorkulturen var på väg på 1990-talet. Den gick tillbaka till några sedan länge övergivna bilder av emergent AI. [Sådan AI som uppkommit av sig själv, min anm.] Bilderna var biologiska och sociala. (Turkle,1997:159).

Searles försvar mot systemargumentet är att det inte hjälper om mannen i rummet lär in alla de kinesiska reglerna utan att förstå dem. Ty fastän de är inlärda så förstår han dem ändå inte. Om han inte förstår reglerna så kan inte det system reglerna bildar förstå, eftersom detta system bara är en del av mannen i rummet (Searle,1980:419-420).

3.5 Searles argumentation och diskussionen i denna avhandling

Skall man ta ställning till systemsvaret bör man först tänka efter hur man själv faktiskt kommer fram till förståelse. I första fasen av ett försök att förstå ett sammanhang tvingas vi reflektera över ett stort antal frågor som vi inte kan ha någon exakt förståelse av. Frågornas betydelse framkommer så småningom genom de sammanhang de framställs i. Vi tvingas hantera en mängd begrepp eller fenomen som får en mening bara i relation till andra fenomen. Det är en uppfattning som stöds av Luhmanns påstående att systemen måste förändra sin referensram innan de kan förstå något som finns utanför en given systemreferens. Det Searle diskuterar hör egentligen till en annan viktig fråga när det gäller att värdera AIs kapacitet. Den tar upp om människan behöver internalisera sin förmåga att använda redskapen för att kunna skapa kunskap med dem, eller om det räcker med att förhållandet mellan människa och redskap handlar om en externaliserad relation. Vi återkommer till den frågan genom en diskussion som pedagogen James Wertsch gör om konnektionistiska systems förmåga att fungera som redskap. I det perspektivet ser det ut som att Searle avvisar detta systemsvar alltför lätt. Det är snarare dessa systemrelationer som skapar förståelse än den faktiska tankeverksamheten. Vi kan inte säga att det är systemet som förstår, men

det är heller inte den väsentliga frågan då den inte handlar om förståelse ty frågan handlar om att systemen skapar **förutsättningar** för förståelse.

Searles avståndstagande från systemsvaret bygger på att det krävs komplex förståelseförmåga för att kunna använda den komplexitet som kommer från omgivningen. Systemet kan inte alls påverka människan om inte denna människa med en komplex funktionalitet kan förstå vad som är värt att ta in från det systemet exponerar. Det förefaller som att intelligens då måste ha att göra med en biologisk egenskap som utvecklats under en lång tid. Det människan upplever under sin livstid genom sin kulturella tillhörighet kan då inte utveckla människans kapacitet för intelligens. Och då kommer man tillbaks till att intelligens enbart handlar om genetiskt arv. Om vi skall undvika detta måste vi arbeta med ett förståelsebegrepp som innebär att vi först i efterhand förstår. Dessutom måste vi skilja begreppet intelligens från begreppet förståelse på så sätt att vi måste ha förutsättningar för att utföra intelligent beteenden om vi alls skall kunna hoppas på att förstå något som innebär att vi har tagit till oss ny kunskap. Det synsättet skall vi utveckla genom Luhmann som menar att förståelse inte är det centrala för sociala processer. Det som är centralt är istället kommunikation och det är bara sociala system som kan kommunicera.

Det är rimligare att det är relationerna i kontexterna, eller systemens förmåga att framhäva de viktiga relationerna i kontexterna, som bildar den viktigaste förutsättningen för förståelse. Det blir då inte den egna hjärnans förmåga som är avgörande utan möjligheterna att exponeras för intryck utifrån.

Även om förståelse inte behöver vara fullständig, eller om förståelsen uppnås först efter ett tag, kan man hävda att strukturella/kulturella förutsättningar för förståelse bara kan skapas av en mänsklig kunskapsproducent. Endast en sådan kan exponera ämnet på ett sådant sätt att mottagaren ges tillräckliga förutsättningar för att någonsin uppnå tillräcklig förståelse. I påståendet ligger att en artificiell förmedlare inte kan ha en sådan pedagogisk kapacitet. Detta är en central fråga som ändå inte har något givet svar. Dessutom skall vi vara medvetna om att det är ett tämligen sällsynt tillstånd att direkt samtala med kunskapsförmedlare i form av en människa som själv kan skapa kunskap. Oftast när människor är inblandade handlar det om intermediära förmedlare som också måste ses som en form av redskap. Och samtidigt bör vi vara uppmärksamma på risken med att en personlig förmedlare bara värnar sina egna kunskapsintressen.

Att Searle avvisar systemsvaret på grund av att det inte erbjuder något som är större än människan, antyder också att han missar det icke-formellas betydelse. Dess betydelse har att göra med alla intryck som vi kan reagera på, men som vi ännu inte har internaliserat i våra medvetanden eller formaliserat i vårt språk. Det består av de kontexter som vi har vant oss så mycket vid att vi kan ha utbyte av dem, fastän vi ännu inte har en fast bild av innehållet i alla dess detaljer. Frågan är om det inte är så att denna osäkerhet är det som gör att vi verkligen har material med vilket vi kan skapa något nytt.

Att Searles argument emot systemsvaret är dåliga behöver förstås inte innebära att systemsvaret är giltigt. Vid första tanken på systemsvaret tycker säkert de flesta att det svaret är orimligt. Turkles lite skeptiska inställning ger en bra beskrivning av de tankar man får kring systemsvaret. Detta systemsvar är vad som bör utredas bättre. En naturlig sociologisk argumentering för AI bör handla om att människans förmåga att förstå har att göra med den sociala värld hon lever i. Det är något som liknar systemsvaret. Dessutom är det i sig inte ett särskilt kontroversiellt sociologiskt påstående. En stark AI kräver emellertid att systemsvaret inte bara stannar vid att det är enskilda aktörers kollektiva handlingar som skapar systemet. Stark AI kräver att systemet kan ses som uttryck för en socialitet som fungerar fristående från de mänskliga aktörernas sociala handlingar. Stark AI skall då kunna skapa kontexter på egen hand.

3.6 För konnektionism och mot symbolisk AI

Den som tidigast kritiserade den klassiska AI-forskningen är filosofen Hubert Dreyfus. I boken *What Computers Cant Do*, första gången utgiven redan 1972, framhåller han att stark AI i princip är omöjlig (Dreyfus,1972:197-203). Dreyfus har senare hävdad att denna kritik inte träffar den konnektionistiska varianten av AI.

Dreyfus framhåller två frågor angående AI, vilka han menar är beroende av varandra: 1) Följer en människa som behandlar information formella regler på samma sätt som en dator? 2) Kan mänskligt beteende, oberoende av hur det har genererats, beskrivas på ett sådant formalistiskt sätt att det kan manipuleras av en dator?

Han menar att icke-programmerbart beteende är inblandat i alla former av intelligent beteende. Därför är det osannolikt att det kommer att ske några signifikanta framsteg med kognitiv simulering av artificiell intelligens (Dreyfus,1972:197).

Dreyfus huvudargument är att en del beteenden är så komplexa att de inte går att automatisera på det sätt som klassisk AI förespråkar. Argumenten kan tyckas likna Searles, men då Dreyfus tar sikte på komplexitet i beteenden istället för förståelseförmåga, blir hans ställningstaganden gentemot AI annorlunda än Searles. Det viktiga är att Dreyfus tycks omfatta tanken på att komplexiteten kan komma från det som ligger utanför den mänskliga hjärnan. Hans argument går ut på att det är för att komplexiteten just ligger utanför hjärnan som det inte går att formalisera den såsom strategierna bakom symbolisk AI tänker sig. Skapandet av ny kunskap handlar om informella beteenden. Därför går det inte att skapa en stark AI som bygger på symbolisk programmering. Det betyder att Dreyfus

argumenterar mot att frågan om huruvida hjärnan är den enda tillräckligt effektiva maskinen är relevant. Ty de som argumenterar för att hjärnan har en speciell kapacitet knyter behov av hjärnans komplexitet till att endast den kan processa symboler. Om komplexiteten söks i det som finns utanför människan bör det vara naturligare att tala om beteenden än om tänkande, eftersom det senare naturligtvis måste vara knutet till processer i människans hjärna. Det gör att Dreyfus argument bättre överensstämmer med att vi här talar om sociala beteenden och inte om mänskligt tänkande. Vi återkommer i nästa kapitel till Dreyfus argumentation.

3.7 Den situationsberoende handlingen, rationalismen och AI

Ett av de större problemen med diskussionen om AI är att den gängse uppfattningen kopplar ihop AI med en slags planerad rationalitet som måste ha att göra med medvetna handlingar. Det är därför viktigt att visa att utgångspunkten för AI-strategier inte behöver handla om en aktivitet som kräver planer. Det kan vi göra genom att ta del av en numera rätt stor diskussion kring AI som bygger på att det är ett redskap som behövs just därför att mänsklig aktivitet inte karaktäriseras av målrationalitet. Det är en diskussion som svänger på uppfattningen om AI. En uppfattning som gör att AI kan fungera just därför att vårt kunskapsinhämtande inte bestäms av målrationala strategier. Den diskussion om situationernas betydelse som Lucy Suchman för i *Plans and Situated Actions*, skapar något av en vetenskapsteoretisk förklaring till varför situationen är så viktig. Hon diskuterar framför allt relationen mellan planer som uttryck för medveten aktivitet, och de situationer individerna befinner sig i och försätter sig i.

Suchman börjar med att uppmärksamma att en traditionell intelligensuppfattning gör ett stort nummer av att det finns en skillnad mellan att navigera efter en planerad kurs och att navigera efter en vilja att komma fram till ett mål. Den första kopplar hon till det traditionella europeiska tänkandet. Exemplet på det andra beteendet hämtar hon från en artikel av Thomas Gladwin (1964) som har skrivit om hur en trukesisk navigatör tar sig fram till sitt mål.⁷ Om den europeiske navigatören stöter på problem måste han först ändra planen. Trukesen använder den information som erbjuds genom vind, vågor, tidvatten, faunan, stjärnorna och molnen. Efter dem styr han sin kurs. Hans ansträngningar handlar om att göra det som är nödvändigt för att nå målet (Suchman, 1987: vii-viii).

⁷ Truk är en stillahavs ö i ögruppen Carolinerna i Mikronesien.

Suchman menar att dessa två olika exempel kan leda till **tre** olika konsekvenser. Suchman menar att de första två är felaktiga.

Den **första** handlar om att det finns skilda sätt att handla som favoriseras olika mellan kulturer. Den europeiska kulturen favoriserar analytiskt tänkande där idealet är att resonera utifrån generella principer till särskilda fall. Trukeserna däremot, utan sådana ideologiska bindningar, lär sig en massa konkreta och förkroppsligade svar som bygger på minne och erfarenhet under år av faktiska resor. Suchman menar att detta sätt att skilja analytiskt tänkande från erfarenhet är felaktigt eftersom all mänsklig aktivitet, även den mest analytiska, bygger på erfarenhet. På så sätt bör vi nedtona tänkandet.

Den **andra** konsekvensen hävdar att om våra handlingar är planerade eller inte, beror antingen på aktivitetens natur eller på vår grad av expertis. På så sätt kan vi kontrastera instrumentella och målorienterade aktiviteter med kreativa och uttrycksfulla aktiviteter. Eller kontrastera nybörjarbeteende med expertbeteende. Suchman opponerar sig mot en sådan syn och menar att vi måste se att trukesen använder en instrumentell handling och litar till sin expertförmåga i lika hög grad som europén när han tar sig från den ena ön till den andra.

Den **tredje** konsekvensen är att avsiktliga handlingar nödvändigtvis är situerade handlingar. Med situerade handlingar menar hon handlingar som utförs i den kontext som bildas av speciella och konkreta omständigheter. Det förklarar att det som vi tror är resultat av tänkande, alltså avsiktliga handlingar, i själva verket är beroende av den situation de framställs i.

Suchman argumenterar för att denna tredje konsekvens är den rätta genom att hävda att den inte är ett symmetriskt alternativ till de andra två. Den är av en annan kvalitet och mer seriös därför att den ifrågasätter om det som representerar en plan verkligheten kan vara det som är grunden för handlingar i specifika situationer. Den är mer seriös därför att den handlar om grundförutsättningar. Därmed inte sagt att inte de andra två innehåller sanningar (Suchman, 1987:viii-ix).

Suchman försöker använda denna syn på handling för att diskutera hur vi skall designa intelligenta maskiner. Det går inte helt att översätta hennes resonemang till den problematik vi har tagit upp eftersom hon diskuterar expertsystem som är en del av den klassiska AI:n. Vid den tiden då hon skrev sin bok hade diskussionen om den konnektionistiska AI-varianten just startat. Det är dessutom missledande att diskutera situationerna i ett handlingsperspektiv. I överensstämmelse med det resonemang vi förde i kapitel 1 bör det snarare handla om ett processperspektiv.

Suchmans argument passar ändå för att understödja DAI-strategier eftersom de bygger på att den tredje konsekvensen som handlar om situationens betydelse är den mest betydelsefulla. Och eftersom konnektionismen just är en slags artificiell positionering kan Suchmans resonemang snarast ses som en argumentering för att konnektionism är något som representerar en form av artificiell social aktivitet. Den kan vara en analytisk modell för hur de sociala faktorerna i

kunskapsförmedlingen fungerar. Suchman säger också att hon i boken speciellt försöker utreda de fenomen som medger att datoriserade artefakter erbjuder en möjlighet att ifrågasätta distinktionen mellan det psykiska och det sociala. Suchman menar att ett sådant ifrågasättande också handlar om hur vi kan redovisa den delade förståelse, eller ömsesidiga intelligens, som vi som människor upplever i våra interaktioner med andra. Kan vi svara på det menar hon att vi också kan förklara hur interaktionen ser ut mellan människor och vissa maskiner [som AI-datorer, min anm.] (Suchman,1987:6)

Ovanstående stödjer tanken på att det sociala inte behöver grundas på individuella handlingar. Mer konkret uppmärksammar hon att interaktionen mellan personer är lika problematiska att förklara som den process där en person interagerar med datorer. Det betyder alltså att social funktionalitet mellan personer är lika svår att få grepp om som en social funktionalitet mellan datorsystemen och mellan ett datorsystem och en människa.

Suchman tangerar frågan om hur datorerna själva interagerar med varandra. Hon tar då sikte på att den relativa socialiteten hos datoriserade artefakter baseras på det faktum att medlen för att kontrollera datoriserade maskiner och deras beteende i allt högre grad bygger på lingvistik och i mindre grad på mekanik. Det betyder att interaktioner med maskiner mer och mer sker genom ett naturligt språk och inte så mycket genom att trycka på knappar (Suchman,1987:11).

Lingvistiska teorier brukar anges som argument för att hävda att den starka AI:n inte kan vara möjlig, eftersom man menar att det bara är människan som kan hantera språket. Det intressanta med Suchmans bok är att hon visar att de lingvistiska teorierna innehåller element som inte bara har att göra med människors förmåga. Det leder henne till att se att det är situationerna som avgör interaktionen mellan människa och maskin, och inte människornas språkbehandling.

Problemet med talhandlingar är att de kräver planer. Och planer, menar Suchman, representerar ofta bara det som andra människor tillåter. På så sätt blir planer styrda av intressen. Vilket bör vara en av orsakerna till den negativa socialitet vi har talat om. Skall vi studera intelligenta handlingar måste vi studera hur människor använder situationerna snarare än hur de planerar handlingar. Istället för att inordna handlingarnas detaljer under studiet av planer skall vi inordna planerna under det större problem som situerade handlingar erbjuder (Suchman,1987:50).

Enligt Suchmans argumentation är det situationerna som skapar planerna istället för att planerna skapas av människorna. Det finns ingen absolut objektivitet. Det som finns är uppfattningen av situationerna. Frågan om de konnektionistiska systemens kapacitet är av en dubbel natur. Å den ena sidan kan vi säga att om de bara kan hantera situationer som är bestämda genom representationer så är de lika begränsade som de positioneringar som sker med hjälp av språket. Om situationerna å den andra sidan representerar något mer än det begränsade språket behöver vi redskap som kan tillåta situationerna att

kommunicera. Det råder en viss tveksamhet huruvida de konnektionistiska systemen kan omprogrammera sig själva så att de kan fungera utanför de ramar som det till människan knutna språket ger. Men om de kan det, så är de sannolikt inte ett sämre redskap för kommunikation än det talade språket. Här tangerar vi en fråga som är väsentlig för att värdera den luhmannska systemteorin. Den frågan handlar om huruvida det är rimligt att anta att andra system än de psykiska, d.v.s. människor, kan observera. Vi återkommer till denna centrala problemställning.

Om vi återgår till Suchmans resonemang så menar hon att planer är resurser för situerade handlingar, men de bestämmer inte vilka handlingar som blir resultat av planerna. Det slutgiltiga resultatet ges av villkoren i situationen. Suchman illustrerar detta med hur man betar sig om man vill åka kanot genom svåra forsar. (Suchman,1987:52). Man börjar med att stanna upp och ta sig en titt på fallet och planera en strategi.

The plan might go something like "I'll get as far over to the left as possible, try to make it between those two large rocks, then backferry hard to the right to make it around that next bunch." A great deal of deliberation, discussion, simulation, and reconstruction may go into such a plan.

Men hur detaljerade planer man än gör så tvingas man ändå att reagera på alla oväntade situationer som uppstår då man faktiskt åker iväg. Man får då falla tillbaks på de färdigheter man sedan tidigare har.

The purpose of the plan in this case is not to get your canoe through the rapids, but rather to orient you in such a way that you can obtain the best possible position from which to use those embodied skills on which, in the final analysis, your success depends.

Detta betyder att planer bara kan hjälpa oss att förbereda oss för att ta emot det som situationen ger oss. De kan bara vara en förberedelse för att se de situationer som ger mening. De kan bara sätta igång den kommunikativa kapaciteten. De kan inte i sig vara det som omformulerar livsvärlden. Att vi tror att det är den mänskliga handlingen som skapade våra möjligheter att se, har att göra med att vi inte kan tolka det som funnits i situationen innan vi har varit igenom den. Eftersom det var vår plan som förde oss in i situationen tror vi att vi har följt en plan (Suchman,1987:52).

Vi har i detta kapitel sett att det är ifrågasatt om intelligens handlar om tänkande eller om ett beteende. Med hjälp av konnektionistisk AI, Luhmanns systemteori och idéer om situationernas betydelse har vi visat på att de sociala processer som är inblandade i skapandet av ny kunskap åtminstone inte handlar om tänkande.

Vi har visat att klassisk AI måste jämföras med människornas tänkande och att den formen av intelligens knappast kan artificialiseras. Konnektionistisk AI står däremot mycket nära sociala processer kring kunskapsskapande. Vi har sett att det finns mycket som talar för att konnektionistisk AI kan vara förenlig med sociala processer just därför att sådana processer måste använda den kulturella sociala struktur som består av olika typer av artefakter.

Vi har också visat på att Searles argument mot AI blir ifrågasatt just därför att sociala processer kring kunskapsskapande har att göra med de möjligheter som ges av situationer. Förståelse är något som uppstår först efter att situationerna gett oss möjligheter att förstå.

Dessutom har vi diskuterat huruvida det finns anledning tro att konnektionistisk AI som en modell för social AI kan ha förmågan att programmera om sig själv. Med hjälp teorier om situationernas betydelse för att se nya möjligheter har vi visat på att det är en fråga som inte så lätt går att avfärda. Talhandlingar som uttryck för socialitet kräver planer vilket gör att den typen av socialitet kräver formalism.

4. Det sociala i AI-diskussionen

4.1 Metodologisk individualism eller Holism och AI

Vi har redan nämnt att den kritiska inställning som de flesta sociologer visar mot AI oftast bygger på uppfattningen om att sociala processer måste förklaras genom de handlingar som individerna utför. Alltså den position som brukar kallas MI. Vi har också visat på att argument för stark AI bör höra samman med en holistisk uppfattning om sociala processer även om mycket av kritiken mot AI kommer från den holistiska sociologin. Innan vi går vidare bör vi se vad metodologisk individualism och holism egentligen innebär.

Sociologisk uppslagsbok definierar metodologisk individualism som *en princip enligt vilken den sociala verkligheten i grunden utgörs av enskilda individer som handlar mer eller mindre ändamålsenligt i ljuset av sina dispositioner och sin uppfattning om den egna situationen*. Men man menar också att situationen påverkas av fysiska resurser och miljö (Bengtsson&Hjern,1973:175).

Vi kan jämföra detta med den åsikt som är motsatt MI - den metodologiska holismen. *Filosoflexikonet* anger: *Enligt den metodologiska holismen bör sociala helheter undersökas och förklaras som sociala helheter och inte reduceras till något annat, t.ex. enskilda personers handlingar*" (*Filosoflexikonet*, 1988:235).

Encyclopedia of the Social & Behavioral Sciences förklarar möjligheterna av att jämföra individualistiska förklaringar med holistiska förklaringar genom att peka på att MI kan omfatta att det finns sociala faktorer som inte är reducerbara till individuella kapaciteter. Men eftersom det bara är individer som har mål och intressen är det dessa som förklarar mänskligt beteende. Dessutom hävdar MI att de sociala villkoren bara kan förändras genom individers handlingar (*Methodological Individualism in Sociology*,s.9751-9755).

Ovanstående visar på att det inte är så enkelt att verkligen se skillnaden mellan individualistiska och holistiska förklaringar. Det är något som också framhålls av Lars Udéhn som hävdar att det finns en stark och en svag variant av MI. Den första medger bara att sociala fenomen kan tolkas i termer av individer och deras interaktioner medan den svagare versionen också tillmäter sociala institutioner och sociala strukturer en förklarande roll (Udéhn,2002).

4.2 AI och praktisk kunskapsförmedling

4.2.1 Randall Collins experiment

Det inlägg i den sociologiska debatten som inom traditionell sociologi tycks har banat vägen för en mer intresserad analys av AIs möjligheter är Randall Collins kapitel med titeln *Can sociology create an artificial intelligence* i andra utgåvan av boken *Sociological Insights* (Collins, 1992).

Collins menar att utmaningen är att få en dator att bete sig som en verklig människa. Det innebär en dator som kan lära sig att presentera nya idéer. Den skall alltså inte bara kunna lösa problem utan också kunna skapa nytt, göra vetenskapliga upptäckter, skriva litteratur och kanske skapa musik (Collins, 1992:159). Collins hävdar att vi vet att människor kan göra allt detta. Men också att det finns en särskild typ av social interaktion bakom allt sådant tänkande. De som skapar nya vetenskapliga teorier interagerar med andra människor. Collins påstår att vi faktiskt vet ganska mycket om på vilket sätt vissa nätverk skapar kreativa vetenskapsmän (Collins, 1992:159-160).

Det blir lite oklart vad det är vi vet om detta nätverkande. Vi bör tänka på att människor sällan i kunskapsförmedling kan interagera med varandra utan att använda sig av artefakter som texter. I det perspektivet är det inte säkert att kommunikation som går genom läsande och skrivande bara är avledningar av direkt kommunikation mellan personer ansikte mot ansikte. God kunskapsutveckling kräver också social kommunikation. Det luhmannska perspektivet på kommunikation ger vid handen att den inte baseras på de enskilda individernas handlanden. När vi studerar kunskapsskapande ser vi gärna interaktioner mellan olika personer men vi glömmer att studera den artefaktiska struktur som faktiskt gör det möjligt för människorna att interagera med varandra! Vi glömmer ofta att situationer som skapas av annat än människor också bildar en del av förutsättningarna för en god kommunikation. Resonemanget antyder att det finns en skillnad mellan sådana interaktioner som sker när människorna direkt möter varandra och sådana som sker där människor kommunicerar med varandra via olika former av artefakter.

Collins text bygger på en diskussion av hur datorer kan utföra de funktioner som finns i en konversation. Collins menar att han grundar resonemanget på hur den sociologiska teorin ser på hur sociala interaktioner utvecklas genom att medvetandet tar del i en konversation. Collins säger inte i texten vilka sociologiska teorier det handlar om, men litteraturreferenserna till kapitlet anger mikro-orienterade teorier om symbolisk interaktionism och om konversationsanalys. Collins anlägger i slutet av kapitlet en rätt så nyfiken attityd till vilken möjlighet det finns att få datorerna att fungera på samma sätt som människor. Han anger att han personligen tror att vi kan bygga en dator som uppför sig som en människa. Han menar också att AI-forskningen måste lära av sociologin.

Collins Experiment bygger på 13 regler (Collins,1992:165-181). De fyra första reglerna handlar om att datorerna skall kunna reagera på personligt tilltal. Jag uppfattar heller inte att Collins menar att dessa första regler inte kan hanteras av datorer.

Däremot verkar det som att Collins är mer tveksam till om datorer klarar av det som han menar handlar om emotionell energi, EE. Regel 5-10 handlar om denna emotionella energi. Det handlar om sådant som att den emotionella energin ökar när man samtalar med folk som har mycket status.

Utifrån Luhmanns teori kan vi utveckla hur denna emotionella energi leder till något nytt. Han menar att system alltid strävar efter att upprätthålla sina funktioner. Det som tvingar systemen till förändring är de irritationer de utsätts för. Utifrån det perspektivet kan vi argumentera för att emotioner kan leda framåt men att de också kan fresta oss att stanna kvar vid det invanda. Människors emotioner gör ofta att vi försöker stöta bort sådant som irriterar oss. Då får emotionerna snarast en roll som binder oss samman med den gamla kunskapen. Skall vi hitta framåt är det bra om vi finner det väsentligt att också kommunicera med sådan kunskap som vi inte redan på förhand har värderat positivt. Enligt systemteorin är det just de olikheter som gör oss irriterade som gör att vi så småningom börjar leta oss framåt. De emotioner som uppstår som irritationer måste sannolikt bearbetas med något som liknar klarsyn om vi skall se att det som irriterar oss istället erbjuder en möjlighet att observera från en ny synvinkel. Poängen är att det inte är emotionerna i sig som erbjuder nya möjligheter. Det är istället möjligheten att hamna i nya situationer som skapar förutsättningar för kreativitet.

Reglerna 11-13 förefaller vara de som Collins uppfattar som mest kritiska om vi skall kunna utveckla AI. De handlar om förutsättningar för kreativitet och speciellt om hur den AI-dator som Collins kallar SOCIO kan komma i kontakt med lärare som redan är kreativa. Datorn behöver kunna göra detta eftersom nätverket runt kreativa individer består av kontakter med andra kreativa individer (Collins,1992:180). Datorn måste då kunna konstruera en konversation med det kulturella kapital som används i flera olika grupper och omorganisera idéerna så att de blir meningsfulla (Collins,1992:181-182). Denna avhandling försöker, genom att använda Luhmanns systemteori, visa att kreativa situationer skapas genom systemens kommunikation. På så sätt finns det möjligheter att argumentera för att det just är en kommunikation som går via de sociala systemen som gör att människorna kommer i kontakt med sådant som andra kreativa individer skapat. Collins visar att det också handlar om att komma i kontakt med de unga forskare som i egenskap av revolutionärer försöker störta gamla idéer (Collins,1992:181). Det är inte helt klart vilken betydelse Collins menar att emotionell energi har för möjligheterna att nå kontakt med kreativa individer. Men vi kan ifrågasätta att emotionell energi gör att datorerna eller människorna dras till revolutionärer. Det

måste nog ske genom att systemkommunikationen först förmår exponera de möjligheter som finns i oppositionella idéer.

4.2.2 Relevansproblematiken, människan och möjligheterna för observation

För att illustrera att emotionell energi ofta kan leda oss fel, skall vi bekanta oss med något som i den biblioteksvetenskapliga litteraturen kallas *psykologisk relevans*. Den forskningen fokuserar på hur vi uppfattar relevansen i de texter för vilka vi exponeras. Den visar på hur komplexa relationerna är mellan individernas medvetande och den omgivande strukturen. Dessutom visar den på att vi oftast är hänvisade till att ganska blint försöka hitta något som kan få oss att se ny kunskap.

I artikeln *Psychological relevance and information science* menar Stephen P. Harter att för att en artikel skall vara psykologiskt relevant måste den skapa kontextuella effekter på en individ. Han menar att det inte är säkert att ett dokument som behandlar det ämne som en sökning [i en databas] frågat efter, åstadkommer en sådan effekt (Harter,1992:607).

Detta bör betyda att den emotionella energi som styrde sökfrågan från början, måste förändras om personen skall uppleva texten som relevant. Det blir då ett glapp mellan olika sätt att se på relevans. Det som man emotionellt förväntar hör till ett ämne blir ur ett psykologiskt perspektiv irrelevant. En sökning efter sådana referenser som man redan vet att man tycker om erbjuder ingen ny kunskap. Den konfirmerar bara tidigare kunskap. En referens till en artikel som personen bara tycker är svagt relevant, kan däremot visa sig ge kontextuella effekter om personen verkligen bryr sig om att läsa artikeln (Harter,1992:607-608).

Vi kan ta detta som ett tecken på att individernas emotioner bara förmår visa på i vilken riktning ett sökande efter kunskap skall gå. Vi vet oftast inte vad vi söker efter. Vi är i ett tillstånd som i litteraturen om relevans brukar betecknas som "Anomalous state of knowledge" eller förkortat ASK (Belkin et.al,1982). Även om vi hittar relevant information kanske vi bara uppfattar relevansen långt senare när själva exponeringen för det nya har förändrat vårt medvetande. Om en individ som gjort en sökning läser den litteratur som referensen syftar på och inte upplever kognitiv förändring kan ändå en sådan förändring uppstå långt senare. Uppstår en förändring beror den inte bara på att själva artiklarna har lästs. Den beror också på att läsarens förmåga att ta till sig innehållet har förändrats då tid och rum exponerat dem för annat. Först då har de sett relevansen i de tidigare artiklarna (Harter,1992:608). Harter menar dessutom att det som man inte genast uppfattar som relevant är viktigast, därför att det är inkörsporten till nya intellektuella sammanhang (Harter,1992:612).

Vi kan hävda att detta sätt att uppfatta relevans ligger nära en fenomenologisk uppfattning. En av den sociologiska fenomenologins företrädare, Alfred Schütz, har redan tidigt försökt sociologisera relevansbegreppet. Schütz talar om hur den

tolkade relevansen är beroende av det han kallar systemet av ämnesmässig relevans. Han framhåller att det inte finns någon tolkad relevans i sig, utan bara sådan tolkad relevans som förhåller sig till ett givet ämne. I överensstämmelse med detta finns det heller ingen speciell typ i sig. Det finns bara typer som relaterar sig till särskilda problem. Schütz menar att denna karaktäristik av typer är av fundamental betydelse för samhällsvetenskapens metodologi (Schütz, 1970:63-64).

Luhmann själv menar att hans användning av funktionsbegreppet också är ett försök att utveckla den fenomenologiska traditionen. Och vi kommer att redovisa texter som hävdar att det finns en möjlighet att se Luhmanns teori som ett komplement till Habermas användning av livsvärldsbegreppet. I *International Encyclopedia of the Social & Behavioral Sciences* (s.13606) talas också om att Schütz har omformulerat den fenomenologiska uppfattningen av mening, så att den handlar om en självskapande egenskap hos sociala och psykiska system. Artikeln påpekar också att det är en av influenserna till Luhmanns systemteori.

4.3 Den sociologiska AI-diskussionen

Vi har redan sett några exempel på hur AI-frågan har behandlas inom sociologin. De texter som presenteras i detta avsnitt är sådana som kan ge en bild av de allmänna tendenserna inom den sociologiska AI-forskningen och de presenteras kronologiskt.

4.3.1 Den tidigaste diskussionen

En av de första artiklarna om AI i sociologisk litteratur är skriven 1985 av Steve Woolgar med titeln *Why not a sociology of machines? The case of sociology and artificial intelligence*. I den artikeln diskuterar Woolgar existerande AI-forskning. Woolgar menar att möjligheterna för en koppling mellan sociologin och AI-forskningen, är avhängig av i vilken mån sociologin ansluter sig till de grundläggande frågor som AI-forskningen ställer. Han pekar på att sociologin kan ha intresse av AI-forskning på två sätt. Det första handlar om att AI-forskningen erbjuder en intressant modell för att testa om vetenskapssociologin kan användas för att studera förhållandet mellan människor och maskiner på ett mer generellt sätt. Det andra handlar om huruvida maskiner tillräckligt mycket liknar människor för att kunna betraktas som subjekt i en sociologisk analys (Woolgar, 1985:567-568).

Den första sociologiska konferensen kring AI (Gilbert&Heath,1985)⁸ hölls samma år som Woolgar skrev sin artikel, men sociologin har varit mycket dålig på att ge teoretiska bidrag till AI-forskningen. Ett faktum som också den till USA emigrerade svenske sociologen Bo Anderson påpekade redan 1989. Han menade att andra beteendevetenskaper varit bättre på att teoretisera frågan om AI (Anderson,1989). Ett påstående som understryks av Douglas Heckathorns uppfattning att den interdisciplinära forskning som mest har undersökt AI, kognitionsvetenskapen, bygger på sex traditionella vetenskapliga discipliner: neurovetenskap; psykologi; artificiell intelligens; filosofi; lingvistik och antropologi (Heckathorn,1989:97). Hans diskussion handlar om hur sociologin kan dra nytta av att inkorporeras i denna forskning.

4.3.2 Senare diskussion

I början av 90-talet började sociologin visa ett större intresse för AI och för den distribuerade varianten av AI. Framför allt handlar det om vad konnektionistisk programmering betyder för möjligheterna av att använda AI för att modellera sociala processer.

Bland artiklar från denna andra våg av sociologisk AI-diskussion finns en av de mer kritiska artiklarna. Den är skriven av Alan Wolfe i *American Journal of Sociology* med titeln *Mind, Self, Society, and Computer: Artificial Intelligence and the Sociology of Mind*. Wolfe ifrågasätter starkt både konnektionismen och det luhmannska systemtänkandet som ett sätt att förena AI-strategier med sociala processer. Wolfe tar avstånd från möjligheterna att grunda stark AI i den luhmannska teorin därför att människornas kapacitet för intellektuella aktiviteter relativiseras i den luhmannska teorin. Wolfe menar att detta antagande bygger på ett fel i denna teori (Wolfe,1991:1090ff.). Wolfes argument går ut på att människorna har ett så speciellt medvetande att det inte kan kopieras med symbolisk databehandling eller med de neurala nätverk som konnektionismen bygger på.

Vi kan också hänföra Randall Collins diskussion och den diskussion om artificiell social intelligens som Carley m.fl. gjorde till denna period. Både Carley och Collins måste räknas till dem som förhåller sig positiva till AI, men ingen av dem har argumenterat för DAI.

4.3.3 Den senaste diskussionen och forskningsprogrammet *Socionics*

Kopplingen mellan DAI och sociologisk systemteori syns tydligast i en tysk sociologisk forskning om AI som går under benämningen Socionics. Den är mycket relevant för den sociologiska diskussionen om AI eftersom den är fokuserad

⁸ Surrey conferences on sociological theory and method, 3

på att reda ut vilken betydelse begreppet social har. För att utreda begreppets betydelse lutar man sig mot det sätt att uppfatta dess innebörd som Niklas Luhmann använder. Denna forskning finns bland annat samlad i antologin *Sozionik: Soziologische Ansichten über künstliche Sozialität* utgiven av Thomas Malsch. I en av artiklarna som ursprungligen publicerades i *Zeitschrift für Soziologie* argumenterar Malsch för att sociologin kan lära sig mycket av forskningen kring hur vi kan modellera artificiella samhällen.

Malsch skriver att det anmärkningsvärda är att denna nya DAI-forskning också använder metaforen social på ett nytt sätt. Genom den konnektionistiska idén har AI-forskningen kunnat ta sig ifrån den klassiska AI-forskningens bindning till den metodologiska individualismen (Malsch,1997:6). På så sätt har AI-forskningen också kunnat komma med nya idéer för den så kallade mikro-makro länken inom sociologin. Den länk som sociologin ser som förbindelsen mellan delar och helhet.

Malsch frågar om det finns någon skillnad mellan hur DAI-forskningen och sociologin använder det grundläggande begreppet social (Malsch,1997:8-9). Hans inställning är att ett par viktiga modeller för att bedöma AI är fast i en försociologisk uppfattning om den individuella människan. De modeller han nämner är *Turingtestet* och Searles *kinesiska rum*. Om sociologerna använder dem menar Malsch att sociologin blir inspärriad i definitioner av problemet med social intelligens som bestäms av hur filosofer och psykologer uppfattar vad det sociala är. Sociologin måste då omfatta den psykologiskt grundade kognitionsvetenskapens uppfattning om AI (Malsch,1997:9). Malsch menar alltså att den MI grundade uppfattning om vad som karaktäriserar den socialt orienterade agenten, innebär att sociologin retirerar från naturliga sociologiska positioner, vilket reducerar sociologin till en kognitionsvetenskap.

I artikeln *Naming the Unnamable* menar Malsch att den semantik som DAI använder måste bli genomsyrad av sociologin för att den skall kunna ge ett bidrag till den sociologiska teorin och vice-versa. Malsch varnar för att argumenten kring AI lätt blir cirkulära och att vi inte får blanda ihop de frågor som hör till sociologin med de frågor som hör till datavetenskapen (Malsch,2001:160-161). I denna avhandling ligger betoningen på att diskutera vilken betydelse sociologins uppfattning om begreppet social har. Genom att diskutera detta centrala begrepp bör vi kunna undanröja en orsak till sociologins cirkulära hållning till frågan om AI. Förhoppningsvis skall vi då också ge ett bidrag som underlättar för den datavetenskapliga AI-forskningen att göra sin strategi mer verklighetsinriktad.

Malsch hävdar att varken klassisk AI eller konnektionism kan stödja en sociologisk teoriutveckling. Han ser en skillnad mellan konnektionism och DAI och hävdar att det bara är DAI-forskningen som kan hjälpa till att utveckla sociologin (Malsch,2001:161-162). Malsch framhåller att den klassiska AI inte kan utveckla sociologin eftersom den bara ser intelligent handling som något en ensam aktör utför. Konnektionismens svaghet är att den bara ser handling som en

funktionalitet i nätverken och saknar därmed ett begrepp för självständig handling. Här finns det en skillnad mellan det systemteoretiska perspektiv på DAI som används i denna avhandling och det perspektiv på DAI som Malsch använder. Trots att Malsch och många andra av de forskare som deltar i diskussionen om socionik använder Luhmanns systemteori som ett analytiskt redskap, ser de på DAI som en modell där flera agenter agerar tillsammans. Därför kan det se ut som om Malsch ansluter sig till en modell som brukar kallas Multi Agent Systems (MAS)⁹. Det är en modell som fortfarande är nära förknippad med att intelligens handlar om en egenskap som har så komplex funktionalitet att den måste höras samman med människans hjärna. Ett systemteoretiskt perspektiv på DAI, som är mer i överensstämmelse med att Luhmann tonar ner individernas betydelse, bör istället utgå från det perspektiv på intelligens som säger att intelligens är en egenskap som finns lagrad i kontexterna vilket innebär att betoningen inte behöver ligga på kapaciteterna i olika agenter. Det är det perspektivet som medger att vi kan gå ifrån påståendet att människans natur representerar något som i ontologiskt hänseende är av en sådan reell natur att vi aldrig kan ersätta henne med artificiella aktörer. Det finns även i en systemteoretisk modell anledning att hävda att det finns aktörer, eller observatörer, inblandade. Men i den teorin har dessa aktörer en enkel funktionalitet och inte den komplexa funktionalitet som vi kopplar till den mänskliga hjärnan. På så sätt kan det vara en definitionsfråga om det finns en skillnad mellan en MAS-modell och en DAI-modell. Det är dock mest i överensstämmelse med en systemteoretisk modell att använda benämningen DAI.

Detta systemteoretiska perspektiv innebär också att det finns mindre anledning hävda att det finns en skillnad mellan konnektionism och DAI. I systemteorin behöver inte den distribuerade intelligensen finnas nedlagd i artificiella aktörer. Det räcker med att se på funktionaliteten i nätverken. Om själva programmeringstekniken inte är det viktigaste, utan att det istället är möjligheterna att registrera komplexiteten i omvärlden som är avgörande, så finns det mindre anledning att bekymra sig för att konnektionismen skulle sakna ett begrepp för självständig handling. Det viktiga är inte att programmen i sig innehåller en komplex funktionalitet. Det viktiga är istället att programmen hjälper till att lyfta fram den komplexa funktionalitet som blir resultatet av att kontexterna genom programmen bättre kan påverka kunskapsproduktionen.

Socionics-forskningen har i Sverige bara nämnts i en uppsats. Simon Flyman redovisar den skillnad som finns mellan den AI-forskning som diskuterar hur AI-aktörer fungerar utan inblandning av människor, och den som diskuterar hur AI-aktörer och människor fungerar tillsammans (Flyman, 2001). Flymans syfte är att diskutera den senare. Frågan om stark AI innebär att vi måste diskutera den förra.

Det kan synas som ett naturligt sociologiskt perspektiv att diskutera hur teknik och människor fungerar tillsammans. Och att intresset för hur artificiella

⁹ Bygger på att många små men kompetenta aktörer samarbetar.

aktörer fungerar tillsammans bara är en fråga som datatekniken bör intressera sig för. Det är här som de sociologiska teoriernas ontologiska dimension blir viktig. Utifrån ett MI-perspektiv bör naturligtvis sociologin inte intressera sig för mer än den första frågan. Problemet med detta perspektiv är att det innebär att fokus ligger på frågor som handlar om de inre mentala processer som andra vetenskaper, framför allt psykologin, är bättre på att analysera. Förklaringen till att sociologin varit så dålig på att ge teoretiska bidrag till AI-forskningen kan handla om att den diskuterat svag AI istället för stark.

Frågan om hur AI-redskapen fungerar tillsammans har, som sagts, istället mycket gemensamt med den holistiska sociologi som försöker förklara hur sociala processer fungerar utan att de behöver grundas på individuella handlingar. Bägge måste leta efter andra funktioner än sådana som har sitt ursprung i människornas medvetanden. Det förklarar att ett av sociologins intressen för AI-frågan har handlat om att med AI kunna modellera sociala processer eller hela sociala samhällen. I denna avhandling bygger en del av resonemanget på att det är de artefakter, av vilka AI är en del, som den humanistiska kulturen har skapat, som skapar den funktionalitet som medger att sociala processer kan fungera fristående från mänskliga handlingar.

4.4 Dreyfus, konnektionismen och den sociologiska kritiken

Dreyfus utvecklar tillsammans med sin bror Stuart sin inställning till konnektionismen i en artikel i tidskriften *Daedalus* (Dreyfus&Dreyfus,1988). Där hävdar de att de neurala nätverk som är grunden för de konnektionistiska strategierna, i princip inte är fångna av samma formalism som den klassiska AI:n. Det beror på att de åtminstone i teorin inte bygger på kognitiv simulering. Även om de delvis programmeras utifrån formella regler kan dessa regler förändras av processen själv. Dreyfus position bör därför möjliggöra att de processer som exponerar komplexiteten i världen kan modelleras genom datorprogram.

What Computers Can't Do gavs i början av 90-talet ut i en nyutgåva med en delvis ny titel *What Computers (Still) Can't Do* (Dreyfus,1992). En diskussion kring den fördes i ett nummer av tidskriften *Artificial Intelligence*. Den intressantaste kritiken kommer från kunskapssociologen Harry Collins (Collins,1996). I det nyskrivna förordet till den nya upplagan uttrycker Dreyfus än en gång att konnektionistisk AI kan vara stark. Det är den hållningen som Collins starkast kritiserar i sin artikel. Den kritiken tar sikte på att det är de sociala processerna som omöjliggör en stark AI.

Collins kritik av Dreyfus angående neurala nät handlar om att det är tveksamt om konnektionistisk programmering innebär något annat än gammaldags AI, "Good Old Fashioned AI" (GOFAI), alltså den klassiska AI:n. Collins menar att Dreyfus bygger sitt argument på att konnektionismen innebär att programmen inte först måste ha beskrivningar som representerar en lagbunden värld innan de kan fungera. Den kritik Collins riktar handlar om att det måste vara något fel på detta argument eftersom det redan finns generationer av intelligenta maskiner som inte uttryckligen behöver regler för att fungera. Till exempel expertsystem som istället upptäcker regler. Och de programmen uppfattar vi inte som några genombrott för AI. Enligt Collins gör Dreyfus inte klart vilken kvalitativ skillnad det finns mellan sådana program och neurala nätverk (Collins, 1996:114).

Collins argument angående den tidigare generationen av intelligenta maskiner visar på hur viktigt det är att de konnektionistiska datorerna fungerar på ett annat sätt än genom en naiv associationism. De måste kunna använda komplexiteten i världen till att göra mer än att avspegla det som de ser vid en viss tidpunkt. Vitsen med neurala nätverk är just att de skall vara kapabla att förändra sina egna beteenden genom den kontakt de har med omgivningen. Exemplet visar på hur svårt det är att faktiskt diskutera den konnektionistiska AI:n på dess egna premisser istället för på premisser som kritikerna tillskriver den. Det är viktigt att vi observerar att Collins här missar poängen i vad som konnektionismen försöker vara. Han avfärdar den på grunder som den själv avfärdar. De grunder som den avfärdar är ju att kunskap skapas genom att upptäcka regler. I konnektionistiska system handlar det istället om att regelverket hela tiden skall kunna förändras tack vare aktiviteten i systemet. Vi kan kritisera konnektionismen för att den inte förmår visa hur en sådan funktionalitet kan skapas. Men vi kan inte kritisera den för en AI-design som den själv försöker utveckla ett alternativ till. Vi bör också tillägga att de regler som expertsystem upptäcker är naturlagar. Inte sådana regler som skapar dynamik i noder. Regeln i ett konnektionistiskt system kan bara vara ett hjälpmedel (heuristik) för att skapa funktionell aktivitet. Inte det som avgör vad som blir resultatet av processen. Möjligen går det att tolka Collins som att han menar att konnektionismen i praktiken kräver samma fasta regler som den klassiska AI:n. I så fall är det förstås ingen skillnad mellan dem. Det är därför som vi måste diskutera förutsättningar för funktionalitet i den processande enheten.

Collins argumenterar här utifrån ett antagande om att det skulle vara lätt att veta hur sociala processer fungerar. Han gör det utifrån ett antagande om att människor verkligen klarar av motsvarande processer utan att använda regler och att det är något som bygger sociala processer. Vi har redan varit inne på att en förklaring till hur social intelligens fungerar är lika svårt att beskriva som att förklara hur konnektionistiska system fungerar. I det perspektivet bör vi inse att människornas tidigare erfarenheter och kunskapsnivåer gör att de gärna fastnar i ett fördomsfullt sätt att betrakta världen genom att de försöker klassificera det nya med sådana begrepp som de redan känner till. De möter världen genom de begrepp

de redan delat in världen i. Den insikten talar för att frågan om människans möjligheter att förändra sig, inte heller är oproblematisk. Det kan vara så att också människan fastnar i fasta regler. Det öppnar för att det kan vara så att tekniken är det som gör att människorna inte alltid behöver fastna i sin tidigare kunskapsuppfattning. Det kanske är så att Harry Collins ändå är fast vid en normativ definition av begreppet social som vi kunde undvika genom att diskutera sociala processer utifrån Luhmanns systemteori.

Frågan handlar om huruvida det finns någon avgörande skillnad mellan människors och datorers förmåga att förändra sig i relationen med omvärlden. Den centrala frågan blir då om de konnektionistiska nätverken kan förändra sina representationer av omvärlden så att de kan ta emot omvärlden på ett nyprogrammerat sätt. Vi får väl anta att människan klarar av att göra detta - klarar av att se världen med nya ögon, som det brukar heta. Men vi bör hålla i minnet att vi människor också har svårt att omprogrammera oss så att vi tillåter oss att se det nya. Vi kan här säga att vi står inför någon form av parallell till det problemet. Det finns skäl att tro att det artificiellt intelligenta systemet har samma svårigheter till omprogrammering som människan. Däremot bör datorernas kapacitet att skanna igenom stora mängder texter innebära att de har bättre chanser att försätta sig i nya situationer. I sammanhanget kan vi fråga oss om människan kan se på något med nya ögon utan att ha redskap som ger nya perspektiv. Svårigheterna handlar inte enbart om att analysera de konnektionistiska programmens kapacitet. Det är lika svårt att veta vilken kapacitet som behövs för att människan skall kunna skapa ny kunskap.

Collins kritik bygger på att han anklagar Dreyfus för att göra ett **professionellt misstag**, en **filosofisk utelämnning** och ett **sociologiskt felslut**:

Det **professionella** misstaget handlar om att Dreyfus använder filosofiska ståndpunkter selektivt. Speciellt menar Collins att Dreyfus inte inser att en wittgensteinsk analys lika mycket gäller för vetenskap, teknologi och matematik som för andra kulturella områden. På så sätt skapar Dreyfus en falsk skillnad mellan formella kunskapsområden inom vilka datorerna är kapabla och informella kunskapsområden där de inte har någon kraft (Collins,1996:100). Collins argumentation för påståendet ovan bör innebära att det finns mycket mer som är formaliserat än vad Dreyfus anser. Det blir inte riktigt klart till vad Collins knyter denna formalisering. Det förefaller ändå som att han använder Wittgensteins analys av språkregler. Det finns visserligen flera möjligheter att tolka Wittgensteins språkregler. Här ser det ut som att Collins tar fasta på att dessa regler är skapade i social interaktion mellan människor vilket gör att endast människor kan uppfatta dem. Det betyder då att väldigt många kunskapsområden är formaliserade av dessa språkregler som endast människor kan uppfatta. Det betyder i sin tur att om datorer kräver formalisering för att kunna fungera så handlar det knappast om samma kapacitet för formalisering som människor har. Det är alltså ett argument

för att den formalisering som är inblandad i de sociala processer som har sitt ursprung i människors handlande inte kan åstadkommas på något artificiellt sätt.

Utifrån Dreyfus ståndpunkt finns det ändå möjlighet att hävda att de mänskliga insatserna handlar om att försöka formalisera de intryck människorna möter när de exponeras för allt informellt i världen. Då skulle den sociala formalisering som Collins talar om istället handla om det sätt som människorna använder för att kunna kontrollera kunskapen. Denna sociala formalisering skulle då snarast ses som det som förminskar möjligheterna av att kunna använda den informella världen för att kunna se något nytt. På så sätt kan vi argumentera för att människan också kräver formalisering. Vi bör då också observera att vi i inledningskapitlet såg att den AI-positiva sociologins hållning menar att det är sociala processer som gör att människan kan komma längre än hon kan med de begränsningar som finns i hennes eget medvetande. Hon behöver alltså något mer än sitt eget medvetande för att inte bli inestängd i de möjligheter som redan är formaliserade.

Språkregler kan vara något som formaliserar världen. Vi bör uppmärksamma att bl.a. Bloor menar att Wittgensteins uppfattning om språkregler skall tolkas som att de är formade av kausala samband.¹⁰ Men man borde också kunna hävda att språkregler kan vara de artefakter vilka vi kan använda för att skapa de positioner vi behöver för att nå längre. De språkregler som är formade av kausala samband tycks vara statiska regler och påminner därför om de som symbolisk AI använder, medan det också måste finnas en möjlighet att uppfatta språkregler som en form av heuristik. Vi kommer senare att se att Vygotskys diskussion om begrepp antyder att deras utveckling skapar redskap med vilka vi kan utvidga medvetandets kapaciteter. Collins tycks blanda ihop dessa bägge möjligheter att se på språkreglerna när han värderar deras betydelse för att nå det icke-formella. Det kan vara heuristiska språkregler som ger människan kapacitet att reflektera över möjligheterna i det icke-formaliserade. Det viktiga är att det då inte är helt klart om språkreglerna bara kan skapas av människor eller om de kan skapa nya positioner på egen hand. På det senare bygger ju konnektionismen. Det kan vara så att det behövs en viss autonomi i språkreglerna för att de skall kunna visa på mer möjligheter än vad människorna själva kan se. För att kunna göra nya kunskapskopplingar behöver vi språkregler som förmår referera till möjligheter istället för språkregler som talar om var gränserna går.

Den stora fråga kring AIs kapacitet handlar då **inte** om att människorna skapar en social formalism. Istället handlar den om i vilken mån det behövs människor för att exponera de möjligheter som finns att se något nytt i en kaotisk värld. Det är naturligtvis troligt att andra människor är en viktig orsak till att en individ kommer i en position som tillåter henne att se något nytt. Men om det inte handlar om att skapa de regler som bara människor kan uppfatta, är det faktiskt

¹⁰ Se till exempel Bloor, 1992

inte så provocativt att påstå att även datasystem har kapacitet att sätta människor i nya positioner. Poängen är att det som en människa gör och som inspirerar en annan människa, varken handlar om en särskilt medveten handling, eller handlar om att skapa några svårförståeliga regler. Det handlar istället om en ganska enkel ompositionerande aktivitet. De regler som styr denna sociala process är ganska enkla eftersom det inte handlar om tänkande.

Vi bör framhålla att Bloors uppfattning om språkregler som resultat av kausala processer ligger rätt nära den position som den symboliska AI:n arbetar med. Collins diskussion blir oklar eftersom han inte skiljer kausala regler från funktionella regler. De regler som krävs för att få konnektionistiska system att fungera i en DAI-struktur är funktionella regler. Därmed bör de vara mindre sårbara för Collins kritik.

Den **filosofiska** utelämnningen gäller ett argument mot AI som tar sikte på att mycket av människans kunskapsskapande kräver kapaciteter som är förkroppsligade i människan själv. Det blir lite oklart varför Collins riktar denna kritik mot Dreyfus men det bör handla om att Dreyfus skulle bygga sitt argument på att det går att bygga in förkroppsligad kunskap i konnektionistiska system.

While the physical embodiment of entities may give rise to their collective "form of life", this does not mean that every individual that is to share that form of life must have the same type of physical body. Immobile cuboid boxes in general may never develop human-like ways of being; but this argument is not enough to show that an individual immobile cuboid box could not. The elision leads Dreyfus to overemphasise questions to do with the detailed architecture and physical form of computers and underemphasise the problem of what it is to be a member of a social collectivity (Collins,1996:100).

Collins tycks mena att Dreyfus blandar ihop egenskaper hos människor med egenskaper hos sociala grupper genom att Dreyfus på något sätt tillskriver komplexiteten människor och inte sociala grupper. Det ser ut som att Collins menar att frågan inte handlar om det vi har kallat *embodiement* och istället handlar om att *environment* alltid måste ses som en kunskap som finns i ett socialt kollektiv. Han tycks hävda att den sociala kapaciteten inte har att göra med någon konstitution i den mänskliga kroppen utan med egenskaper i människornas sätt att interagera med varandra. Detta är i överensstämmelse med att han menar att språkregler är kollektiva fenomen som bara människor kan skapa och använda. Detta bör emellertid bygga på en missuppfattning av Dreyfus position, eftersom han knyter det icke-formella till det som finns i den miljö människan möter. Och att det är den miljön som gör att människan når det komplexa.

Om de egenskaper som gör människan annorlunda inte finns i människan själv utan i de sociala liv vi lever bör det ändå finnas anledning att hävda att

människans förmåga till intelligens är mer beroende av kulturella artefakter som kan ha en egen omstrukturerande kapacitet oberoende av människornas kapaciteter. Frågan om de konnektionistiska systemens kapacitet bör då kunna formuleras i termer av om de kan skapa mening eller om de bara producerar obegripligt brus. Frågan bör då handla om vad det är som gör att något blir till en artefakt. Vad är det som gör att en samling avtryck i miljön plötsligt framstår som något som kan påverka oss? Enligt de konnektionistiska modellerna har det att göra med att dessa avtryck positionerar sig på nya sätt genom att de inbördes ändrar sina relationer. Den viktiga frågan blir då om denna positionering kan ske utan att det är människorna som skapar relationerna? Eller så kan vi formulera frågan i konträra termer. *Kan människorna se nya relationer i omvärlden utan att omvärlden på egen hand har positionerat sig så att vi människor har fått nya positioner från vilka vi kan observera?* Collins ståndpunkt är mycket motsägelsefull. Den betonar att det finns en socialitet som har att göra med mänsklig interaktion, men den har samtidigt inte att göra med egenskaper i människans kropp.

Det **sociologiska felslutet** handlar om att Dreyfus inte inser att mycket av den uppenbara kapaciteten hos datorerna egentligen består av en slags interaktion i realtid med dem som använder dem (Collins,1996:100).

This means that even if computers are not proper members of social collectivities, they can seem to act like humans with the (usually unnoticed) help of their users. Often, the results are excellent. Thus, some philosophically and psychologically uninteresting computers can do sociologically interesting and useful things in unanticipated ways (Collins,1996:100).

Det skulle då betyda att artefakter bara kan vara redskap för människorna. Det är en logisk slutsats av Collins tidigare argument men vi bör uppmärksamma att det förstärker uppfattningen om att kulturella strukturer inte kan påverka människan av egen kraft. På så sätt blir det svårt att se hur mänsklig intelligens kan utvecklas på något annat sätt än genom biologisk utveckling. Det sociala blir då också något som blir beroende av biologisk utveckling. Det låter som en alldeles för radikal position. För att komma närmare socialitetens grunder bör vi se närmare på hur Darwins utvecklingsbiologi har använts för att förklara det socialas utveckling.

4.5 Människans ursprung enligt utvecklingsbiologin

Ett skäl till varför många sociologer ser en inbyggd motsättning mellan AI och sociala processer kan ha att göra med att användningen av begreppet social har

blivit så trivialiserad. Det sociala har blivit ett begrepp för allt som människorna gör tillsammans. På så sätt har man missat att begreppet i sig innehåller många olika dimensioner. Dimensioner som gör att en analys av sociala processer inte kan göras utifrån en och samma mall. Det finns många begrepp i denna avhandling som har en mångtydig innebörd. Vi får leva med att allt inte går att definiera om inte texten skall bli olidligt detaljerad. Antagligen går det inte heller att ge entydiga definitioner av hur begreppet social används i den sociologiska diskussionen. Vad som är viktigt är att vi inte använder begreppet på ett sådant sätt att det blir svårt att diskutera ett sådant fenomen som artificiell intelligens. Bland andra lyckades Vygotsky förändra synen på sociala processer i lärandet genom att fundera på människans ontologiska utgångspunkter utifrån den darwinistiska utvecklingsbiologin. Vi skall därför fördjupa oss i vad modern darwinistisk utvecklingsbiologi kan säga oss om hur vi bör se på det sociala. Det är också en diskussion som hjälper oss att se hur omgivningen påverkar den sociala sfären.

Vygotsky använde darwinistisk teori huvudsakligen för att visa på det rimliga i tanken på att det unikt mänskliga skapades genom att människans kulturella struktur utvecklades. Genom att språket och begreppen utvecklades kunde människan skilja sig från djuren och bli en kulturvarelse. När darwinistiska argument återigen kopplades till teorier om människans sociala roll var syftet det motsatta. Den så kallade sociobiologin som lanserades av Edward Wilson i boken *Sociobiology* (Wilson, 1975), som första gången utgavs 1975, försökte knyta det sociala till någon form av genetiskt arv. På så vis var det en inriktning som försökte nedtona kulturens betydelse och som försökte hävda att människan trots allt var ett djur. I den teorin blev det sociala något som borde knytas till beteenden som var nedärvda i människans biologiska kropp. Den positionen mötte stark kritik, inte minst från sociologer. Det är därför som det är så förbryllande att sociologer i förhållande till AI-frågan, så ofta bygger sina argument på att det sociala någonstans måste knytas till egenskaper i människans kropp. Det naturliga vore att koppla det till egenskaper som har uppstått genom den kulturella utvecklingen, eller som hör samman med att de kulturella redskapen ger henne möjligheter att bete sig mer intelligent än djuren.

Det är också så att den sociobiologiska litteratur som utgivits efter Wilson mer tar sikte på det som darwinismen säger om miljöns roll för att utveckla människan än på resultatet av denna utvecklingsprocess. Vi kan påminna att Darwin talar om att de mest anpassningsbara överlever och inte de starkaste. De starkaste bör däremot sträva efter att kontrollera miljön så att den är gynnsam för deras maktpositioner. Vi kan också hävda att om miljön hela tiden förändrar människan finns en viss grund för att också tala om en naturvetenskaplig relativism.

Den mest kända efterföljaren till Darwin idag är troligen Richard Dawkins. Han hävdar i boken *Den själviska genen* (Dawkins, 1992) att genen är den enhet som utvecklas. Och genen är alltid självisk på så sätt att den strävar efter sin egen

överlevnad. Dawkins argumenterar för att det naturliga urvalet och dess evolutionära konsekvenser blir begripligare om de enskilda generna, och inte enbart individerna, uppfattas som de enheter mellan vilka urvalet vrakar. Dawkins menar att utvecklingen bygger på kumulativa processer. I *Den Blinde Urmakaren* (Dawkins, 1988) hävdar han att det naturliga urvalet är en nödvändig och tillräcklig förutsättning för organismernas anpassning och utveckling och därmed för jordens biologiska mångfald. På så vis behöver inte en egenskap som intelligens vara gudagiven. Vi bör observera att detta är en förklaring som går via biologisk utveckling. Den visar dock att intelligens kan förklaras med matematiska modeller.

Den kritik som den amerikanske paleontologen Stephen Jay Gould har riktat mot Dawkins handlar framför allt om att han menar att det som utvecklas är enskilda organismer, i frågan om människans utveckling alltid enskilda individer. Dessutom menar han att bland annat hans egen forskning har visat att utvecklingen inte är kumulativ. Istället har han tillsammans med Nils Eldredge lanserat begreppet punkterad jämvikt som står för att utvecklingen är episodisk (Gould&Eldredge, 1977). Korta perioder av utveckling och specificering av nya arter följs av mycket längre perioder av stabilitet, där istället många arter försvinner utan att utvecklas till nya och mer effektiva arter. Framför allt baserar sig Goulds ståndpunkter på den analys av fossilavlagringar i en bergsformation i British Columbia i Kanada som kallas *Burgess Shale* (Gould, 1989). Det skulle då öppna för att det är egenskaper i miljön under en generations livstid som gör att grunden sås till att människans kapaciteter förändras. Vilket skulle öppna för att den moderna datatekniken verkligen kan förändra människans kapacitet i ett kortare perspektiv. Det framstår som oklart om Gould menar att dessa kontextegenskaper förändrar människans kapacitet genast eller om vi ändå behöver vänta på att detta frö till förändring skall utveckla egenskaper efter en lång biologisk process. Går det genast kan vi hävda att AI genast ger människan en större kapacitet till intelligens. Dock kan Goulds variant knappast användas till mer än att förklara varför människornas medvetanden blir mer kapabla just därför att de använder redskap. Den kan inte i sig användas för att förklara varför stark AI kan fungera eftersom den bör kunna utveckla sig utan att göra detta med människans hjälp. Människans påverkan på AI-systemen kan bara förklaras med att hon fungerar som miljö för dessa system. Människans påverkan kan på så sätt bara vara avledd.

Skall vi tala om stark AI har den frågan egentligen inte att göra med biologisk utveckling eftersom det handlar om ett tekniskt redskap. Men den har ändå likheter med Dawkins syn på utvecklingsprocessen eftersom vi får anta att stark AI måste utvecklas enligt statistiska principer och att utvecklingen har att göra med en kumulativ process. Skillnaden gentemot den biologiska utvecklingen är bara att vi kan räkna med att programmets utveckling går mycket fortare därför att en generation i nodens utveckling kan gå blixtnabbt medan den biologiska utvecklingen alltid är beroende av att mätas med den tid som går åt för en mänsklig generation.

Filosofen Daniel Dennett har i boken *Darwin's Dangerous Idea* (Dennett,1995) kraftigt försvarat Dawkins utvecklingstänkande. Framför allt försvarar Dennett Dawkins därför att han menar att man måste konstruera utvecklingsteorin enligt Dawkins modell om man skall kunna undgå förklaringar som kräver någon form av allsmäktig skapare, det Dennett kallar en skyhook.

Dennett påpekar att Dawkins modell bara kräver att vi kan hitta mekanismer, Dennett kallar dem för *Cranes*, som kan lyfta upp arternas komplexitet till nya nivåer. Stark AI bör också lita till mekanismer istället för till någon skyhook, ty i det senare fallet måste människan spela den rollen vilket skulle vara ologiskt.

Dennett framhåller också en annan idé som Dawkins lanserade i *Den själviska genen*. Det är begreppet *meme* som är en motsvarighet till generna men som driver den kulturella utvecklingen framåt. Memen är en slags självpropagerande enhet som kan förklara mänskligt beteende och kulturell utveckling.

Som exempel på memers kan nämnas melodier, idéer, slagord, klädmoden, sätt att tillverka lerkrukor eller slå valv. Precis som generna fortplantar sig i genpoolen genom att hoppa från den ena kroppen till den andra via spermier och ägg, fortplantar memerna sig i mempoolen genom att hoppa från den ena hjärnan till den andra via en metod, som i vid mening kan kallas imitation. Om en vetenskapsman hör talas om eller läser om en god idé, för han den vidare till sina kolleger och elever. Han tar upp den i sina artiklar och föreläsningar. Om idén då slår an kan den sägas fortplanta sig själv genom att den sprider sig från hjärna till hjärna. Som min kollega N. K. Humphrey så elegant summerade ett tidigare utkast till detta kapitel: ". . . memers bör betraktas som levande strukturer, inte bara metaforiskt utan också rent sakligt. När du planterar in en fruktbar mem i mitt medvetande parasiterar du bokstavligt talat på min hjärna och förvandlar den till ett redskap för memens fortplantning, på samma sätt som ett virus kan parasitera på värdcellens genetiska mekanismer. Och detta är inte bara ett sätt att uttrycka sig - memen för till exempel 'tro på ett liv efter döden' förverkligas ju faktiskt fysiskt, miljontals gånger, som en struktur i enskilda människors nervsystem världen över (Dawkins,1992:221-222).

Frågan blir då om just memen är den enhet som kan förklara att stark AI kan utvecklas?

I boken *Den felmätta människan* (Gould,1983) går Gould till angrepp mot dem som hävdar att intelligens är en genetisk egenskap som kan mätas med så kallade IQ-tester. Istället menar han att också intelligens är något som utvecklas och är beroende av ontogenesen. Problemet med Goulds tes är att det blir oklart hur samspelet med den kulturella strukturen gör att människan utvecklas. Talet om

ontogenesens betydelse förefaller öppna för att redskapen påverkar människornas intelligens men det är inte klart om detta kan ske bara genom att människan får tillgång till redskap. Goulds uppfattning om att utvecklingen inte är kumulativ och att det istället förekommer perioder av evolutionära språng har han gemensam med Luhmann (Luhmann,1997:505). Sverre Moe menar att vi idag måste fundera på om elektronisk kommunikation och Internet representerar evolutionära språng (Moe,2003).

Dawkins uppfattning om *memer* är då som mer givande om vi vill förklara människans kulturella utveckling. I sammanhanget bör vi uppmärksamma att Dawkins är rätt så kritisk mot den tidigare variant av sociobiologin som Wilson försökte lansera. Det är något som också betonas av Ullica Segerstråle som framhåller att det kapitel som i Wilsons *Sociobiology* försöker dra paralleller mellan djurens beteenden och människans, har en helt annan framtoning än det kapitel i *Den själviska genen* som behandlar *memerna*. Medan Wilson försökte förlänga sin teori om gener till samhället så menar Dawkins att kulturen bildar en skild storhet (Segerstråle,2000:72). Det understryker att Dawkins mer ser kulturen som en miljöfaktor. Problemet är väl bara att Dawkins teori om genens utveckling och hans tankar om *memer* inte riktigt hör samman med varandra.

Det problematiska med Dawkins uppfattning handlar om att teorin om genens utveckling är utpräglad reduktionistisk och bygger på en realistisk ontologi. Både Dawkins och Luhmann kan ses som reduktionister i det att de försöker knyta tillbaks mänskliga och sociala processer till en enda förklaringsmodell. Vi bör framhålla att det är stor skillnad på en reduktionism som bygger på realism och en som bygger på relativism. Den första reduktionismen, som hos Dawkins, bygger på att en viss enhet fungerar på ett visst sätt enligt en naturlag. Då får Dawkins svårt att förklara hur *memen* kan vara en självpropagerande enhet. Den relativistiska reduktionism som Luhmann utgår från, söker istället utvecklingsmotorn i den funktionalitet som uppkommer därför att världen hela tiden förändras. Det är en reduktionism som skapar variation istället för att sträva efter en grundläggande statiskhet. Om de kulturella strukturerna, t.ex. i form av *memer*, skall kunna ompositionera sig bör de bygga på relativism. Den luhmannska teorin behöver därmed inget annat än en god funktionell mekanik för att fungera. Detta innebär att teorin inte behöver något mera än den byggkran som systemet i sig utgör. Därför är det bättre att grunda en teori om kulturell utveckling på Luhmanns kommunikationsteori än på de ståndpunkter som Dawkins erbjuder. En förklaring till de sociala processernas påverkan på kulturell utveckling enligt både Gould och Dawkins, tvingar oss tillbaks till att det sociala handlar om biologiskt utvecklade och nedärvda egenskaper. Den sociologiska AI-kritik som bygger på att det är något unikt med den mänskliga kroppen, tvingas på så sätt också tillbaks till att det mänskliga måste förklaras med en lång biologisk utveckling. Och då är det svårt att hävda att den kapacitet som nya redskap ger genast kan få en betydelse genom att de förändrar kroppens kapaciteter. Det är mer givande om vi som Vygotsky knyter

ontogenesens utveckling till redskapens betydelse. Detta skapar en större öppning för att AI kan vara stark.

I detta kapitel har vi visat på att sociologin ändå har visat ett ganska stort intresse för frågan om stark AI. Även om det finns anledning att hävda att de flesta sociologer är skeptiska till möjligheterna av en genuint stark AI har vi visat på att det finns sociologer som inte bara är positiva till stark AI utan som också argumenterar för att stark AI kan vara en bra modell för att exemplifiera hur sociala processer kan se ut.

Vi har tagit del av en del av den sociologi som angriper stark AI. Vi har visat på att den sociologin tar miste i sin kritik därför att den överdriver kroppens betydelse i sociala sammanhang. Därmed missar den att det sociala snarast finns i en social och kulturell kropp – den överdriver betydelsen av *embodiement* och missar det som vi har kallat *environment*.

På så sätt hamnar denna sociologi snett i diskussionen kring utvecklingsbiologins betydelse. Genom att betona det kroppsliga tror den att människans kapaciteter bara kan utvecklas genom en lång biologisk utvecklingsprocess och missar därmed att den sociala människan kan kompensera sina kroppsliga brister genom att bygga upp en social och humanistisk kulturell struktur. Denna sociologi missar på så sätt kärnan i den egna disciplinen.

5. Luhmanns socialitet och Distribuerad AI

5.1 Second Order Cybernetics, relativismen och det posthumana

Cybernetiken brukar av många ses som ett program som studerar hur den faktiska världen enligt en mekanistisk lagbundenhet skapar den funktionalitet som driver samhällets processer. Det skulle alltså vara raka motsatsen till sociologins uppfattning om att kunskap skapas som bearbetning av det som inte är formaliserat. Som vi skall se har utvecklingen inom cybernetiken gått från att se information som ett värde i sig till att information bara avspeglar föränderliga positioner i en relativ värld. Det antagande om lagbundenhet i världen, som i början låg till grund för cybernetiken, handlade om att det skulle finnas en verklig eller reell ontologisk nivå, medan den moderna cybernetiken oftast ser på världen som relativ.

Den luhmannska systemteorin bygger också på den utveckling som cybernetiken genomgick under efterkrigstiden. Kopplingen till cybernetiska modeller tycks vara det som framkallar en uppfattning om att också Luhmanns systemteori står för en mekanistisk syn på samhällsprocesser. Nedan skall vi se att cybernetikens utveckling mot den så kallade *andra ordningens cybernetik* (*Second Order Cybernetics*) har gjort att den distanserat sig långt från den tidigare cybernetikens mekanistiska anstrykning.

Den bästa beskrivningen av hur denna *Second Order Cybernetics* har vuxit fram, finns i den amerikanske litteraturvetarprofessorn Katherine N. Hayles bok *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*.

Hayles analyserar relationen mellan människorna och deras redskap. Avsikten är att undersöka hur cybernetiken och speciellt informationsteknologin modifierar människans syn på sig själv och på sin relation till maskinerna. Speciellt studerar Hayles hur informationsbegreppet förändrats, hur idéer om cyborgerna som en teknologisk artefakt och kulturell ikon konstruerades i perioden efter andra

världskriget. Och hur detta ledde till att människan uppfattades som en posthumanistisk varelse. Hayles menar att människans umgänge med litteraturen sedan länge har gjort henne posthumanistisk. Vilket betyder att hon måste vara något mer än enbart ett tänkande subjekt.

Hayles framställning bygger på att redovisa hur ett antal frågor behandlades under de så kallade Macykonferenserna. De ordnades efter andra världskriget för att diskutera cybernetikens möjligheter. En av de stora frågorna gällde hur man skulle uppfatta informationsbegreppet.

Hayles menar att frågan rörde tre fronter: Den första handlade om konstruktionen av information som en teoretisk enhet. Den andra om hur man konstruerar (mänskliga) neurala strukturer så att de ser ut som ett flöde av information. Den tredje om hur konstruktionen av artefakter transformerade informationsflödena så att de kunde bli observerbara och på så sätt framstå som verkliga (Hayles,1999:50).

Hayles skriver att den linje som först vann, ansåg att information stod för ett eget värde som kunde kopplas loss från människan. Motstånd mot denna syn går ut på att det är en skillnad på vad som kan innehållas i detta informationsbegrepp och de komplexa problem som cybernetiken försökte greppa. Hayles menar att uppfattningen om att information är ett värde i sig, som är mer väsentligt än materia eller energi, vann därför att det i den kulturella situation som rådde med tro på (positivistisk) vetenskap inte fanns något bättre alternativ (Hayles,1999:50).

Inom den interna debatten mellan cybernetiker började så småningom åsikter som hävdade att information inte kunde vara ett värde i sig göra sig starkare. Två rivaliserande skolor började träda fram. En som hävdade homeostasis och en annan som betonade reflexivitet. Den förra var väl förenlig med uppfattningen om att information är en egen storhet. Den andra tog fasta på att se information som en process, vilket kopplade frågan till hur reflexivitet uppstår eftersom det då också gällde att uppmärksamma effekten på mottagaren. Denna inriktning förlorade eftersom den ansågs omöjlig att hantera, då den krävde att man skulle kunna analysera den kontext som bedömdes svår att specificera och avgränsa. Den förra påstods däremot vara mycket lättare att hantera i kvantitativa sammanhang (Hayles,1999:57).

För att vi skall kunna anknyta detta till resonemanget om AI behöver vi komma med några klarlägganden: För det **första** innebär detta att cybernetiken under denna tid hade samma hållning till informationsbegreppet som de teorier som menar att kunskap skapas av människan genom att bearbeta objektiv information. Eftersom den symboliska AI:n försöker skapa artificiella aktörer som skall bearbeta information på samma sätt som människan, kan man hävda att både denna symboliska AI och de som hävdar betydelsen av den rationella människans tankeprocesser, ligger väldigt nära det informationsbegrepp som cybernetikerna under denna tid förfäktade. De finns dock en viss skillnad på hur cybernetiken och rationalismen ser på kunskapsprocesserna.

Som vi redan varit inne på menar rationalismen att det är genom en rationell analys av det givna innehållet som information kan omvandlas till kunskap. Den symboliska AIn, som är en del av rationalistiska ansatser, skiljer sig från den traditionella rationalismen genom att den mer betonar att villkoren för den rationella analysen bestäms av lagbundenheter i världen. Därför tonar man ner medvetandets analytiska förmåga och hävdar att datorernas insats mest handlar om att registrera den lagbundna informationen och om att behandla den med redan kända lagar. De rationalister som är kritiska mot stark AI betonar däremot att den lagbundna informationen är så svår att tolka och upptäcka att det krävs mänskliga hjärnor. Bägge varianterna av rationalism har gemensamt att den kunskap man kan komma fram till är begränsad av vad sanningen tillåter. Att vetenskapen går framåt har bara att göra med att det finns så många felaktigheter i själva processen att människan sällan når fram till sanningen. Det är befogat att hävda att de rationalister som argumenterar för symbolisk AI, uppfattar att datorerna är mer tillförlitliga än människor och att AI därför lättare kan undvika felaktigheterna i processen.

Cybernetiken under denna första tid liknar den symboliska AIn, men den betonar i ännu högre utsträckning att världen är lagbunden och mekanistisk. Detta gör att man kan tona ner medvetandets roll, vilket i sin tur gör att denna cybernetik har ännu lättare att ta till sig att mänskliga handlingar kan mekaniseras.

För det **andra** så innebär de ståndpunkter som framläggs av den konstruktivistiska kunskapssociologin att man i första hand ifrågasätter möjligheterna till objektiv information. Kritiken handlar således inte om vilken kapacitet för rationell analys som människans medvetande har. Vi har visserligen sett att man ofta anser att informationens subjektiva värde är så knutet till att det är människor som skapat den att det också bara är människor som kan förstå den. Det är skälet till att man oftast är kritisk till stark AI. Här räcker det att visa på att kunskapssociologins inställning till information liknar den kritik som framlades internt inom kretsen av cybernetiker mot det informationsbegrepp som till en början vann. Det är viktigt att betona att denna nya inriktning på cybernetiken, *den andra ordningens cybernetik*, tog avstånd från uppfattningen om världens lagbundenhet. Detta innebär att utvecklingen inom cybernetiken innebar ett förhållningssätt till det mänskliga medvetandets kapacitet som är mer förenlig med den kritik mot informationsbegreppet som den konstruktivistiska kunskapssociologin uppvisar.

Hayles menar också att skillnaden mellan att uppfatta information som ett värde och som en kontext återkom i diskussionen om vilken bild av människan som skulle styra designen av cybernetiska system. Diskussionen handlade om huruvida människan skulle uppfattas som en informationshanterande maskin eller som en del i en kontext (Hayles, 1999:57).

Så småningom blev anomalin i den första modellen alltför uppenbar. Det utlösande var att man ifrågasatte om människan kunde stå i mitten av modellen. Vi

bör uppmärksamma att en sådan position också ifrågasätter om det är med människan vi skall jämföra den starka AI:n. Den som först ifrågasatte den regerande modellen var en man vid namn Frank Fremont-Smith. Han framhöll att en annan människa aldrig kan stå i mitten av modellen eftersom den som studerar en människa förändrar dennes relation till maskinerna (Hayles,1999:68).

Den andra ordningens cybernetik uppstod ur denna övergång till att betona reflexivitet. De centrala personerna var matematikern Heinz von Foerster samt antropologerna Margaret Mead och Gregory Bateson. De försökte lämna frågan huruvida medvetandet måste vara förkroppsligat i en människa, och valde istället att fokusera på observatören och det system i vilken observatören befann sig. Den förskjutningen av fokus är betydelsefull för DAI.

Det som bildar ett steg mellan *Second Order Cybernetics* och den moderna sociologiska systemteorin är teorin om *själstyrande system*, eller med det grekiska ord som används *autopoietiska system*. Teorin är grundad av chilensen Humberto Maturana som var en av deltagarna i de senare Macy konferenserna. Teorin är utvecklad av Maturana tillsammans med en annan chilensare, Francisco Varela (Maturana,&Varela,1987), (Maturana,&Varela,1980). Varela har på senare tid rätt kraftigt omformulerat sina positioner på ett sätt som innebär ett brott mot den sociologiska systemteorin. Den nya positionen går tillbaks till att de icke-rationella komponenterna i kunskapsprocessen finns nedlagda i kroppen (Varela et.al.,1991). Han tar då ställning i den debatt som handlar om huruvida det förkroppsligade finns i mänskliga kroppar eller i den kulturella kropp som den humanistiska civilisationen har skapat.

Hayles visar att Maturana skiljer mellan allopoietiska system och autopoietiska system. Hayles menar att studiet av människan som ett autopoietiskt system är inriktat på att se hur individen har tillgång till de kvaliteter som tillåter henne att leva fritt och utveckla sin individualitet. Studiet av allopoietiska system handlar däremot om att se hur de styrs av mål som är instrumentella och nedlagda i systemet utifrån. Hayles menar till exempel att när hon kör sin bil är dess funktioner underordnade de mål hon har. Istället för att kolvarna i motorn använder energin till att reparera sig själva så använder de kraften till att driva fram bilen så att hon kan komma till sin affär. Hon fungerar då autopoietiskt men hennes bil fungerar allopoietiskt (Hayles,1999:141). Också Dirk Baecker framhåller samma synpunkt i en artikel. Han framställer det som att medan den klassiska maskinen känns igen på att dess delar gör ett arbete, processar den transklassiska maskinen information. Denna maskin styrs av cybernetiska principer på ett kritiskt sätt(Baecker,2006:38). Hayles framhåller att detta ger Maturana en möjlighet att tala om den maktkamp som finns inom samhället. Det samhälle som tvingar människan att leva som en maskin som genom teknisk determination fungerar allopoietiskt.

Humanistiska vetenskaper tycks ha mycket lättare än samhällsvetenskapen att acceptera att kulturella strukturer har en autonom funktionalitet. Framför allt

litteraturvetenskap, men också genusforskning har använt sig av cybernetiska resonemang för att belysa detta. Inom genusvetenskap är det så kallad queer-teori som har använt sig av sådana resonemang. Till exempel har Donna Haraway diskuterat cyborgens som en metafor för den könlösa och teknikberoende människan (Haraway, 1991). Också litteraturvetare har använt posthumanistisk teori för att visa på texternas autonoma funktionalitet. Jonas Ingvarsson som för något år sedan lade fram en litteraturvetenskaplig avhandling om posthumanistiska inslag i svensk skönlitteratur (Ingvarsson, 2003) skriver att författare som Åke Hodell och Harry Martinsson varit påverkade av cybernetiska teorier. Martinssons Aniara brukar ju ses som en stark kritik av det teknologiserade samhället. Det finns förstås mycket i dikten som kan tolkas i den riktningen. Men där finns också ett starkt drag av att det är människorna som använder tekniken för sina egna syften. Det är också miman, en slags superdator, som utsläcker sitt eget liv som en reaktion på det människorna har ställt till med på jorden!

När Dorisburg smält bort led miman flera

Dar.

Hon bad mig säga Ledningen att hon

sen någon tid var lika samvetsöm

som stenarna. Hon hade hört dem ropa

på stenars vis i Doris fjärran dal.

Hon hade sett granitens vita gråt

när sten och malm förgasas till ett dis.

Hon hade rörts av dessa stenars kval.

Förmörkad i sitt cellverk av den hårdhet

som människan visar i sin ondskas tid

kom hon som länge väntat var därhän

att hon på mimors sätt till slut bröts ner.

Indifferent tredje vebens tacis

ser tusen ting som inget öga ser.

Nu ville hon i tingets namn ha frid.

Nu ville hon ej förevisa mer.

De sista ord hon sände var en hälsning

från en som nämnde sig Den

söndersprängde.

Hon lät Den söndersprängde själv få vittna

och stammande och söndersprängt berätta

hur svårt det alltid är att söndersprängas,

hur tiden rusar till för att förlängas.

(Martinsson, 1997:52-54)

5.2 Luhmanns systemteori och kunskapsprocesserna

Luhmanns produktion är rätt omfattande och framför allt svår att tränga igenom därför att han hela tiden definierar om en mängd begrepp samt ser på sociala fenomen på ett sätt som är alldeles annorlunda. Ändå kan man hävda att den gör sociologin enklare eftersom den inte i alla lägen behöver knytas till djupsinniga uppfattningar om människans natur. Den luhmannska sociologin fokuserar på processer externa från människan och på egenskaper i kontexterna, vilket gör den mer analytiskt givande för att studera sociala processer. Denna sociologi har därför gett mig ett förklarings skimmer över vad sociologi egentligen innebär även om mycket av teorin handlar om utveckling av det som också finns i äldre sociologi. Det är för övrigt rätt ont om verk i den svenska sociologin som behandlar den moderna systemteorin. I slutet av 90-talet kom en avhandling på Sociologiska Institutionen i Lund om Luhmanns teori, skriven av Jan Inge Jönhill med titeln *Samhället som system och dess ekologiska omvärld*. Jönhill presenterar Luhmanns teori ända fram till hur den ser ut i det stora avslutande opuset *Die Gesellschaft der Gesellschaft*, som Luhmann han skriva klart innan han dog 1998.

5.2.1 Distinktionen

Jönhill framhåller att den mest grundläggande kognitiva operationen hos Luhmann är att göra distinktioner. Vi bör uppmärksamma att Luhmanns syn på kognitiva operationer gör att det kognitiva inte måste handla om individernas tankeprocesser. Luhmann menar istället att dessa kognitiva processer är knutna till social kommunikation. Därför finns det fog för att hävda att den kognitiva operationen, som vi vanligen ser som en process där individen med sitt medvetande bearbetar information, istället handlar om en social funktionalitet.

Grunden till den annorlunda synen på den kognitiva processen är att Luhmann menar att människan måste uppleva en skillnad för att kunna inhämta och bearbeta information. Det gör att det finns en starkare betoning av att medvetandet blir beroende av externa faktorer. Dessutom betecknar distinktionen hos Luhmann en form, vilket gör att den inte bara behöver vara något abstrakt. Denna teori om formens betydelse har ursprungligen formulerats av en matematiker och filosof vid namn George Spencer Brown i boken *Laws of form*. Spencer Brown försöker visa hur formen kan användas som den grundläggande enheten i meningsskapandet. Spencer Brown bygger upp sin argumentation genom en teori om hur skillnader skapar olika former. Grunden bygger på att vi inte kan indikera något utan att göra en distinktion. Därför är distinktionen den grundläggande formen. Definitionen av en distinktion är att den skall stå för perfekt avgränsning (Spencer-Brown, 1979:1-2).

Naturligtvis kan man tolka detta som att det är individen som måste göra distinktionen och att den bara blir en tankekonstruktion. Det går inte att komma ifrån att distinktionen ändå har att göra med egenskaper i kontexten. Dirk Baecker förklarar att Spencer Browns tankar om distinktioner innebär att man inte kan konstruera något eller ens göra något utan att först göra en distinktion. Därför innehåller distinktionen motivet, innehållet och värdet i sig självt (Baecker,1999:2-3).

Diskussionen pekar på att de processer som sker i medvetandet inte alltid bygger på komplex funktionalitet. Att medvetandet inte har möjlighet att använda annat än en basal funktionalitet understryks av att Baecker skriver att det man kan göra i en sådan här meningssökande aktivitet är begränsat till två alternativ. Antingen kan man göra samma distinktion igen eller så kan man försöka observera vad som inte är indikerat och därmed gå utöver distinktionen. Det viktiga är att bägge tillvägagångssätten måste ses som en operation och inte som ett tillstånd eller ett ting. Det man kan göra är således att antingen följa en distinktion eller att ta avstånd från den (Baecker,1999:3).

Detta visar på hur mycket mekaniska element det faktiskt finns i våra försök att upptäcka världen. Vi kan inte själva tillföra motiv, mening och värde i det vi möter. Hittar vi ingen mening i det vi möter kan vi bara välja att avfärda distinktionen. Om vi skall hitta ny mening måste vi sätta oss i situationer som låter oss se meningsfulla distinktioner. Operationerna blir alltid lika om vi inte hamnar i situationer som gör att vi kan exponeras för nya möjligheter. Visserligen måste vi välja att se dessa nya situationer vilket kan upplevas som att vi tänker. Det viktiga är ändå att vi också tillåts att se de nya möjligheter som ges. Det är förutsättningarna för observationen som är viktigast och inte vad som den som observerar gör. De distinktioner som bara finns som kognitiva element inom en människas medvetande gör snarast att vi fastnar i gamla positioner. En distinktion som upprepas konfirmerar bara identitet. Ett intelligent beteende kräver någonting mera. Det kräver att vi kan gå över gränsen och göra nya distinktioner.

5.2.2 Information som skillnad

Betydelsen av distinktionen kommer tydligast fram i de möjligheter som den ger att skapa ett informationsbegrepp som kan kombineras med en relativistisk kunskapsteori. Raf Vanderstreten framhåller att Luhmann ser information som det som genom kommunikation väljs ut av de möjligheter som exponeras, och att det är denna selektion som samtidigt gör kommunikationen möjlig (Vanderstraeten,2000:586). Vi bör uppmärksamma att detta visar på att kommunikation också har att göra med en förmåga att skapa en relation mellan system och omvärld.

En fråga som är viktig är hur ett yttrande uttrycker informationen. Vanderstreten framhåller att information bör ges i en form som yttraren och

adressaten kan förstå. Detta yttrande kan uppstå avsiktligt eller oavsiktligt. Det kan också uppstå utan ett språk, till exempel via ett tilldragande utseende, eller helt enkelt genom bortavaro. Det viktiga är att yttrandet kan tolkas som en selektion och inte bara som ett tecken för något annat. Dessutom innebär förståelse att någon förändring uppstår hos mottagaren. Att det är viktigt beror på att informationen som ges skall skapa nya möjligheter för kommunikation (Vanderstraeten,2000:587).

Vi skall understryka att Luhmann sätter förståelse i relation till att det är något nytt som skall förstås. Att bara förstå det som man inom en given systemreferens känner till är inte poängen med förståelse. Det är skillnaden mot det som finns i en given systemreferens som skall förstås. Ur det perspektivet blir t.ex. Searles betoning av förståelse tämligen trivialt. Genom den luhmannska teorins raster är förståelse något som först kan uppstå sedan förutsättningar har skapats genom att systemrelevansen förändrats. Och den förändras endast genom att omvärlden skapar en irritation som leder till strukturell koppling, vilket först sker när ett system inte längre klarar av att hantera sin omgivning och därmed tvingas att förändra sig. Enligt Luhmann förstår man bara inom ett system. Det strategiskt viktiga för att skapa ny kunskap bör då inte handla om förståelsen. Det är istället den kommunikation som ändrar systemreferensen som är viktig. På så sätt handlar inte den kritiska processen kring skapandet av ny kunskap om att individerna gör tankekonstruktioner i en förståelseprocess.

Detta leder till en uppfattning om att information exponerar osäkerhet. Loet Leydesdorff menar att vi uppfattar att information har två betydelser: information som osäkerhet och information som mening. Leydesdorff framhåller att det bara är den första varianten som betecknas som information i den luhmannska teorin. Information blir alltid lika med osäkerhet (Leydesdorff,1996:297).

En fördel med den andra ordningens cybernetik är att den inte bidrar till att människorna blir determinerade av givna värden som människorna konstruerat. Ett påstående som går tvärt emot en allmän uppfattning om hur cybernetiken fungerar. Men denna allmänna uppfattning är bara giltig för den tidigare cybernetiken och inte för det alternativ som Bateson m.fl. försökte utarbeta. Det viktiga är istället komplexiteten i systemet. Betoningen av det komplexa gör att Bateson kommer fram till att vi behöver det oklara för att ha något som vi kan skapa något nytt av. *All that is not information, not redundancy, not form and not restraints - is noise, the only possible source of new patterns* (Bateson,1973/1978:386).

De strategier som går ut på att göra ett budskap mer redundant, mer övertydligt, kan göra att mycket av potentialen för att skapa mening går om intet. Detta leder Bateson till att också göra en omdefinition av information som har blivit mycket inflytelserik. Den handlar om att den tekniska termen information definieras ***som vilken som helst skillnad som skapar en skillnad vid något senare tillfälle***. Denna definition är fundamental för alla analyser av cybernetiska system och organisationer. Enligt Bateson skapar definitionen en möjlighet att orsaken till

händelser måste ses som krafter, påverkan och liknande. Genom den cybernetiska analysen blir det möjligt att omforma orsaken till dessa händelser som skillnader (Bateson,1973/1978:351).

På så sätt betonar det cybernetiska perspektivet att meningssökandet blir viktigare än meningen. Det betyder samtidigt att det blir mindre skillnad mellan kunskap och information eftersom kunskap uppstår när en process har gett en observatörsposition som medger att vi kan se mening i informationen. Dirk Baecker knyter kunskapandet till social handling genom att påstå att det handlar om vilket som helst sätt att reflektera över skillnaden mellan samhället och det som inte uppfattas av samhället och om att genom social handling formulera om denna skillnad. Kunskapsskapande handlar på så sätt om att göra skillnaden uppenbar (Baecker,2008:54-55).

5.2.3 Observatören

Vi har redan varit inne på att möjligheterna för att observera är centrala inom den andra ordningens cybernetik. Ändå är det observerande systemet i den luhmannska teorin ett operativt slutet system. Det betyder t.ex. att människan, som är ett psykiskt system, inte kan se och analysera allt. En observatör kan se en annan observatörs blinda fläckar men aldrig den egna blinda fläcken (Jönhill,1997:62). Det innebär alltså ett avståndstagande från att medvetandet själv kan observera något utan att vara beroende av att interagera med andra system och med omgivningen. Det är också en förklaring till varför vi behöver social intelligens. Det psykiska systemet kan inte klara sig på egen hand. Det är något som naturligtvis tonar ner de medvetandeprocesser som sker inom en individ. Istället är det så att fakta i omvärlden måste presenteras och tolkas genom andra system, företrädesvis genom sociala system. Omvärlden kan visserligen påverka systemen varav människan är ett, men det sker inte med en direkt kognitiv modell där människan kan värdera allt, utan dessa fakta måste gå genom sociala processer. Vi kan säga att inte heller människan har den starka intelligens som kritiken av AI utgår från. Detta öppnar naturligtvis för att behovet av social intelligens har att göra med att medvetandet är begränsat, ungefär så som vi såg att Carley och Newell motiverade AI-system i inledningskapitlet. Vi bör observera att vi här tangerar en oklarhet hos Luhmann angående den sociala kommunikationen. Enligt en tolkning handlar Luhmanns kommunikationsbegrepp bara om kommunikationen inom ett socialt system. Relationen mellan systemen och omvärlden omfattas då inte av de kommunikativa processerna. Argumentationen för AIs kapacitet i denna avhandling bygger dock på att det faktiskt finns anledning att utvidga kommunikationsbegreppet till att också omfatta system/omvärldsrelationen. Det finns skrivningar i Luhmanns omfattande produktion som kan användas som stöd för en sådan uppfattning. Och en del av utvecklingen av teorin idag handlar om möjligheterna för en sådan utveckling. För

denna avhandling är den frågan inte avgörande. Det som är avgörande är istället att AI-diskussionen leder fram till att det faktiskt finns mycket som stödjer att system/omvärldsrelationen är central för möjligheterna att skapa ny kunskap. Om sedan den relationen ryms inom Luhmanns kommunikationsbegrepp är inte det mest centrala.

Luhmann löser problemet med den blinda fläcken genom att han inför begreppet observerande system av andra ordningen. Innebörden i detta är att observationer alltid måste relateras till en annan observatörs perspektiv. Observationer grundar sig då på redan företagna observationer och inte på objekten själva. Jönhill menar att detta kan beskrivas som att observerande system av andra ordningen innebär att en observatör alltid är ett system och kan därför betecknas observerande system, och att det som observeras består primärt av observationer av observationer (Jönhill,1997:55).

Luhmann nöjer sig inte med observatörer av första och andra ordningen. Också en observatör av tredje ordningen krävs för att förklara hur systemen observerar distinktionen mellan system och omvärld. Jönhill skriver att det handlar om den företeelse som man bland annat finner i massmedia och som består av att flera observerar samma objekt (Jönhill,1997:57). Luhmann talar bara om traditionell massmedia som tidningar, radio och TV. Men Felix Geyer har i artikeln *Virtual Communities in Cyberspace* visat på att detta perspektiv blir ännu viktigare i Internet-samhället (Geyer,1996).

5.2.4 Kontingens och dubbel kontingens

Det begrepp hos Luhmann som mest har att göra med vilka begränsningar som finns för direkt personlig kommunikation är kontingens. Det är ett fenomen som tonar ner att social intelligens måste bygga på interaktion mellan individer. Det är också ett begrepp som olika tänkare brukar behandla på olika sätt. Hos Luhmann är kontingens det som blir kvar när man utesluter det som är nödvändigt eller det som är omöjligt (Luhmann,1995:106). Alltså det som man vet är möjligt eller det som man vet att är omöjligt. Vi bör tolka Luhmann som att han menar att både det nödvändiga och det omöjliga kan bli osäkert respektive möjligt som resultat av att kommunikativa processer sätts igång. Poängen är att det i alla situationer finns så mycket som är kontingent att det alltid finns möjligheter att hitta nya sätt att se.

Luhmann inför också begreppet dubbel kontingens. Dubbel kontingens framträder i interaktioner. Den dubbla kontingensen innebär att en person inte bara har svårt att tolka en annan persons budskap. Den första personen har också svårt att tolka sina egna positioner i interaktionen med den andra personen. Därför går det inte att direkt förstå allt som en annan människa säger. Det som den andre människan säger kan bara öppna för vad som kan vara möjligt. Och då kan det också handla om sådant som den andre själv inte var medveten om att hans eller hennes uttalanden skapade öppningar för. Det gör att denna dubbla kontingens

förklarar de sociala systemens stora betydelse för kommunikationen. Eftersom människorna inte direkt kan förstå varandra måste kommunikationen gå via ett socialt system.

This modified, non-Parsonian understanding of double contingency has a twofold consequence. It enables the differentiation of a particular *world dimension* for socially distinct meaning perspectives (the social dimension) and it enables the differentiation of particular *action systems*, namely, social systems. The social is then accessible in all meaning as the problem of the similarity or discrepancy of interpretive perspectives. It is simultaneously a specific occasion for selectively coordinating actions within systems that can distinguish themselves from their environment. By modifying Parsons's theoretical approach, phenomenology and systems theory, the analysis of meaning and system/environment analysis, can be united (Luhmann, 1995:106).

Det öppnar för att ett socialt system måste bygga upp någon form av artefaktisk/kulturell struktur som kommunikationen kan gå genom. Och samtidigt är det denna artefaktiska struktur som visar var skillnaden mellan det sociala systemet och omvärlden finns. Vi kan observera att detta sociala system är ett aktivitetssystem. Eftersom det inte är människorna som kommunicerar så bör det vara artefakterna som skapar den aktivitet som ledet till kommunikation.

5.2.5 Strukturell koppling

Orsaken till att funktionella processer ändå på sikt kan skapa förändring i systemen, trots att dessa system är operationellt stängda, är att det förekommer något som Luhmann kallar strukturell koppling, vilket är en teori om förhållandet mellan systemen och deras omvärld.

Jönhill menar att teorin om strukturell koppling är en teori om det komplicerade förhållande som finns mellan människan, det sociala sammanhang hon befinner sig i och den miljö hon lever i (Jönhill, 1997:120). Trots att autopoietiska system är operativt slutna är de samtidigt kognitivt öppna. Denna operativa slutenhet kan synas absurd. Men när vi närmare funderar på begreppet förklarar det varför människorna inte blir determinerade av kausala processer. Det är inte genom kausalitet och symmetri som omvärlden påverkar medvetandet. Medvetandet förändras bara genom att inre processer retas av omvärlden så mycket att dessa inre processer söker efter nya möjligheter att göra de distinktioner som kan förändra medvetandet. Det är alltså genom att det autopoietiska systemet är öppet för kognitioner som det tvingas till förändring. För att funktionerna skall förändras måste systemet (vare sig det är ett psykiskt system med ett medvetande eller ett socialt system) successivt blir kapabelt att observera något nytt i omvärlden.

Teorin om strukturell koppling är en teori om hur vi långsamt anpassar vår egen låsta struktur till omvärldens struktur. I teorin om strukturell koppling finns argument för att människans medvetande är så pass styrt till tidigare erfarenheter att medvetandet inte räcker till för att på egen hand skapa ny kunskap. Människan måste istället, genom kommunikation med andra system, utveckla sitt medvetande så att de selekterande funktionerna anpassas och så småningom kan ta till sig allt mer från omgivningen och från andra system. Det innebär ett motstånd både mot behaviorism och mot naiv associationism.

5.3 Meningsproblematiken

Om kunskap skall kunna ha att göra med strukturella faktorer blir det en central fråga om mening kan handla om strukturella egenskaper. Utgångspunkten för den frågan måste vara att inte ens människan kan se någon mening i det som hon inte förstår, och att en av förutsättningarna för att kunna skapa kunskap därför är att vi måste utveckla något beteende som tillåter oss att hitta mening i det som vi vid en första exponering inte ser som meningsfullt.

Det är här som vi har stor användning av det sätt att behandla mening som Luhmann inför. Vi har redan varit inne på att mening i Luhmanns teori inte innebär någon givet. Även icke-mening inkluderas i meningsbegreppet. Mening fungerar istället självt som medium för de observationer som är beroende av distinktioner. Det finns således ingen meningslös upplevelse. Därför blir mening heller inget som är liktydigt med information. Jönhill ger en bra beskrivning av detta förhållande genom att uppmärksamma att ett meddelande som upprepas kan ha mening utan att för den skull ha informationsvärde. Information måste därför vara förknippat med nyhetsvärde (Jönhill,1997:156).

En annan fråga är om form är något som bara människor kan uppfatta. Vad som bestämmer formen har visserligen att göra med från vilket position vi observerar. Det kan dock bara räcka som ett svar på att AI-redskapen kan vara människan behjälpliga att själva se. Frågan är om artificiella aktörer, eller rättare sagt icke-psykiska system, överhuvud taget kan se. Det svar vi kan ge är att det är möjligheterna att göra nya distinktioner som avgör vad som kan observeras. De möjligheterna skapas genom att formen förändras och det är inte i första hand människor som ändrar dessa förutsättningar. Vi återkommer till den frågan i kapitel 9 där vi bekantar oss med en konversation mellan Hayles och Luhmann om den frågan. Vi kan tillägga att frågan också är behandlad av andra systemteoretiker. Till exempel av Helga Nowotny m.fl. i ett kapitel som heter *How Does Contextualization Happen* i boken *Re-Thinking Science* (Nowotny et.al.,2001).

5.4 Luhmanns funktionalism som en radikal samhällsteori

Luhmann anser att funktionsanalysen måste vara sociologins metod par preferens därför att sociologin måste vara den vetenskap som fokuserar på att upplysning innebär att det är samhällenas strukturer och funktioner som utvecklas, och inte att det innebär att det är individernas medvetanden som skall upplysas. King och Thornhill framhåller att den klassiska uppfattningen om vad som ingår i korrekta och rationella individuella avgöranden inte är relevant för att diskutera det moderna samhällets funktionella komplexitet. Den rationalitet som hör samman med vår bedömning av hur samhället fungerar kan bara utvecklas genom att den process stimuleras som kan konstruera och stabilisera ännu mera komplexa system. Luhmanns position bygger på att han anser att den upplysning som fokuserar på människans förnuftsformåga inte är en egentlig upplysning. Genom att den skiljer ut det mänskliga förnuftet från händelser i den sociala utvecklingen bygger den på metafysik. På så sätt kan den inte förstå innebörden i modernitet eftersom den bygger på en uppfattning om att samhället består av individuella förekomster av förnuft. Det leder till en förenkling av vad social komplexitet handlar om. En sådan uppfattning kan bara reproducera samma metafysiska uppfattningar som den kritiserar.

For Luhmann, therefore, the imputation of rational, causal or moral criteria standing independent of all social reality is in fact only the most rudimentary, underdeveloped expression of Enlightenment. Genuinely postmetaphysical Enlightenment, in contrast, would be prepared to accept the complex and changing forms of reason, and it would reflect on the extent to which the processes of social evolution follow extremely variable and fluctuating rational imperatives. Above all, real (sociological) Enlightenment necessarily rejects the quasi-metaphysical fictions that the social world has an essential structure and that this can somehow be causally divined in the medium of human rationality, or morally prescribed in the medium of human law. (King,&Thornhill,2003:136

King och Tornhill skriver att vi ser att Luhmanns funktionalism inte är en konservatism om vi överger den optimistiska uppfattningen om att varje människa bidrar till välbefinnandet i helheten. Skall en förändring eller progressiv utveckling ske, måste det gå via en förändring av samhällets koder och den funktionella specificeringen som hör till koderna:

This is only 'conservatism' to the extent that it denies the possibility of any social system taking over the function of another or of reforming the other in accordance with its own directives. Whether Luhmann offers this only as a description of modern society, or as a normative model - an account of what modern society ought to be - remains debatable. As far as social change is concerned, however, within the constraints imposed by functional specificity this is not only possible, it is inevitable. Every decision changes the world in some way and millions of decisions are made every day. What Luhmann is fundamentally opposed to is totalizing or dogmatic accounts of society which claim that social change may be managed, controlled or predicted. Change for Luhmann is essentially contingent in nature. Something happens, a decision is made which, through a process of making distinctions - 'it is this and not something else' - gives rise to further decisions and the making of further distinctions. Explanations for such changes which incorporate, as they are bound to do, the identification of causes and rationalizations as to why one course of action was taken rather than another are themselves the result of choices, of selections and rejections. They are nothing other than attempts by communication systems to give meaning to contingent occurrences. Like all system operations, they are selective and exclusive; they see only what they are able to see. Any explanation or rationalization can in sum offer only a partial view of complexity (King, & Thornhill, 2003:34).

Luhmann ger oss anledning fundera på om teorier som bygger på att medvetandet skall bli mer utvecklat, i själva verket bidrar till att konservativa åsikter behålls. Det är bara genom utökad differentiering som samhället utvecklas och skapar utökad solidaritet.

Vi har i detta kapitel visat att Second order cybernetics handlar om något helt annat än den mekaniska cybernetik som den har vuxit fram ur. Vi har istället sett att denna nya cybernetik är en mycket bra modell för att förklara hur människorna interagerar med sin sociala och kulturella struktur. Vi har också tagit upp några av de viktigaste beståndsdelarna i den Luhmannska teorin och visat på varför dessa begrepp är bra att använda för att förklara hur sociala processer kan kopplas samman med den kulturella strukturen. Vi har också visat att detta gör den Luhmannska teorin till en radikal samhällsteori som hjälper oss att se människans möjligheter i det nya och datoriserade samhälle som bland annat använder sig av artificiell intelligens. Genom Luhmanns teori har vi styrkt att den konnektionistiska AI:n verkligen är ett alternativ till symbolisk AI och att det ger anledning att hävda att de AI kritiker vi har mött, framför allt Searle och Harry Collins har fel.

6. Sociala processer som artefaktisk aktivitet

Vi har kommit så långt i analysen att vi har sett att det är tveksamt om de sociala processerna verkligen står i motsättning till att AI-system kan skapa kunskap. Vi har dock inte riktigt visat på hur sociala processer kan uttryckas genom artificiella processer. Vi bör framhålla att det inte är ett problem som endast berör AI-frågan. Sociologin är över lag mycket dålig på att visa på vilken konkret funktionalitet sociala processer innehåller. Durkheim har visat på betydelsen av funktionalitet, och Parsons har utvecklat den diskussionen. Bägge knyter funktionaliteten till en ontologisk realism vilket gör att deras teorier blir rätt så mekanistiska. Ingen har som Luhmann visat på att funktionaliteten också har en betydelse i en relativistisk modell. Det som gjorde mig intresserad av hur funktionalistisk teori kan förklara samhällsprocesser var just Durkheims begrepp organisk solidaritet. Också det begreppet föreföll mig en aning svårförklarigt innan jag blev bekant med den luhmannska varianten av funktionalism.

Vi kan påminna om att ett av de motiv den AI-positiva sociologin anger för att intressera sig för AI, just är att den tekniken ger möjlighet att modellera, pröva och analysera sociala processer. Vi kan också hävda att frågan om AI-teknikens roll i kunskapsprocesserna då inte bara handlar om på vilket sätt AI-tekniken i sig fungerar. Den handlar också om hur de sociala processer ser ut där både AI-teknik, annan teknik och människor deltar. Därför handlar frågan om hur kunskapsskapande processer över huvud taget fungerar.

6.1 Medvetandets behov av redskap

Om vi intresserar oss för hur redskapen kan vara en social kraft som befriar människan har andra vetenskaper än sociologin mer att säga, bland andra antropologin. Bland sådana vetenskapliga framställningar märks den skrift som Tim Ingold redigerat: *Tools, language and cognition in human evolution*. Här skall vi

emellertid gå vidare med vad vygotkskyorienterad pedagogisk teori kan bidra med. James Wertsch talar om att människan inte behöver internalisera förmågan att använda redskapen innan de kan bli intelligenta redskap. Vi skall se hur Wertsch fördjupar den diskussionen mot slutet av kapitlet. Här skall vi se på hur han i boken *Vygotsky and the Social Formation of Mind* redovisar hur Vygotsky ser på sambandet mellan socialiteten och det individuella medvetandet.

Wertsch pekar också han på att Vygotsky skiljer mellan den utveckling som en art går igenom och den utveckling som sker med individen inom en art under dennes liv. Wertsch framställer det som att Vygotsky utgår från att artens utveckling, fylogenesen, inte kan förklara hur individens utveckling, ontogenesen, sker. Den senare utvecklingen är kulturell. Fylogenes och ontogenes är samtidigt det som skiljer elementära mentala funktioner från högre mentala funktioner. Den naturgivna utvecklingen förklaras med biologiska principer. Den kulturella utvecklingen har däremot att göra med faktorer som handlar om förmedlande redskap, inklusive dekontextualisering (Wertsch,1985:42). I en diskussion om AI betyder det att det som hör till ontogenesen inte bara behöver handla om utvecklingsmöjligheter som direkt skapats av människor. Det kan också handla om att en kulturell struktur kan påverka människorna under ontogenesen genom att de artefakter som skapar strukturen utvecklas. På så sätt påverkas människornas förmåga till intelligens av att en utveckling av kulturella redskap skapar en bättre miljö.

Wertsch menar att Vygotsky antog att naturgivna krafter upphör att spela en roll för individen ganska tidigt i den personliga utvecklingen och att det därefter är kulturella krafter som tar över den ledande rollen. Dessutom fungerar de naturliga och kulturella krafterna helt oberoende av varandra under de tidigare faserna av ontogenes (Wertsch,1985:44). Det är inte helt klart hur dessa kulturella faktorer skapas. Vygotsky framhåller att det är redskap som förmedlar dem, vilket antyder att det finns en betydande grad av artificialisering inblandad i den process där individerna utvecklas.

Det kan tyckas som mycket radikalt att en kulturell struktur kan vara del av det som skapar människans kapacitet. Funderar vi närmare på vad det skulle betyda att denna struktur inte alls har någon sådan kapacitet så hamnar vi också i en mycket radikal position. Då måste vi hävda att det bara kan vara människan som en biologisk organism som kan utvecklas. Det blir tveksamt om det ens kan handla om att människor påverkar varandra eftersom människors interaktioner uttrycks genom kulturella artefakter som språk och metaforer. Knappast går det heller att hävda att det är människornas interaktioner som skapar artefakterna som sedan underlättar fortsatta mänskliga interaktioner. Ty om artefakterna inte någonstans får en kapacitet att utveckla förutsättningarna för dessa interaktioner blir människornas interaktioner instängda i en given systemreferens. I så fall hamnar vi i en position där det som skapar intelligens är egenskaper som utvecklats under artens utveckling, alltså fylogenesen. Då kan vi heller inte knyta den sociala

kapaciteten till ontogenesen eller för den delen till redskapsanvändandet. Då blir det svårt att påstå att människans kultur kan skapa det som gör att människan till skillnad från djuren blivit en varelse som kan utveckla intelligens.

Om intelligens hör samman med fylogenesen blir Searles uppfattning att människan är den enda maskin som är tillräckligt komplex för att kunna tänka logisk. Men då får också Searle svårt att förklara kulturell variation i intelligent beteende. Vi bör tillägga att argumentationen för stark AI bara kräver att den kulturella strukturen kan ha en viss omstrukturerande kraft i sig. Den behöver inte vara den huvudsakliga orsaken till att människans intelligens utvecklas. Ty den mänskliga intelligensen beror på många faktorer. Det låter rimligt att intelligensens utveckling handlar om att många faktorer samverkar.

Vi kommer då tillbaka till att frågan om stark AI kräver att vi utreder om redskap och kulturella artefakter kan förändra sig utan interaktion med en människa. För att fortsätta utreda den frågan bör vi se närmare på internaliseringens processer. Wertsch framhåller att Vygotsky liksom flera andra teoretiker, bl.a. Piaget, ser internalisering som en process som handlar om hur vissa externa aktiviteter omvandlas och används på individernas interna plan. Men olika andra teoretiker definierar Vygotsky extern aktivitet i termer av semiotiskt förmedlade sociala processer. Eftersom semiotiken är läran om hur kulturella artefakter får betydelser, betyder det att det är artefakterna, eller de kulturella redskapen, som skapar förutsättningarna för det individuella. Och samtidigt innebär det att det är dessa externa processer som erbjuder förklaringen till hur de interna processerna fungerar. Wertsch menar att Vygotsky framhåller att all högre kunskap som finns inom en individ, tidigare har förekommit i extern form. Det betyder att varje högre tanke måste gå genom ett externt stadium under sin utveckling eftersom den ursprungligen hade en social funktion. Det betyder också att det som är externt i processen också är socialt. Vilken som helst mental funktion har varit extern eftersom den var social någon gång före den blev en intern mental funktion (Wertsch, 1985:62).

Detta betyder inte att Vygotsky såg internaliserade mentala processer som en kopia av externa processer. Istället menar Vygotsky att internaliseringen ändrar **hela** processen. Det finns en inbyggd relation mellan extern och intern aktivitet. Det kan tolkas som att vi ändå behöver det mänskliga medvetandet någonstans. Och det kan tolkas som att de externa processerna bara kan uppstå som resultat av interna processer.

Wertsch menar att Vygotskys positioner är något oklara. Det som framför allt framstår som oklart är vad som ändrar hela processen. Utifrån Luhmanns teori kan vi hävda att det som ändrar hela processen är att psykiska system, tack vare att de konfronteras med sociala system, efter ett tag börjar se på sig själva på ett annat sätt. Det är först efter att de psykiska systemens hela process har blivit ändrad som de har förutsättningar för att använda sina djupsinniga själsliga kapaciteter till att tänka i nya banor. Till exempel till att ta emot kunskapen så att den internaliseras i

deras medvetanden. Det viktiga är att detta inte kan ske utan att sociala processer har gett människorna förutsättningar för att tänka i nya banor. Det vi bör uppmärksamma är att den komplexa funktionaliteten då först kommer efter att människorna blivit exponerade för kunskap.

Wertsch framhåller att den mekanism som ligger bakom högre mentala funktioner är en kopia av mekanismer från sociala interaktioner. Alla högre mentala funktioner är internaliserade sociala relationer. I sin egen privata sfär bevarar människorna funktionerna från den sociala interaktionen (Wertsch, 1985:66). Det är ståndpunkter som mer liknar de relationer som den luhmannska systemteorin talar om. Högre mentala funktioner har att göra med externa faktorer vilket bör innebära att bl.a. intelligens hör samman med de yttre faktorer som föregår internalisering.

Problemet med Vygotskys teori i jämförelse med Luhmanns är att vi fortfarande inte får en uppfattning om hur det sociala skapas, och vi får heller inte veta hur det sociala når individen annat än att det handlar om en kopia av mekanismer från sociala interaktioner. Det förefaller underförstått att dessa sociala interaktioner i sig inte kan vara en kopia av hur tänkandet fungerar. Vi lämnas således kvar med en osäkerhet om hur det sociala når människorna. Det vi får veta är bara att det finns något speciellt med de sociala funktionerna som inte grundas på mänskligt tänkande. Dessutom innebär detta att tänkandet i sig kan ha en funktionalitet som söks i sociala processer. Om det är så att sociala processer kan innehålla artificiella element kan vi också hävda att intelligenta processer kan vara artificiella. Det stärker möjligheterna för att den enkla funktionalitet som behöver finnas i AI-systemen, och som vi hos människorna upplever som tänkande, inte behöver vara så svår att efterlikna. Åtminstone behöver den inte ha att göra med den mänskliga hjärnans naturgivna kapaciteter.

Det är alldeles klart att Vygotsky är en tänkare som mer eftertryckligt än andra betonar att det individuella är beroende av det sociala. Att han betonar att medvetandets interna nivå är kvasisocial gör att vi känner igen en luhmannsk argumentation. Eftersom språket, eller externa tecken, är det semiotiska redskap som förmedlar det sociala till individen, bör vi komma närmare frågan om artefakternas kapacitet genom att se hur Vygotsky uppfattar språkutvecklingen. Vi vet så här långt inte vilken funktionalitet denna process innehåller. Om högre mentala funktioner är en kopia av sociala interaktioner antyder det att funktionaliteten också är en kopia av externa faktorer. Innan vi återkommer till språkutvecklingen skall vi se på en fråga som ofta dominerat AI-diskussionen. Det är frågan om problemlösning.

6.2 Expanderings och redskap

Yrjö Engeström, som är professor i pedagogik i Helsingfors, diskuterar i sin bok *Learning by expanding* medvetandets förhållande till omgivningen i situationer där aktören vill utvidga existerande kunskap. Engeström hävdar att ett problemlösande angreppssätt inte är särskilt bra i en sådan situation. Problemlösning menar han är en reaktiv form av lärande därför att den förutsätter att en given kontext ger uppgifterna åt individerna. På det sättet bortdefinieras lärandet som ett sätt att hitta nya kontexter (Engeström, 1987:2). Diskussionen är intressant eftersom den hjälper oss att få ett nytt perspektiv på den AI-diskussion som ofta fastnat i diskussionen om problemlösning. Den visar på att skapandet av ny kunskap kräver mer än problemlösning. Dessutom utvecklar Engeström de vygotsky-inspirerade teoriernas förhållande till redskapsanvändande.

Engeströms argumentation tar sikte på hur människan kan ta till sig det nya som hon inte vet att hon söker efter. Han diskuterar därför framför allt hur människan beter sig för att expandera kunskapen. Det som Engeström tillför diskussionen är att han menar att expansiva processer är beroende av materiella förutsättningar, vilket ger ytterligare en möjlighet att se att en sådan materiell storhet som datateknik kan knytas till en social funktionalitet.

Engeström knyter denna kritik av problemlösande strategier till att ifrågasätta den induktiva logiken. Han menar att den per definition inte förmår gå utöver den givna kontexten. Lärandet kan då inte utvecklas längre än vad kontexten tillåter. Engeström kommer fram till två påståenden: a) att uppfattningen om att skapande består av induktioner som kombineras är fundamentalt falsk, och b) att uppfattningen om att den högsta formen av lärande består av ett induktivt och kombinerande problemlösande också är fundamentalt felaktig (Engeström, 1987:4).

Engeström framhåller att alternativet till reaktiva former av lärande är expansion, vilket han menar är det som förmår överskrida de givna kontexterna. I det sammanhang vi har diskuterat skapandet av ny kunskap, handlar det om en kapacitet att överskrida det som är den socialt accepterade kunskapen. För att skapa en strategi för ett expanderande lärande undersöker Engeström hur individen kan närma sig sådant som hon är omedveten om. Engeström menar att dessa processer också måste ha någon form av materiell bas (Engeström, 1987:6). Det antyder att det sociala bör innefatta andra processer än sådana som på sätt eller annat har ursprung i människors handlingar och människors tänkanden. Engeströms uppfattning om expansion påminner om Peirces begrepp abduction. Att Engeström hellre använder Vygotsky som teoretisk inspiratör har att göra med att Vygotsky bättre knyter sin diskussion till materiella förutsättningar.

Problemet med att se att konnektionistiska program kan skapa dessa materiella förutsättningar är naturligtvis att de gärna blir induktiva. Den kritiska frågan för de konnektionistiska systemen är då om de kan vara redskap som ökar

den sfär som kan påverka oss. Engeströms kritik av induktionism kräver att dessa system måste ha någon egenskap som gör att de själva kan skapa nya kunskapsstrukturer. Annars är de bara redskap för associativa beteenden. Därför måste vi ta den kritik mot konnektionismen på allvar som går ut på att den bara bygger på associationism. Konnektionismen måste kunna exponera människorna för kunskapsrelationer som de tidigare inte sett. Vi bör dock inse att det är skillnad på under vilka förhållanden induktioner görs. Vi har den form av induktion som en människa gör utifrån sin bristfälliga förmåga att observera över stora områden. Och en annan form som handlar om att människorna ges möjligheter att göra associationer genom att datasystem exponerar dem för likheter och skillnader i hela den världsvida litteratur som finns elektroniskt publicerad. När människan måste utgå från det hon redan känner till blir både induktionen och deduktionen bunden till det mänskliga medvetandets kapacitet. Ett datasystem kan faktiskt på ett helt annat sätt pröva observationsmöjligheterna utifrån många fler positioner. Det är en empirisk fråga om detta bara leder till kaos. Vi bör uppmärksamma att det finns en praktisk skillnad på vilken funktionalitet vi verkligen kan förvänta oss av människors observationsmöjligheter och av datasystemens motsvarande möjligheter. Det som konnektionistiska system kan utföra, bygger på en funktionalitet som liknar den som finns i induktion, men datasystemen är inte lika bundna som människorna till tidigare erfarenheter. Det har att göra med att de inte på samma sätt som människan styrs av emotionella bindningar till tidigare erfarenheter. Istället har datasystemen större möjligheter än människorna att skanna igenom all litteratur som finns elektroniskt lagrad. De har därför en egenskap som fungerar bättre än motsvarande egenskap hos människorna. Den egenskapen bör göra det lättare för datasystemen att kunna gå utöver den givna kontexten.

Engeström menar också att fördelen med Vygotskys teoribildning är att den bättre låter oss se hur skapandet av kunskap avspeglar sig i det materiella. Engeström ställer frågan om detta betyder att det bara är kognitionen eller kommunikationen som reorganiseras, medan den materiella praktiken blir oberörd. Engeström menar att sann expansion alltid innehåller både en mental och en materiell sida. Dessutom menar Engeström att expansiva processer kan analyseras och ritas upp. Engeström understryker att vi måste analysera den aktivitetsprocess som leder till expansion om vi skall nå insikt om hur dessa processer fungerar. Den processen har att göra med hur expansion kan kopplas till användning av olika typer av redskap.

Kopplingen mellan sociala processer och tekniska redskap har att göra med att Vygotsky gör en skillnad mellan två relaterade typer av förmedlande instrument i människans aktivitet, redskap och tecken. De senare hör till kategorin psykologiska redskap. Vi har således två typer av redskap, tekniska och psykologiska. Om det är så att bara psykologiska redskap kan förmedla kunskap måste sociala processer vara knutna till mänsklig fattningsförmåga, medan saken kommer i ett annat läge om

tekniska redskap har denna förmedlingskapacitet. Hur dessa förhållanden ser ut bör vi kunna komma närmare genom att undersöka om psykologiska redskap kan vara knutna till kulturella kontextfaktorer eller om de helt måste knytas till mänskliga handlingar.

Engeström menar att både tekniska och psykologiska redskap förmedlar aktivitet. Men endast psykologiska redskap innebär och kräver ett reflekterande och förmedlande medvetande från individen över sig själv eller andra. Det skulle faktiskt innebära att redskapen måste vara knutna till mänskliga handlingar. Det bör dock inte vara Engeströms poäng eftersom expanderingskräver att vi kan reflektera över sådant som vi inte kan göra med medvetandet. Det är därför som de psykologiska redskapen inte är tillräckliga utan vi behöver också den materialitet som tekniska redskap erbjuder. I sammanhanget bör vi tänka på att även boken är ett tekniskt redskap. Det bör vara en ståndpunkt som överensstämmer med Engeströms argumentering för att expansion kräver någon form av materialitet. Det kan vara så att enbart psykologiska redskap inte räcker till för att expandera kunskapen. Expansionskräver att vi också ges tekniska förutsättningar. Människan måste ha ett tekniskt redskap som kan användas för att observera även sådant som inte hon själv eller andra redan har reflekterat över. Problemet med psykologiska redskap är att de är kopplade till medvetandet. Om de inte förlängs med mer materiella egenskaper är de inte användbara i vidare sammanhang än sådana där kommunikation sker direkt i mötet mellan två eller flera människor.

6.3 Semiotiken som det artificiella hos Vygotsky

Det som blir oklart är hur vi kan argumentera för att tekniska redskap blir laddade med en sådan social aktivitet att de själva kan skapa sociala processer. Det som närmast tangerar denna fråga hos Vygotsky är det han säger om den semiotiska analysen. Vi skall därför fördjupa oss i den redovisning som Wertsch gör av Vygotskys semiotik för att få en bättre möjlighet att se på relationen mellan psykologiska och tekniska redskap. Och samtidigt får vi en möjlighet att få svar på frågan om ett tekniskt redskap på samma gång kan vara ett psykologiskt redskap. I det sammanhanget bör vi framhålla att DAI-perspektivet anger att den funktion vi söker i det tekniska redskapet inte behöver vara komplex.

Wertsch framhåller i *Vygotsky and the Social Formation of Mind* att Vygotsky understryker att denna distinktion mellan psykologiska och tekniska redskap har sina begränsningar. Ett tekniskt redskap fungerar som en dirigent för den mänskliga påverkan på objekten. En funktion som är inriktad mot externa objekt. Ett psykologiskt redskap influerar däremot ett beteende – antingen hos individen själv eller hos någon annan (Wertsch, 1985:77-78). Det kan antyda att de

psykologiska redskapen handlar om att etablera kunskapsintressen medan det är de tekniska redskapen som ger förutsättningar för expansion.

Wertsch framhåller också att Vygotskys diskussion av psykologiska redskap utvecklades under åren och blev mer semiotiskt orienterad, vilket leder fram till att psykologiska redskap ges en mer teknisk natur. År 1930 angav Vygotsky språk, olika system för att räkna, minnestekniker, algebraiska symbolsystem och konstverk som exempel på psykologiska redskap (Wertsch,1985:79).

Det ser alltså ut som att det finns mindre skillnad mellan tekniska och psykologiska redskap i Vygotskys senare texter. Wertsch framhåller emellertid att denna lista bara kan förstås om vi ställer den i ett sammanhang där vi talar om mening. Wertsch framhåller dessutom att vi måste se på ett antal egenskaper hos psykologiska redskap om vi skall förstå deras funktion för medvetandets meningsprocesser. Om vi ser på denna fråga utifrån Luhmanns teori kanske vi ändå inte skall tala om mening eftersom Luhmann menar att det inte finns något som inte har mening. Däremot kan vi se på redskapen som det som kan omskapa kontexterna så att vi kan se ny mening.

Wertsch återkommer till att flödet och strukturen hos mentala funktioner förändras av psykologiska redskap. På så vis är inte psykologiska redskap något som bara underlättar mentala funktioner. De är också något som förändrar de mentala funktionerna. Det bör alltså betyda att de psykologiska redskapen föregår det mänskliga tänkandet. Detta bör innebära att de inte bara kan vara ett resultat av mänskliga intentioner. Det bör medge att tekniska redskap åtminstone i vissa situationer kan ge upphov till att psykologiska redskap utvecklas. Det är t.ex. sannolikt att användningen av datorer redan har lett till att våra psykologiska redskap förändras.

En annan egenskap hos psykologiska redskap är att de är sociala och inte organiska eller individuella. Här finns det en motsättning hos Vygotsky. Den har att göra med att han ser att dessa redskap inte bara är en produkt av den sociokulturella utvecklingen utan också hör samman med social interaktion och personlig kommunikation. Den sociokulturella utvecklingen av redskapen har vi tidigare varit inne på. Betydelsen för den sociala interaktionen har att göra med att ett tecken alltid är ett medel för sociala avsikter. Tecken påverkar således den sociala interaktionen och den sociala kommunikationen mellan två människor (Wertsch,1985:79-81). Vygotskys otydlighet kommer av att det inte framgår om redskapen i sig blir påverkade av den sociala kommunikationen och i vilken mån det är människans handlingar i denna kommunikation som förändrar redskapen. Som vi flera gånger har pekat på framhåller Luhmann att det bara är sociala system som kan kommunicera. Interaktion hör till de begrepp hos Luhmann som är något oklara. Vi bör uppmärksamma att Luhmanns snarast knyter interaktioner till relationer mellan människor. Luhmanns begrepp innebär således att vi snarast skall vända på Vygotskys ordval och tala om att tecken påverkar den personliga interaktionen och den sociala kommunikationen. I det perspektivet kan personliga

interaktioner inte handla om kommunikation eftersom människor inte kommunicerar. För att kommunikation skall kunna uppkomma inom ett socialt system måste det tillföras en funktionalitet som har att göra med någon annan aspekt av det sociala systemet än att människor interagerar. Det är därför som det finns större förutsättningar hos Luhmann att knyta förutsättningar för att expandera till tekniska redskap än hos Vygotsky som ändå ser en större roll för människornas handlingar. I det sammanhanget bör vi uppmärksamma att social kommunikation hos Luhmann är en förutsättning för att ny kunskap skall kunna skapas. Hos Luhmann är interaktioner mellan människor alltid bundna till det som redan kan förstås. Förslagsvis kan vi knyta Vygotskys uppfattning om betydelsen av de tekniska redskapen till att de har en funktion som förmedlar mellan systemen och deras omgivning. För att kunna diskutera det förhållandet behöver vi ta upp Luhmanns diskussion om självreferens.

6.4 Reflexion, social kommunikation och omvärlden

Vi har talat om att Luhmann menar att systemen är operationellt slutna. Det som gör systemen slutna är att de är *självreferentiella*. Systemet ser på sig självt på det sätt som det har vant sig att uppfatta sig självt. För att det psykiska systemet, i.e. människan, inte skall fastna i sitt eget medvetandes begränsningar behöver det således något som gör att det kan se längre än till sig självt. Det viktiga är de processer som är igång, och att förståelse är något som uppstår efter den självreferentiella processen. Förståelsen blir då reducerad till något som faller på plats efter självreferentiella processer. De meningsskapande processerna föregår förståelsen och i dessa processer är det reflexionen som är det väsentliga. Vi kan naturligtvis inte säga att dessa andra system inte behöver vara system där observatören är av mänsklig natur. Dock betonar Luhmann att dessa system skall använda skillnaden mellan system och omvärld. Det innebär att vi inte kan se på det sociala som något som bara är skapat av mänskliga handlingar. Frågan är snarast i hur hög grad det är de villkor som har att göra med det artificiella som är avgörande för den självreferentiella processen. Det kan förklara varför tekniska redskap behövs för kunskapsutveckling.

Vi kan uppmärksamma att vi här också får en förklaring till att det finns en viss giltighet i det systemsvar som riktades mot Searles experiment. Det som skapar meningen är inte själva förståelsen utan den distinktion som uppstår som ett resultat av den skillnad som finns mellan system och omvärld.

Den självreferentiella process som skapar kunskap är social. Det innebär inte detsamma som att alla former av självreferens är sociala. Och inte heller det samma som att självreferens aldrig behöver den funktionalitet som finns i de psykiska

systemen. Därmed kan vi heller inte säga att AI-system kan leva upp till alla krav på självreferens. Skall vi diskutera dessa frågor bör vi se vad Luhmann själv säger om självreferens. Han anger att vi kan skilja mellan tre former av självreferens (Luhmann,1995:443-444):

Vi kan tala om **enkel självreferens** när den grundläggande distinktionen är mellan element och relation. Det jag som refererar till sig själv kan då också vara en händelse, eller i sociala system, en kommunikation. Denna enkla självreferens är den minimala formen av självreferens utan vilken autopoietisk reproduktion skulle vara omöjlig.

In basal self-reference, the self that refers itself is also an element, for example, an event or, in social systems, a communication. Basal self-reference is the minimal form of self-reference, without which autopoietic reproduction of temporalized systems would be impossible.... Basal selfreference is a constitutive requirement for forming self-referential systems, but it is not a system reference, since the indicated self is intended as an element, not as a system, and since the guiding distinction is element/relation and not system/environment. This of course does not deny that the concept of element presupposes a system and vice versa. But that does not negate the distinction between different forms of self reference; it merely grounds the expectation that they correlate to each other.

Den andra formen är **reflexivitet**. Här består den grundläggande distinktionen av det som sker före och efter. Det jag som refererar till sig själv är inte bara en aspekt av distinktionen utan också en process som konstitueras av distinktionen. Den blir då beroende av att skillnaden mellan före och efter blir större genom att selektionen leder till ökande förutsättningar för fortsatt selektion. På så sätt ökar reflexiviteten samtidigt som den intensifierar egenskaperna som karaktäriserar denna process.

A process emerges with the help of the before/after difference if the additional condition of an increase in selectivity is fulfilled. Thus communication as a rule is process, namely, is determined in its elemental events by the expectation of a reaction and the reaction to an expectation. One can speak of reflexivity whenever a process functions as a self to which the operation of reference belonging to it refers. Thus within the course of a communicative process one can communicate about that communicative process. Thus reflexivity takes advantage of a unifying formation that combines a multitude of elements (often a countless number) within which the self-reference includes itself. Above all, this means that the self-referential operation must comply with the characteristics of belonging to a specific process, must be communication in a communicative process (communication about communication), observation in a process of observation (the observation of observation), and an application of power in a process of applying power (the application of power to the powerful). In this sense reflexivity increases and intensifies the features that typify process.

Slutligen kan vi tala om **reflexion** när den grundläggande skillnaden är mellan system och miljö.

Only in reflection does self-reference exhibit the characteristics of system reference; only here do the two conceptual domains overlap. The self is the system to which the self-referential operation attributes itself. It is an operation by which the system indicates itself in contrast to its environment. This occurs, for example, in all forms of self-presentation that assume the environment does not immediately accept the system in the way it would like itself to be understood.

Det framstår som om den **första** formen, den som handlar om självreferens, väl beskriver det som sker i förståelseorienterade processer. Det är en form av referens som bara riktas inåt mot det som individen eller systemet redan känner till. Därmed är det en form av referens som inte kan expandera kunskapen. De redskap som är inblandade bör bestå av allmänt vedertagna begrepp som möjligheten kan betecknas som ett slags psykologiska begrepp som beskriver existerande kunskap.

Den **andra** formen, som handlar om reflexivitet, är starkt beroende av att den erfarenhet som en person skaffar sig påverkar en tidigare kunskapsnivå. Samtidigt är detta också en form av referens som bara riktar sig inom systemet. Luhmann talar om att tidsdimensionen är viktig. Det bör innebära att tiden gör att många av de osmälta kunskapsrelationer som individen bär på så småningom mognar. Tidens gång gör att individen har kunnat ta till sig något som hon vid första kontakten tyckte var intressant men inte helt lätt att förstå. På så sätt kan vi också tänka oss

att psykologiska begrepp kan användas som ett slags referenspunkter till vad som var tidigare och vad som kom efteråt. På så sätt kan de förstås fungera som expanderande redskap men denna expansion blir ändå begränsad till vad systemreferensen redan tillåter eftersom det måste handla om att reflektera över sådant som redan finns inom systemet.

Det är först med den **tredje** formen av referens, den som handlar om reflexion, som vi får en tillfredställande förklaring till varför ny kunskap kan tränga sig på oss trots att vårt medvetande inte kan se den kunskapen, och trots att det är en kunskap som går emot den som är allmänt accepterad inom vår egen sociala tillhörighet. Det sker genom att systemet blir förändrat så att det blir uppmärksammat på skillnaden mellan dess egen struktur och det som finns i omgivningen. På så sätt ändras också systemreferensen och tillåter människorna att se nya kunskapskopplingar. Det blir ganska rimligt att detta i första hand handlar om att något som är fristående från det mänskliga utvecklas. Om vi ser på materialitet som det som är på sidan av vad som finns i de mänskliga medvetandena så bör det just handla om att en materiell struktur utvecklas. Om redskapen utvecklas så bör det handla om tekniska redskap. Det är genom dem som vi först blir varse nya skillnader. Först därefter kan de formuleras som psykologiska begrepp/redskap.

Därför finns det anledning hävda att det är denna form av referens som anger att det finns anledning tro att AI-system verkligen kan skapa kunskap på egen hand. Den socialitet som utvecklar kunskapen handlar om en process som uppstår av differensen mellan systemet och miljön. Genom den processen uppstår de nya situationer som ger oss möjlighet att göra nya observationer. Vi bör uppmärksamma att det finns mindre anledning tro att en sådan process skapas av interaktioner med andra människor eftersom de vi personligen kan interagera med befinner sig inom samma system som vi själva. Och de som finns inom andra sociala system kan vi inte förstå utan att tekniska redskap som förmedlar först utvecklas. Människor i andra sociala system kan vi bara nå genom att gemensamma referensramar uppstår. Innan vi kan samtala med människor från andra kulturer (eller kanske även från andra vetenskapliga discipliner) måste ett kulturellt kapital i form av gemensamma benämningar och gemensamma berättelser uppstå genom att våra respektive sociala system närmar sig varandra. Vi bör tillägga att Luhmann här beskriver en process som handlar om något mer än systemens egen kommunikation. Eftersom Luhmann menar att kommunikation bara sker inom sociala system så handlar det om något mer än kommunikation. Som redan sagts finns det skäl att försöka utveckla Luhmanns kommunikationsbegrepp så att det också omfattar denna relation. Min egen uppfattning är att jag bättre förstår Luhmanns uppfattning om att det endast är systemen som kan kommunicera om det handlar om just relationen mellan systemen och deras omgivning.

Vi kan också uppmärksamma att denna tredje form av självreferens ger oss en möjlighet att förklara hur de mentala processerna är kopior av sociala processer

såsom Vygotsky hävdade. Vi kan inte med medvetandet ta till oss det som finns i omgivningen. Ny mening måste alltid gå via den sociala kommunikationens reflexion och när människans medvetande som en kopia av denna sociala kommunikation. Det är också en förklaring till varför social intelligens kan ta oss längre än medvetandets intelligens.

Sociologins debatt med naturvetenskapen om kunskapens natur är ett exempel på att människor som finns inom olika system har svårt att kommunicera med varandra. Det är först genom att hela sociala system har brutits mot varandra som en viss samsyn börjar skönjas. Den består inte av att olika personer inom de olika grupperna kommunicerar med varandra. Istället handlar den om att de texter som produceras inom de olika disciplinerna så småningom visar på var kommunikationsmöjligheterna finns. Vi kan också hitta många andra exempel på att förståelse över systemgränserna är mycket svår att uppnå. Många av konflikterna i världen ger exempel på detta. Inte minst konflikten mellan västvärlden och den muslimska världen. Ett tydligt och konkret exempel är Israel Palestina konflikten.

Vi kan utveckla denna synpunkt med att Luhmann menar att kommunikation inte handlar om en process som förmedlar färdig mening eller information. Han säger själv att det inte handlar om en aktualisering av en delad uppfattning om mening, som informerar åtminstone en av deltagarna i kommunikationen. Istället handlar det om en fördelning av överraskningar. Inte heller handlar det om att överföra information eftersom informationens identitet skapas vid en särskild tidpunkt och inte något som består över tid (Luhmann,1990c:32). På så sätt bör olika system åtminstone kunna överraska varandra.

Det är viktigt att framhålla att Luhmann menar att språket, som ett gemensamt överenskommet språk, inte är den viktigaste delen i dessa processer. För att kommunikation skall ske räcker det med en rynkad panna eller att några objekt på t.ex. en tallrik arrangeras om (Luhmann,1990c:32-33). Språket är på så sätt alltför bundet till den allmänna uppfattningen om vad mening är. Luhmann framhåller att detta innebär att vi måste ifrågasätta om den verbala kommunikationen är den grundläggande formen. Istället menar han att det är själva skillnaden som måste föras fram och födas i verkliga och pågående erfarenheter. Skillnaden måste vara en del av det som presenteras för medvetandet. Det gör att tidsdimensionen blir viktig eftersom den skapar en skillnad mellan de observationer som görs vid en tid och en annan. Det innebär att det är denna skillnad som uppstår mellan en tid och en annan som skapar den sociala dimensionen och inte den delade meningen mellan olika sociala aktörer. Luhmann menar att en delad mening som manifesteras genom språket, i själva verket gör att tidsdimensionen blir oviktig eftersom det inte kan finnas någon tidsskillnad mellan två subjekt som erfar samma mening (Luhmann,1990c:40). Om vi skall försöka oss på en enkel tolkning av Luhmanns något abstrakta resonemang bör det betyda att sociala interaktioner mellan människor är alltför knutna till nuet, för att vi

verkligen skall kunna se vilka möjligheter framtiden ger oss. De framtida möjligheterna upptäcks bäst om vi använder oss av den sociala kommunikationens artefaktiska möjligheter till synkronisering.

6.5 Det sociala som text/artefakt

Vi har tagit upp att tekniska redskap bland annat består av texter. I den bibliotekskontext som vi har fört diskussionen handlar det om hur texter kan representera en artificiell socialitet. Det är intressant att den luhmannska teorin har använts för att analysera det litterära systemet.

I artikeln *Systems Theory and Literary Studies in the 1990s* (Schmidt&Jahraus,1999) påpekas att Gerhard Plumpe och Niels Werber skiljer mellan litteraturen som ett system och litteraturen som en miljö för andra system (Plumpe&Werber,1993). Plumpe och Werber menar att litteraturen uppfattad som en miljöfaktor blir ett symboliskt genererat medium för kommunikation inom det konstnärliga området. På så sätt skapas litteraturen som en skillnad mellan medium och form Vi bör uppmärksamma att uppfattningen om att de litterära systemen kan fungera som **miljö** för andra system innebär att vi överger analysen av kommunikationen inom ett system. Istället talar vi om en kommunikation som sker mellan ett system och dess omvärld.

Det intressanta med Plumpes och Werbers diskussion är att de försöker göra Luhmanns syn på koden mer öppen. Luhmann menar att vi använder olika koder när vi värderar en distinktion. Koden vacker/ful är den vi använder för att värdera konstnärliga budskap. Men den koden tycks vara mycket knuten till en individs värderingar. På så sätt blir det som att den bara kan kombineras med psykologiska redskap. Plumpe och Werber ersätter Luhmanns kod vacker/ful med koden intressant/tråkig. Koden intressant/tråkig handlar mera om att något kan fånga ett intresse utan en aktiv värdering. I enlighet med diskussionen kring psykologisk relevans bör det som framstår som intressant mer stimulera till att observera något nytt än om något upplevs som vackert som mer har att göra med tidigare värderingar. Därför bör den vara mer förenlig med att det inte behöver finnas något medvetande som kodar eftersom det som är intressant tränger sig på utan att en enhet med komplex funktionalitet först behöver göra en bedömning, medan det som vi upplever som vackert sannolikt först måste kopplas till en människas egna preferenser. Vi har redan uppmärksammat att det riktas kritik mot att Luhmanns teori blir för mekanistisk och konservativ. Det är nog en missuppfattning som bygger på att man inte ser hur viktig relativismen är i hans teori. Det kan också vara en uppfattning som hänger samman med att man inte ser att teorin om

strukturell koppling inte är en teori om att system aldrig kan förändras. Istället är denna del av Luhmanns teori det som förklarar hur system kan förändras, trots att systemreferensen tenderar att bidra till att systemen sällan förändras. Luhmanns teori kan ändå ibland vara lite motsägelsefull. Den omtolkning som Plumpe och Werber vill lansera gör att vi kommer undan en motsägelsefullhet. Med koden intressant/tråkig blir det också lättare att förstå vilken roll de kulturella artefakterna har som självständiga faktorer som kan påverka individernas utveckling. Genom att koden blir mer fri från mänsklig värdering blir det också som om förändringar i koderna kan förmedlas genom tekniska redskap.

Tanken på att litteratur fungerar som miljö för andra system kommer bäst fram i den så kallade Leidenmodellen. I den fokuserar Berg och Prangel på skillnaden mellan texten och kontexten (Berg&Prangel,1993). Här framstår det som om texten har en situationsskapande roll och att den inte är bunden av den form den själv har. Textens mening ges då av skillnaden mellan vad texten säger och vad den förnekar. Schmidt och Jahraus menar att texten blir en händelse som frilägger dess sociala dimension. Textens mening skapas mot den bakgrund som den förnekar. På så sätt får texten bara en mening genom att den temporärt förnekar den kontext som den befinner sig i (Schmidt&Jahraus,1999).

David Roberts nämner i *Self-reference in literature* (Roberts,1999) att Dietrich Schwanitz fokuserar på texter som symboliska strukturer (Schwanitz,1990). Den modellen framstår som ännu mera i överensstämmelse med AI eftersom den innebär att inte heller texten i sig behöver ses som en enhetlig struktur. Texten blir istället något som förmedlar mellan system och omvärld. Schmidt och Jahraus menar att Schwanitzs fråga är hur erfarenhet kan kommuniceras fastän den inte kan observeras som annat än aktivitet:

In both cases, similar techniques of communication are used: a combination of ambivalence, opaqueness, authentication, and paradox. In the communication of love, consciousness and communication diverge: There is the unconscious, but there is also the unspeakable. But also in the novel (paradigmatic example: Tristram Shandy) experience must not be recognized as action because otherwise it would obstruct itself. Consciousness and communication are rendered indistinguishable, as in free indirect speech. Literature can thus make intimate communication public without losing its intimate character. These examples demonstrate how Schwanitz uses systems theory to analyze genuinely literary phenomena without claiming to make statements about the social dimension of literature. (Schmidt&Jahraus,1999:245)

Det finns mycket i detta sätt att se på relationer mellan texter som påminner om hur Merton resonerar i *On the Shoulders of Giants*. Det är rimligt att hävda att den texten handlar om relationen mellan system och den omvärld som litteraturen

representerar. Och Luhmanns teori är till stor del en utveckling av Mertons idéer om oavsiktliga konsekvenser. Det är uppenbart att människor inte upptäcker sådana konsekvenser genom medvetandet.

Det är inte bara inom av Luhmann inspirerad litteraturteori, eller av Vygotsky inspirerad pedagogik, som artefakternas betydelse för kunskapsutveckling poängteras. Också inom mer idéhistoriska studier ser man artefakternas betydelse. Paul Blom har bland annat i artikeln *Intention, history, and artifact concepts* tagit upp vilken betydelse bilder, prototyper och former har (Bloom,1996:6,8,12).

Den diskussionen pekar fram till att sociologer också har funderat på metaforernas roll. De tyska sociologerna Sabine Maasen och Peter Weingart har i artikeln *Metaphors - Messengers of Meaning* (Maasen&Weingart,1995) hävdad att kunskapsförmedling bör studeras genom att vi använder oss av begreppet metafor såsom det har använts inom litteraturvetenskapen. Genom att introducera en okänd analogi i en ny kontext kommer både metaforerna och kontexterna att förändras. Dessutom kommer en sådan metod i kombination med diskursanalys att leda till att vi frilägger den institutionaliserade mekanism genom vilken speciella kunskapsstorheter utväljs. Det bör medge att metaforen är något som är laddat med social funktionalitet. Den får en funktionalitet som ligger utanför det som orden betyder.

Maasen & Weingart citerar bland annat Luhmanns användning av metaforer som en del av semantiken. Det handlar om en semantik som möjliggör att texternas betydelse skapar en lösare struktur än vad som är fallet om vi ser texterna som en komposition av låsta begrepp. En sådan lösare struktur exponerar möjligheter istället för att representera en fast organisation av erkänd kunskap.

En annan text som de citerar är en som är en klassiker inom området, George Lakoffs och Mark Johnsons *Metaphors we live by* (Lakoff&Johnson,1980). Maasen & Weingart ställer bland annat frågan hur den nya mening som metaforerna förmedlar stabiliseras, och svarar med hänsyftning till Lakoff & Johnson att det handlar om en process som innehåller flera diskurser. En metafor som har hittat in i en diskurs blir på så sätt attraktiv för andra diskurser (Maasen&Weingart,1995:26). Metaforernas kraft kommer då inte direkt från människorna utan de måste stabiliseras under ett flertal diskurser, vilket naturligtvis också understryker att kommunikationen inte är direkt avhängig av att personer samtalar med varandra. Lakoff & Johnson har senare utvecklat sina idéer om metaforernas betydelse. I boken *Philosophy in the flesh: the embodied mind and its challenge to western thought* (Lakoff&Johnson,1999) har de dessutom anknutit detta tänkande till frågan om AI. De hävdar i den boken att det inte finns något medvetande som är skilt från kroppen, liksom det inte finns några tankar som existerar oberoende av kroppar och hjärnor. På så sätt är de nog motståndare till stark AI. De framhåller dock att kognitionsvetenskapen har lärt oss att medvetandet inte bara består av detta. Via metaforen ges medvetandet en insida men det får också en utsida. Frågan är väl bara om detta betyder att kontexten är

helt beroende av medvetandet eller om det kan ha en sådan autonomi att vi också kan hävda att medvetandet är beroende av kontexten.

6.6 Den sociala kommunikationen och handlingsbegreppet

Vi har använt Luhmanns teori om social kommunikation till att argumentera för möjligheterna av stark AI. Vi har stärkt de argumenten genom att peka på redskapens betydelse för kunskapsutveckling. Vi har föreslagit att texter är något mer än mänskligt handlande men vi har inte riktigt klargjort att kunskapsprocesser hör samman med social kommunikation framom mänskligt handlande. Vi återkommer i ett senare kapitel till en jämförelse mellan Luhmanns kommunikationsteori och Habermas talhandlingsteori. Här skall vi se hur Luhmann själv motiverar att kommunikation bara sker inom sociala processer. Vi kan börja med att redovisa att Luhmann ser att de sociala systemen existerar i sig själva och att de inte bara är en avspeglning av sociala interaktioner mellan individer. Raf Vanderstreten understryker att tankeverksamheten inte är det centrala för hur de sociala systemen fungerar (Vanderstraeten,2000:582). Den medvetna tankeverksamheten använder vi för att fundera över det system som vi redan är operativt slutna i. För att vi skall kunna nå fram till en kognitiv öppenhet som exponerar oss för möjligheterna att expandera vår kunskap behöver vi kommunikation.

Luhmann själv stödjer sin uppfattning om att det är sociala system som kommunicerar med att det inte finns något som talar för att medvetna hjärnor kan åstadkomma kommunikation.

Neurophysiological studies, difficult enough in themselves, are not very helpful in terms of the mind. We no longer assume, as did the ancient Indians, that we can alter conditions by concentrating. The more common idea that the mind effects physical behavior or even communication is equally mysterious. The assumption that this occurs is nothing more than an observer's causal attribution. We have to start any clarification with the observer. Once this is recognized, what follows is the question whether and - if the answer to that is affirmative - how the mind participates in communication. The fact that the mind does participate is undisputed, since communication could not exist without the mind, just as life could not exist without a molecular organization of matter. But what is participation? (Luhmann,2002:169-170)

Utifrån vårt perspektiv är frågan om detta deltagande handlar om att man tillför en funktionalitet eller om man bara deltar enligt en given funktionalitet.

Eftersom det är kommunikation som är social medan handling inte är social, innebär det att sociologin behöver ersätta handlingsteori med kommunikationsteori (Luhmann,1990b:6). Vilket också innebär att relationen mellan handling och kommunikation måste ställas på huvudet. Det blir då kommunikation som är en förutsättning för handling och inte tvärt om eftersom endast kommunikation kan förändra systemreferensen. Varken psykologiska motiv, samtal eller kapacitet för att argumentera föregår handling. Endast genom den kommunikation som länkar valmöjligheterna och gör dem hanterbara kan fortsatt kommunikation komma till stånd (Luhmann,1990b:6-7). Det leder också till att sociologin bör lämna betoningen av subjektet eftersom subjektet är en del av helheten och kan således inte vara det som helheten skall baseras på (Luhmann,1990d:113-114).

Luhmann menar också att det faktum att samhället inte består av kroppar och medvetanden, innebär att förändringar i media och i teknik inte är marginella. Det antyder naturligtvis att det finns en potential också i tekniska redskap. Luhmann framhåller istället att samhället består av ett nätverk av kommunikation:

Therefore, if media and techniques of communication change, if the facilities and sensitivities of expression change, if codes change from oral to written communication ' and, above all, if the capacities of reproduction and storage increase, new structures become possible, and eventually necessary, to cope with new complexities (Luhmann,1990a:100).

Luhmann har själv i en artikel i antologin *The Micro-Macro Link* kopplat diskussionen om förhållandet mellan mikro- och makro-nivåer till kommunikationsmöjligheterna. Han förklarar att den traditionella uppfattningen om sociala system behöver göra en skillnad mellan sådana system som har monopol över kommunikation och sådana som inte har det. Han menar att sociala system fungerar som samhällen om de innehåller alla operationer som handlar om kommunikation. Systemen innehåller då allt. Om kommunikationen finns i systemens omgivning fungerar de bara som interaktioner (Luhmann,1987:114). Det är inte fullvärdiga sociala system. Det leder till att psykiska system som människor inte kan bilda sociala system eftersom kommunikationen finns i det samhälle de ingår i. Människor kan på så sätt bara interagera med varandra inom en given systemreferens. Vi kan således inte förklara att sociala system uppkommer genom individuella handlingar.

Detta betyder inte att kommunikation är oberoende av människor. Människor deltar i kommunikationen men de kan inte styra eller determinera kommunikationen kausalt. Människor är inte orsaken till kommunikation. Det är

endast kommunikationen som kommunicerar. Att människor kommunicerar handlar bara om ett påstående som är en konvention. Kommunikationer och inte handlingar är de minsta beståndsdelarna i det sociala eftersom det i en kommunikation deltar minst två människor, eller två psykiska system. Handlingar är däremot inte element i sociala system utan en produkt av sociala förutsättningar (Nassehi&Kneer,1997:92-94).

Skillnaden mellan begreppen interaktion och kommunikation blir inte helt klar. Luhmann använder begreppet interaktion på olika sätt. Oftast benämner han de samtal som personer har med varandra som interaktioner till skillnad från kommunikationer. Kommunikationer sker bara inom sociala system. Däremot tycks Luhmann också benämna relationer mellan olika sociala system som interaktioner. Det handlar just om att han är oklar över om olika sociala system kan kommunicera med varandra. Det bör vara så att det inte bara är inom ett socialt system som kommunikationen kan ske, utan att också olika artefaktiska system kan påverka varandra i en process som påminner om kommunikation.

Formens betydelse kunde i en bibliotekskontext handla om att texterna får en egen kraft beroende på deras positioner. Detta kan exemplifieras med en tanke som tillskrivs Foucault:

The frontiers of a book are never clear cut: beyond the title, the first lines, and the last full stop, beyond its internal configuration and its autonomous form, it is caught up in a system of references to other books, other texts, other sentences: it is a node within a network (Radford,1992:419).

Det imaginära formas på detta sätt inte i opposition till verkligheten:

it grows among signs, from book to book, in the interstice of repetitions and commentaries; it is born and takes shape in the interval between books. *It is a phenomenon of the library.* (citrat från Foucaults text "Fantasia of the Library", I: "Language, Counter-Memory, Practice") (Radford,1992:420)

Vi kan tillägga att det ändå bör gå att ge individernas deltagande i kommunikationen en mer formell funktion än att det bara handlar om en konvention. Om vi istället för att uppfatta individerna som enheter som bygger sin roll på en komplicerad funktion, ser dem som en form av mänskliga noder som liksom artefakter har en enkel funktionalitet, kan vi se dem som omvärld till de sociala systemen, alternativt noder inom de sociala systemen. Det kan vara så att deras kapacitet för komplext tänkande i vissa avseenden gör att de kan reta systemen mer än kulturella artefakter kan, t.ex. genom att kritisera rådande

systemreferens. Men det vi kallar konstruktiv kritik bör i första hand handla om att man försöker visa på nya möjligheter att se på problem. Då bör det inte så mycket handla om en kapacitet för kritiskt tänkande utan om att en annan människa tack vare sin sociala position alltid ser på saker utifrån en annan observatörsposition. Det centrala är då vad som ges av situationen. Om intelligens handlar om att expandera är inte det centrala att medvetna individer tänker kritisk. Vi bör istället se på den människa som ger konstruktiv kritik som någon som fungerar som ett redskap för andra människor och att den funktionalitet som då är inblandad, på grund av kontingensen, inte kan vara särskilt komplicerad. Om vi menar att artefakten tillför funktionaliteten i den sociala kommunikationen bör också människan kunna spela denna roll.

6.7 Redskapsanvändning och konnektionism

Vi skall avrunda detta kapitel genom att återvända till Wertsch. I boken *Mind as Action* för Wertsch ett resonemang om förhållandet mellan människa och redskap i de sociala processerna. Wertsch diskuterar en viktig sak som har att göra med redskapens förmåga att kunna fungera självständigt. Nämligen frågan om i vilken mån människan måste kunna internalisera kunskapen om hur man använder redskapen. Den diskussionen har relevans för frågan om vi måste se på redskapsanvändning som ett led i en process som styrs av mänsklig handling eller som en aktivitet i ett system. Om vi inte behöver internalisera i vårt medvetande hur vi skall använda redskapen bör det också tala för att skillnaden mellan psykologiska och tekniska redskap inte är så stor. Wertsch uppfattar nog den form av decentraliserad interaktion som blir resultatet av att redskapen ges en mer självständig status, som ett led i en handlingsorienterad process. Wertsch argumentation pekar emellertid på att det finns många detaljer kring redskapsanvändandets förhållande till den mänskliga handlingen som är motsägelsefulla.

Wertsch ser tänkandet som den decentraliserade interaktion som blir följden av att vi använder förmedlande redskap (Wertsch,1998:20-21). Frågan som blir aktualiserad av Wertsch resonemang om förmedlande redskap är snarast vilken roll dessa redskap har i tänkandet (Wertsch,1998:24).

Den centrala frågan för Wertsch är hur man lär sig behärska de förmedlande redskapen. Det är här som vi kommer in på om externa processer är inblandade i det som vi vanligen uppfattar som ett tänkande som utförs med människans hjärna. Wertsch menar att individernas användning av redskap i utvecklingsprocesser brukar ses som en slags internalisering. Alltså en process där hjärnan använder redskap för den tankeprocess som ändå sker inom hjärnan. Om

man först tvingas internalisera i sitt medvetande hur AI-systemen fungerar, så är det svårt att hävda att dessa kan fungera frikopplade från mänsklig aktivitet. I så fall är de bara redskap som människan kan använda för sitt eget tänkande. Wertsch menar att termen internalisering är missledande (Wertsch,1998:48).

Wertsch menar att det som man bedömer som internalisering för det mesta handlar om sådant som **inte** sker på ett internt medvetandeplan. Forskning om kognition visar att mycket handlar om processer som aldrig har varit ägnade att internaliseras. Wertsch refererar till boken *Cognition in the Wild*. I den diskuterar Edwin Hutchins den redskapsanvändning som behövs för att navigera in ett stort fartyg i en hamn. Wertsch tar upp att Hutchins analys visar på att det som vi uppfattar som internalisering inte är det. Det leder till att Wertsch hellre talar om "knowing how" och "mastery" än om "internalization" (Wertsch,1998:50-51).

Wertschs argumentering kan användas för att hävda att redskapen måste kunna utföra något som liknar "tänkande" om vi skall kunna använda dem. Vi kan inte internalisera kunskapen om hur de fungerar utan vi måste lita till att redskapen istället kan "tänka" så mycket att de kan interagera med oss. Det är inte helt klart om det innebär att de samtidigt är kapabla att klara av de sociala processer som är inblandade i kunskapsskapandet. Om social kapacitet handlar om förmåga att interagera med människor bör detta vara möjligt eftersom vi här diskuterar en situation där initiativet kommer från datorerna. Vi bör då se interaktion som en relation där människan eller datorn kan ta initiativet i en interaktion. Frågan är om det faktiskt är rimligt att tala om att redskapen kan tänka. Det luhmannska synsättet på sociala processer, som tonar ner tänkandets roll och som istället betonar möjligheterna att observera, ger en naturligare förklaring till att redskapen har stor betydelse för kunskapsprocesserna fastän människorna inte fullt ut har lärt sig att hantera dem. Redskapen behöver inte vara kapabla att tänka för att kunna interagera eftersom interaktionen mellan människor inte kan bygga på särskilt komplicerade processer p.g.a. den dubbla kontingensen. Redskapens betydelse kommer istället av att de är delar av det som bildar den funktionalitet med hjälp av vilken den sociala kommunikationen fungerar. De hjälper till att skapa nya observationsmöjligheter. Där har vi ännu ett möjligt svar på frågan om varför redskapen ändrar hela processen kring skapandet av ny kunskap. Det är för att de ger oss större möjligheter att observera och därmed gör det lättare att nå den komplexitet som finns i vår omvärld.

Wertsch insikt om det begränsade behovet av internalisering leder honom till en positiv uppfattning om de konnektionistiska modellernas kapacitet. Han stödjer sig på en av de mer framträdande företrädarna för konnektionistiska modeller, Andy Clark (1993), som vi redan har mött i samband med begreppet *Cyborg*. Med stöd av Clark hävdar han att konnektionismen erbjuder en beskrivning och en analys av mentala processer (Wertsch,1998:51). Wertsch menar att konnektionistiska modeller är mycket suggestiva när det gäller att förstå de färdigheter som är inblandade i förmedlande redskap och hur dessa färdigheter kan

utvecklas. Han framhåller att konnektionismen hjälper oss att utveckla system som bygger på kunskap om hur vi skall använda kulturella redskap utan att vi behöver placera in explicita regler i huvudet på aktörerna (Wertsch,1998:52).

Wertsch försöker i *Mind as Action* hitta ett alternativ till metodologisk individualism och social reduktionism. Wertsch menar att sociokulturell analys bör försöka förstå relationen mellan mänsklig handling, inklusive mentala funktioner, å den ena sidan och kulturella, institutionella och historiska sammanhang å den andra sidan. Han menar att oftast leder vår vana att tänka i konträra termer oss till att uppfatta denna relation som antingen eller. Metodologisk individualism försöker sålunda få oss att anta att de kulturella, institutionella och historiska kontexterna kan förklaras med egenskaper hos individerna. Och social reduktionism försöker få oss att anta att individers beteenden kan förstås med sociala fakta.

Hans råd för att råda bot på sådana motsättningar är att försöka hitta ett sätt för att leva i mitten. Det betyder att vi måste försöka hitta en analysenhet som fokuserar på hur dessa krafter får en dynamisk kontakt med varandra.

Rather than seeking the key to individual mental processes in sociocultural setting, or vice versa, I have argued that we should employ a unit of analysis that focuses precisely on how these forces come into dynamic contact. Although there are several conceivable ways to formulate this issue, I have done so by examining mediated action, defined in terms of an irreducible tension between cultural tools and active agents. Following Vygotsky, I argue that this unit may be profitably examined by invoking more basic analytic elements - namely, agents and mediational means - but it cannot be reduced to either of these elements in isolation. Following Burke, I argue that these elements must be understood as aspects, or dimensions, of mediated action rather than as independently existing essences (Wertsch,1998:179-180).

Kenneth Burke som Wertsch nämner sist i citatet återkommer vi till i kapitel 9.

I detta kapitel har vi gått närmare in på hur människorna använder redskap för att bete sig intelligent. Vi har visat att redskapsanvändandet sedan länge varit något som intresserar pedagogiska teorier och att bl.a. Vygotsky knöt detta redskapsanvändande till darwinistisk utvecklingsteori. Vi har visat att Vygotsky var inne på att sociala processer externa från det mänskliga tänkandet skapade förutsättningar för att människornas kapaciteter för intelligens kunde utvecklas. Däremot har vi också visat på att det blir lite oklart hur dessa sociala processer hos Vygotsky ser ut och att de går att utveckla med hjälp av Luhmanns systemteori.

Vi har också visat på att intelligens inte bör diskuteras i samband med problemlösning. Istället skall vi diskutera intelligens som en förmåga att expandera

vår kunskap utanför det som vi redan känner till. Vi har också visat att det leder till att vi måste se att intelligenta processer först och främst handlar om att de sociala systemen kan interagera med sin omgivning. För människan betyder detta att det viktigaste inte är att människorna kan internalisera i sina medvetanden hur de skall använda de olika redskapen. Vi har också visat att detta ger en möjlighet att utveckla samhällsvetenskapens uppfattning om mikro- och makro-processer.

7. Social kommunikation som personlig interaktion

Vi såg i förra kapitlet att det finns anledning hävda att stark AI kan fungera eftersom kunskapsprocesser handlar om kommunikation och sådana processer behöver materiella funktioner. Det påståendet är nog så kontroversiellt att vi bör jämföra det med den kritik mot AI som kommer från en sociolog som betonar att kommunikation måste vara knutet till personlig interaktion. Vi skall därför ta del av en kritik av AI från Johan Asplund som definitivt knyter de sociala processernas meningsskapande till den interaktion som uppstår då människor samtalar med varandra.

Asplund argumenterar utifrån social interaktionism. Vi skall kontrastera hans resonemang med uppfattningar som kommer från en annan social interaktionist som mer betonar att interaktioner är så bundna till en strukturell formalism att människornas interaktioner med varandra tenderar att begränsas till vad dessa strukturer tillåter. De synpunkterna kommer från Susan Leigh Star.

7.1 Personlig interaktion och AI

Asplund bygger i boken *Genom huvudet* upp sin kritik mot AI genom att argumentera mot idén om att kognitionen kan vara distribuerad, vilket vi har sett är en av grunderna för DAI-strategier. Asplunds argumentation handlar om den sociala process som skapar det icke-formaliserade. Asplund fokuserar på att denna funktionalitet uppkommer när människor samarbetar. Argumentet är inriktat på att kontexterna inte på egen hand kan skapa mening. Asplund argumenterar också uttryckligen mot distribuerad kognition vilken är den bärande tanken i en DAI-modell. Han hävdar att föreställningen att tänkande och problemlösning sker i huvudet utgör kärnan i ett mycket gammalt och livskraftigt paradigm.

Under sin långa historia har detta paradigm naturligt nog emellanåt ifrågasatts. Den "solipsistiska" uppfattningen om tänkandets natur har dock inte rubbats. Det är uppenbarligen mycket svårt att på allvar betvivla det sanna och riktiga i att tänkandet sker i huvudet. Ifrågasättanden av bilden "i huvudet" har aldrig, vad jag vet, utvecklats därhän att man kunde skönja ett fungerande alternativ till bilden ifråga. Det senaste ifrågasättandet går under beteckningen "distributed cognition". Det centrala påståendet är att tänkandet består i eller uppstår ur interaktionen mellan individer samt mellan individer och diverse redskap eller artefakter. Tänkandet är ett emergent fenomen på kollektiv nivå. Detta påstående kan se lovande ut men visar sig i realiteten vara tämligen intetsäggande. Edwin Hutchins och hans efterföljare har empiriskt studerat olika former av teamwork, t.ex. navigering av ett fartyg och verksamheten i förarkabinen i ett flygplan. I sådana sammanhang handlar det ovedersägligen om flera individer som interagerar med varandra samt med allehanda redskap eller instrument. Navigering och flygning är, närmast per definition, kollektiva aktiviteter som innefattar ett "utspritt" informationsutbyte och beslutsfattande. I sitt val av studieobjekt riskerar anhängare av "distributed cognition" att förutsätta det som skulle bevisas. Bilden "i huvudet" behöver ingalunda implicera ett förnekande av möjligheten och existensen av kognitivt samarbete. Härtill kommer att Hutchins m.fl. har övertagit centrala begrepp och postulat från det förhärskande paradigmet, dvs. begrepp och postulat som är förankrade i bilden "i huvudet", vilka de sedan utan vidare applicerar på olika former av teamwork. Detta är knappast sättet att genomföra eller ens påbörja ett paradigmskifte. (Asplund, 2002b:12-13).

Asplund bygger sin framställning på att försvara begrepp som kan förefalla självklara. Men det går ändå att hävda att bilden av tänkandet i huvudet bara avspeglar det faktum att människan enbart kan uppleva något med hjälp av sitt eget huvud. Medvetandet kan naturligtvis inte uppleva de processer kring skapandet av ny kunskap som inte sker i huvudet. Bilden i huvudet kan därför ses som en slags naivistisk tolkning av vad som ligger bakom kunskapsprocesserna. Denna metafor ligger sannolikt bakom idén om att AI-system bör konstrueras som artificiellt tänkande aktörer. Det finns en stark betoning av att det någonstans måste finnas något som motsvarar det mänskliga tänkandet.

I detta sammanhang är det värdefullt att uppmärksamma att bilden i huvudet har att göra med ett stort metodologiskt problem i diskussionen om stark AI. Frågar vi människor hur de har kommit fram till kunskap, kan de bara redogöra för hur de själva upplever att de kommit fram till kunskapen. Det naturliga svaret blir då att det skett genom tankeprocesser, eftersom vi människor inte kan uppleva

dess processer på annat sätt än genom hjärnan. Det här är något som inom AI-forskningen brukar benämnas *The Frame Problem* (Korb,1998),(Haselager&Rappard,1998). Skall vi kunna tänka oss att dessa processer handlar om något utanför människan måste vi först börja se de anomalier som denna bild av tänkandet i huvudet för med sig. Först då kan vi vara mogna att pröva andra vägar. Och först då kommer vi i en situation som medger att vi kan se poängen med en DAI-strategi.

Ser vi närmare på hur Asplund argumenterar så framstår han som mycket kluven till bilden i huvudet. Asplund säger egentligen inte vad han anser att intelligens är, men det framstår som att han argumenterar för att intelligens består av sociala interaktioner mellan ett antal personer som alla tänker. Intelligens kräver att många människor kommunicerar sitt tänkande. Vi bör därför tolka Asplund som att intelligens inte utvecklas genom solitärt tänkande, men intelligensen kan samtidigt inte vara utan det tänkande som de olika individerna i sin ensamhet utför. *Föreställningen "i (ett) huvud" utgör i och för sig inget nät men den är den första noden i det nät som har vuxit sig allt större och blivit allt finmaskigare i kognitiv psykologi* (Asplund,2002b:11).

Utifrån Luhmanns sätt att se på kommunikation bör man kunna hävda att det som sker i huvudet kan vara en nod i en kommunikativ process men den behöver inte nödvändigtvis vara den första noden. Det är viktigt att framhålla att Asplund här argumenterar mot grundläggande antaganden hos GOFAI-modellen. Men vi har sett att inte ens en robotforskare som Rodney Brooks menar att aktören behöver vara kapabel till komplext tänkande. Och DAI-modellen bygger på att funktionaliteten i noderna inte behöver vara komplicerad (Anderson,2003:96,100,102).

I sammanhanget kan vi peka på en fråga som både Michael Polanyi och Gregory Bateson har ställt och som Tarja Susi redovisar (Susi,2006:21,26). Frågan handlar om var gränsen för den blinde mannens eget system går när han använder en vit käpp för att ta sig framåt. Innehåller detta system käppen eller inte? Batesons svar är att frågan om systemet avgränsas av huden eller av någon artefakt som människan använder är en nonsensfråga. En viktigare fråga är vad man försöker förstå och därmed vilken analysenhet man skall rikta in sig på. I denna avhandling argumenterar vi för att analysenheten skall vara det sociala systemets kommunikation och inte tänkandet hos psykiska system.

Asplunds kluvenhet visar på hur krånglig frågan kring AI är. Han bygger sin argumentation på många antaganden om medvetandets funktioner som vi redan har sett att det finns oklarheter kring. Asplund använder dessa oklarheter till att angripa AI, men det är lika rimligt att hävda att Asplunds ståndpunkt blir grumlig därför att han inte kan visa hur den sociala kommunikation som förmedlar tänkandet mellan människorna fungerar.

Det är märkligt att det kan finnas ett kognitivt samarbete om inte tankarna kan förmedlas genom någon typ av artefakter. I så fall måste tankarna flyga till

människorna från hjärna till hjärna. Som vi tidigare varit inne på kan en människa säkert med sin tankeverksamhet skapa en omvärld för en annan person. Denna tankeverksamhet kan dock knappast den andra personen ha insyn i. Den ene människan kan bara observera de uttryck den andre med sin tankeverksamhet levererar. Den ene människan kan säga något som den andre bara kan tolka från sin observatörsposition. Tänkandet hos den första personen är då bara ett sätt att skapa kontexter. Det är inte tänkandet som en annan människa från sin observatörsposition uppfattar som meningsfullt. Det som ger mening är istället den kontext som tänkandet skapar. Detta gör att vi bör nedtona tänkandets betydelse och se att den kontextskapande processen också kan fungera t.ex. genom att konnektionistiska system ändrar relationerna mellan sina noder. Det viktiga är kommunikationen, och den är beroende av att sociala processer kan exponera möjligheter. Det låter osannolikt att dessa möjligheter kan manifesteras bara genom det som sägs mellan människor. För att de skall ha någon räckvidd och någon beständighet måste de också kommuniceras med hjälp av artefakter. Vår kultur, inklusive vetenskapen, är beroende av att vi kan kommunicera på distans. Det centrala i detta är att de kommunikativa processerna måste omtolkas. Om de mänskliga insatserna begränsas till att skapa kontexter så kan vi heller inte hävda att kommunikativa processer i första hand har att göra med medvetna talhandlingar.

Asplunds ståndpunkt ger uttryck för samma motstridiga uppfattning kring det sociala som Harry Collins. De tar bäge avstånd från att intelligens är något förkroppsligat i människan själv, och hävdar istället att den består av en egenskap hos sociala grupper. Men den kan samtidigt inte handla om något artificiellt eftersom det sociala alltid måste grundas i mänskliga handlingar. På så sätt bygger de sina positioner på cirkelresonemang. Både Asplund och Collins kan till slut inte förklara hur den sociala interaktion som utvecklar kunskapen fungerar. Också Luhmann knyter kunskapsutveckling till kommunikation men hos honom hittar vi faktiskt en mer utarbetad förklaring till hur den processen ser ut.

I en diskussion om datorers förmåga att slå stormästare i schack (vilket de faktiskt idag kan) kritiserar Asplund ett påstående av Dreyfus att en skicklig schackspelare förlitar sig på "intuition" eller "instinkt" snarare än beräkningar. Han framhåller istället att schack spelas av två med varandra interagerande spelare. *Schack är ett kognitivt spel men också en dialog* (Asplund, 2002b:80).

Asplund tolkar exemplet enligt sin åsikt att intelligens är något socialt samtidigt som tänkandet inte kan ske någon annan stans än i människan. Exemplet kan dock lika väl tolkas som att intelligens inte kräver personlig kommunikation eftersom en duktig schackspelare knappast visar sina avsikter. Det kan istället tas som ett exempel på att kommunikation inte alltid bygger på förmedling med ett språk som människor gemensamt förstår. Ett exempel från ett spel där motparten har all anledning att inte kommunicera sina rätta avsikter ger alla möjligheter att illustrera att kommunikation även med människor, inte alltid sker under förhållanden där människans medvetande har de bästa förutsättningarna för att

förstå! Motspelaren måste alltså läsa sig till den andres avsikter genom de regler och strategier som styr schackspelandet. Den tolkningen säger att det handlar om att motspelaren försöker göra det mesta av situationen. Dessutom handlar schackspel också om att kunna hitta på egna kreativa drag. Det bör handla om att kunna skapa nya situationer. Visserligen är schackspel inte något speciellt bra exempel på distribuerad kognition eftersom det är styrt av regler som är givna. Därför handlar det inte i första hand om att hitta på något nytt utan om att tänka sig fram till hur givna regler skall utnyttjas. Vi bör observera att dessa regler är tämligen enkla. Den intelligens vi behöver för att kunna hitta på något nytt, bör då handla om att reglerna blir kontextfaktorer, vilka vi kan utgå ifrån för att kunna hitta på nya sätt att lägga upp en strategi för att vinna partiet. Den som själv har spelat schack vet att tjugningen med spelet just handlar om att situationer ger möjligheter att utveckla strategin. Vi kanske kan uttrycka det som att kommunikation skall exponera möjligheter istället för att vara en del i tänkandet.

Vi kan också ta schackspelet som ett exempel på att regler ibland bara skapar en rätt så enkel funktionalitet. När vi diskuterar skillnaden mellan symbolisk AI och konnektionistisk AI handlar det inte alltid om att den ena är helt beroende av regler medan den andra inte alls styrs av regler. Skillnaden är antagligen bara att processerna i symbolisk AI är helt förutbestämda av regler, vilket gör att de som kontextfaktorer blir oföränderliga. I en konnektionistisk modell är istället reglerna skriva så att de skall förändras. På så sätt skapar de kontextfaktorer som hela tiden förändras.

Vi kan hävda att regler och strategier måste vara skapade av människor. Vi kan då också hävda att schackspelare inte styrs av motspelarens tänkande utan av andra människors tänkande under århundraden av schackspelande. Det skulle då kunna vara något som visar att kulturella strukturer måste skapas av människor. Asplund argumenterar för sin sak genom att diskutera intelligens som problemlösning. Problemlösning är också kommunikation enligt Asplund: *Jag har beskrivit problemlösningsprocessen som en kommunikationsprocess som går fram och tillbaka i bestämda turer mellan problemställare och problemlösare* (Asplund,2002b:46). Också schackspelande får vi anta att Asplund vill att vi skall se som ett exempel på problemlösande. Här står vi emellertid inför ett fall som karaktäriseras av den begränsning kring intelligensens innebörd som uppstår genom att vi diskuterar begreppet i samband med problemlösning. Problemlösning är, som Engeström säger, något som förutsätter en given kontext. En given kontext är säkert en av människor bestämd kontext. Schackspelet representerar i hög grad en sådan given kontext. Synen på intelligens blir annorlunda om vi istället diskuterar intelligens som det beteende vi behöver för expanderings. Alltså det beteende som strävar efter att gå utöver de givna kontexterna. Det säger oss att människorna behöver det sociala just för att deras medvetanden är begränsade. Då blir det svårt att som Asplund hålla fast vid att intelligens måste hänga samman med solipsistiskt tänkande samtidigt som det är ett socialt beteende. Asplunds

försök att grunda det sociala i tänkandet strandar ändå på att han återvänder till psykologiska förklaringar till de sociala processerna. Han beskriver kanske hur en gemenskap byggs upp mellan två eller flera människor, men han beskriver inte hur dessa människor också agerar utifrån den kunskap som finns lagrad i litteratur och andra humanistiska artefakter. Sådana artefakter som expanderar kunskap istället för, liksom schackspelets regler, bestämmer vad som gäller. Det finns skäl anta att sådana strukturer som människor tillsammans tänker sig fram till är just sådana som definierar vad som är tillåtet.

Asplunds argumentation håller bara gentemot sådana AI-strategier som bygger på att skapa artificiella aktörer. Sådana lösningar kan inte fungera eftersom argumentationen visar på att individuella aktörer är beroende av att kunna tänka tillsammans för att interaktionen dem emellan skall leda till att de kan lösa de problem de samtalar om. Asplund blandar dock ihop olika scenarion genom att han så knyter upp sin diskussion till problemlösning. Artificiella aktörer kan jämföras med mänskliga aktörer när den kontext i vilken problemlösningen skall ske, är styrd av regler. När vi talar om det som sker i expanderande situationer är det sannolikt att både mänskliga aktörer och artificiella aktörer kommer till korta. I sådana situationer är det istället egenskaper i kontexterna som är avgörande. Och då finns det mindre anledning att anta att människors kapaciteter är större än artefakternas. Sannolikt så utvecklas också kunskap i en interaktion mellan olika personer, men det beror inte på kapacitet för komplext tänkande inom dessa individer utan, som vi tidigare föreslagit, att varje individ representerar en relation till de artefakter med vilka den sociala kunskapen representeras, eller själv fungerar som en mänsklig artefakt. Denna interaktion mellan två eller flera personer handlar då om att olika sociala kunskapsrelationer omstruktureras genom dessa personers handlingar. Att bete sig pedagogiskt är ett exempel på en situation där människan bättre förmedlar ett budskap genom att framföra det enkelt än genom att försöka erbjuda komplicerade och djupsinniga förklaringar. Och eftersom det är mycket annat än dessa handlingar som kan omstrukturera dessa kunskapsrelationer, behöver inte den sociala kunskapsutvecklingen vara knuten till dessa mänskliga handlingar.

Asplund anger i sin bok att den har en lång och svindlande tillkomsthistoria. Han säger också att han känner sig långt ifrån nöjd och att han helst skulle vilja omarbota allt ytterligare minst en gång (Asplund, 2002b:7). Det framstår ändå som om Asplund har lockats till förhastade slutsatser därför att hans egen interaktionistiska teori är så kopplad till att människor samtalar personligen med varandra. Det framstår som om hans egen, och riktiga, uppfattning om att intelligent beteende hör samman med kommunikation då blir svår att försvara. Det är som om hans egen teori hindrar honom att se att detta kommunikationsperspektiv bara kan försvaras om man hävdar att det är sociala system som kommunicerar. Och som hindrar honom att se att det är en argumentation som talar för distribuerad kognition.

7.2 Sociala system som statiska eller föränderliga

Den andra symboliska interaktionisten, Susan Leigh Star, framhåller att det är vår fastlåsning vid det gamla som gör att människornas valmöjligheter ofta minskas på ett drastiskt sätt när vi inför tekniska lösningar. På så sätt förstärker dessa tekniska lösningar en negativ socialitet istället för att förlösa en positiv socialitet.

I ett kapitel med titeln *Power, technology and the phenomenology of conventions* i boken *Sociology of monsters* menar Star att en viktig del i analysen av tekniska system handlar om att försöka se hur de nätverk som skapar systemen stabiliseras över en längre period. Det finns en del förändringar som inte kan göras ogjorda. Till exempel kan inte det ursprungliga valet av rött som stoppsignal på trafiksignaler bytas ut (Star,1991:40). Star menar att detta förhållande har uppkommit genom att människor har byggt in dessa begränsningar i systemen. I fallet med stoppsignaler handlar det dessutom tydligt om att det är människans uppfattning om att rött betyder stopp som har lett till att systemet fått denna egenskap. Systemets referensram tillåter således bara att rött betyder stopp.

Star framhåller att när väl ett arrangemang blivit standard är det svårt eller omöjligt att etablera alternativa standarder. Ofta krävs det att ett alternativt samhälle av en eller annan orsak etableras (Star,1991:41). Det kan tolkas som att det är oppositionella människor som måste visa vägen till ett alternativ. Men det kan också tolkas som att det är människorna som ser till så att ingen omstrukturering sker, och att det är teknologiska förändringar som gör att några människor får förutsättningar att skapa ett nytt samhälle. För att alternativ skall kunna uppstå måste de komma via nya system. Resonemanget kan användas för att hävda att teknikens nackdelar kommer från människan, därför att vi människor strävar efter att kontrollera vår omvärld genom att göra den endimensionell. Och när tekniska förändringar har kunnat införas beror det på att en ny situation har tillåtit förändringar att ske. Den som följer med informationsteknologins utveckling idag ser nog ett exempel på detta. När en teknologisk nivå som webben har etablerats, skapar den en situation utifrån vilken vi kan se hur vi också kan använda tekniken för att utveckla andra tjänster. Till exempel till att publicera och förmedla elektroniska texter.

Stars resonemang visar på att det är nödvändigt att vi, samtidigt som vi kritiserar den konventionella användningen av datortekniken, även måste vara mer observanta på att den nya tekniken också kan ge oss en chans att bryta upp gamla strukturer. Den konnektionistiska varianten av AI kan vara en sådan teknik. Därför bör vi åtminstone ta den på större allvar än att bara avfärda den som något som standardiserar kunskapen. Stars resonemang ger en insikt om att en negativ attityd till AIs kapacitet inte bara beror på en negativ känslomässig attityd. Den kan också ha att göra med att vi hittills inte har haft en teknik som kan låta oss få erfarenhet av att AI faktiskt kan vara något mera än ännu en symbol för överdrivna drömmar

om den nya datorteknikens välsignelser. Googles framgång visar inte minst vilken betydelse en ny teknik faktiskt kan ha. Vi kan nog sammanfatta AI-teknikens utveckling på så sätt att dess tidigare bindning till GOFAI innebar stora risker för att tekniken blev ett redskap för olika kunskapsintressen. Stark AI kan ses som skrämmande därför att den kan ta över mycket av sådant som vi har uppfattat att bara människor kan göra. Men vi bör inte glömma den stora emanciperande kraft som denna teknik kan ha just därför att den blir fri från mänskliga intressen.

Star har för övrigt visat att hon intresserar sig för hur AI-tekniken kan hjälpa oss att etablera ett alternativt samhälle. Hon har bl.a. lanserat en variant av Turing testet som hon kallar Durkheim testet (Star,1989:38-41). Det ifrågasätter om tester av intelligens verkligen skall göras i en jämförelse med hur en individ beter sig intelligent. Hon framhåller att en kritik av sådana individuella jämförelser har framförts av dem som sysslar med DAI. Hon framhåller att en test av ett öppet system, som DAI, kräver att vi inte bara testar en nod i systemet. Att testa alla noder i systemet på en och samma gång är ändå inte möjligt. Om vi följer Durkheims uppfattning om karaktären av sociala fakta bör vi testa intelligens i ett DAI-system genom kollektiva fenomen. Det vi bör analysera är hur DAI-systemet förmår vara användbart för arbetet i samhället, för samhällets förmåga att förändras och dess möjligheter för att anpassa sig till nya förutsättningar. Samt hur DAI kan hjälpa till att exponera flera synpunkter och samtidigt öka kommunikationen mellan olika uppfattningar eller mellan olika delar av ett samhälle.

Star visar att en mer nyanserad syn på tekniken hjälper oss att se att det är begränsningarna i det mänskliga medvetandets förmåga att se nya möjligheter som är en viktig orsak till tendenserna att allt fler områden blir standardiserade. När vi människor brister i vår förmåga att hantera alla de intryck som vi översvämmas av i dagens informationstäta värld, försöker vi ordna världen enligt de referensramar vi redan har. Skall vi ta oss ur detta är lösningen knappast att vi avstår från teknik och istället tillber den kommunikation som sker i direkt dialog mellan individerna. Den lösningen berövar oss möjligheterna att förändra systemen. Utan tekniken blir vi fast i den givna systemreferensen.

Vi skall avsluta med att framhålla att mycket av den forskning som bedrivs kring det som kallas HCI, Human Computer Interaction, bygger på sociologisk eller antropologisk teori. I boken *Where the Action Is* framhåller Paul Dourish att forskningen är inspirerad av Erving Goffmans bok *The Presentation of Self in Everyday Life*. Intresset handlar framför allt om hur människor uppträder beroende på hur andra ser på dem i olika situationer (Dourish,2001:91-92). Den andra sociologen som har haft stor betydelse, framför allt för utvecklingen av begreppet embodiment, är Alfred Schütz. Det är framför allt hans diskussion i boken *The Phenomenology of the Social World* om hur intersubjektiviteten fungerar som har väckt intresse (Dourish,2001:110-113). Dourish menar att Schütz är den som pekar framåt mot det stora intresse för Garfinkels etnometodologi som HCI

forskningen idag visar. Dourish själv har utvecklat en form av metodologi, inspirerad av etnometodologin, som han kallar *teknometodologi*.

The word itself is a little flippant, but the model it proposes is not. By a "deeper" relationship, we mean one that satisfies two criteria. The first is that it attempts to draw not simply on a set of observations of a specific working setting, but rather on ethnomethodology's fundamental insights about the organization of action as being a moment-to-moment, naturally occurring, improvisational response to practical problems. The second is that it attempts to relate these understandings not simply to the design of a specific interactive system aimed at a specific setting, but rather, at the basic, fundamental principles around which software systems are developed - ideas such as abstraction, function, substitution, identity, and representation (Dourish,2001: 77-78).

Ovan ser vi något som kan tyda på att språkregler skall tolkas som tillfälliga anvisningar om relationer mellan representationer såsom etnometodologin förespråkar, istället för som kausala regler. Vi kan också se att begreppet *coupling*, som är centralt för Luhmann, också har betydelse inom HCI forskningen.

Coupling is how an intentional reference is made effective. In the physical world, my local actions can have a remote effect, through a chain of couplings, from one thing to another to another - perhaps from my hand to a lever to a rock I want to move. I act on the rock by acting on the lever. As far as I am concerned, I am acting on the rock; from my point of view, the rock and the lever are coupled. This idea of coupling is not simply a physical phenomenon but an intentional one too. My actions are outwardly directed, through a chain of associations (including, as we will see, social and linguistic associations as well as physical ones). By coupling, I mean the way that we can build up and break down relationships between entities, putting them together or taking them apart for the purpose of incorporating them into our action (Dourish,2001:138).

7.3 Biblioteken som ett standardiserat system

Vi såg i inledningskapitlet att det är svårt att ge praktiska exempel på vad det är diskussionen handlar om. För min egen del har intresset för att diskutera AI uppstått via de observationsmöjligheter som praktiska erfarenheter av

biblioteksarbete gett mig. Jag skall därför ge några praktiska exempel från biblioteksinstitutionens verksamhetssätt för att åtminstone ge en antydning om vad det teoretiska perspektivet kan betyda i verkligheten. Diskussionen om standardisering är definitivt något som berör biblioteksinstitutionens praktiska verklighet. Dessutom handlar biblioteksarbete om en förmedling av kunskap där det inte finns någon direkt mänsklig handling mellan kunskapsproducent och konsument. Vi kan påstå att standardisering som en form av regel har behövts därför att förmedling på distans kräver tekniska redskap. Ser vi på hela biblioteksinstitutionen som ett tekniskt redskap är det lärorikt att veta att den mänskliga handling som författandet av en bok innebär, ganska snabbt kommer i skymundan när den kontext skapas som biblioteket representerar. Den förmedling som går via biblioteket är så beroende av många andra handlingar att det är rimligare att se på hela förmedlingsprocessen som en social process. Och de handlingar som utförs i senare led av förmedlingsprocessen har blivit så mycket beroende av regler och tekniska redskap att dessa handlingar knappast kan ses som autonoma. De handlingar som kan göras är istället så beroende av artefakterna att det blir rimligt att istället för social handling tala om en process i ett system.

Vi kan illustrera detta förhållande genom att ge exempel från de handlingskedjor som sker mellan det att en bok köpts in till biblioteket och till att den lånats ut. En del av inköpen handlar om att köpa in sådant som föreslagits av andra. På ett universitetsbibliotek kommer förslagen i de flesta fall från forskare och lärare på olika institutioner vid universitetet. Vem är det då som förmedlar? Den lärare som vill att en viss bok skall finnas på biblioteket så att studenterna kommer åt att läsa den, eller bibliotekarien som faktiskt sett till att den rent fysiskt har förflyttats från förlagets lagerutrymmen till bibliotekets hyllor? Så långt kan vi se att det handlar om mänskliga handlingar. Men man kan också hävda att bibliotekets förmedlande handling börjar först när den katalogiserande bibliotekarien ser till så att boken fått sådan klassificering eller sådana ämnesord att den kan exponeras för den som söker efter kunskap? I sammanhanget kan vi nämna att den handlingen oftast startar på andra håll genom att det är någon på ett nationalbibliotek som har katalogiserat eller klassificerat boken redan då den utgivits.

Själva klassificeringsarbetet visar att det kan gå lång tid mellan handlingar i förmedlingen. Någon eller några har för länge sedan bestämt att ett ämnesinnehåll skall klassificeras på ett visst sätt i bibliotekens kataloger. En bibliotekarie som katalogiserar den nyinköpta boken kan läsa sig till vad den handlar om. Däremot kan inte bibliotekarien klassificera den med de ord bibliotekarien själv anser lämpliga. Denne måste lägga på en klassifikation som är accepterad av det klassificeringssystem som biblioteket håller sig med. De handlingar som från början varit avsiktliga har automatiserats därför att det finns ett behov av att standardisera vissa delar av förmedlingen. Behovet av att frigöra kunskapsförmedlingen från den omedelbara personliga relationen mellan förmedlare och mottagare gör att den

styr av en mänsklig intentionalitet som ligger långt bak i handlingskedjan. I fallet med klassificeringssystemen kan man hävda att en slags samhällelig uppfattning om kunskap styr de kunskapsförmedlande handlingarna. Den uppfattningen är oftast knuten till uppfattningar som gällde för mycket länge sedan. De svenska bibliotekens klassificeringssystem SAB (Sveriges Allmänna Biblioteksförenings system) etablerades i början av förra århundradet. På så sätt är bibliotekarien fast i traditionella sätt att se på kunskap. Vi kan säga att bibliotekarier fastnar i hur det kulturella kapitalet manifesterades för mycket länge sedan. Att SAB systemet har sin grund i förhållanden kring förra seklets början har diskuterats i en svensk biblioteksvetenskaplig avhandling (Hansson, 1999).

Med de termer som vi har använt i denna avhandling handlar det om att biblioteken är exempel på att expertsystem funnits redan innan datoriseringen. Problemet med deras funktionssätt liknar också de som vi har sett att riktas mot moderna datoriserade expertsystem. Att de har svårt att genom sina egna förmedlingssystem visa på kunskapskopplingar utanför de som redan är formaliserade. Idag försöker man utveckla de datoriserade katalogerna bl.a. med något som inom biblioteksvetenskapen kallas ontologier. Det är ett sätt att skapa någon slags regler för hur formaliserade begrepp inom olika vetenskapliga discipliner skall översättas till varandra. Vi skulle då få redskap som försöker överbrygga svårigheterna att utifrån en vetenskaplig disciplins formaliserade begreppsvärld kunna nå kunskap inom andra ämnesområden som är formaliserade med helt andra begrepp. Alltså att nå något som man utifrån ett betraktelsesätt upplever som icke-formaliserat.

Min egen erfarenhet av att försöka använda söksystem som är uppbyggda på detta sätt är att det man hittar för det mesta framstår som icke-relevant. Mitt avhandlingsarbete startade en gång i tiden som ett försök att via empirisk analys av bibliografiska poster hitta gemensamma beskrivningar mellan olika ämnen. Det mest påfallande i det arbetet var att det inte ens inom sociologin fanns många indexerings av litteratur som tydde på att sociologer använder en helt entydig begreppsapparat.

Det har emellertid alltid i funnits ett exempel på att biblioteksinstitutionen också har haft en förmåga att exponera icke-formaliserade relationer. Inte genom rationella söksystem eller gemensamma begrepp utan genom att böckernas fysiska placering i bibliotekens rum skapar informella strukturer. Det handlar om den *serendipity* funktion som vi redan sett att Merton har sociologiserat. Alltså det fenomen som bland annat uppstår då vi botaniserar bland hyllorna på ett bibliotek och upptäcker något som vi inte visste att vi letade efter.

Vi bör uppmärksamma att denna funktion faktiskt har att göra med att en struktur gör att regler inte alltid determinerar vad som är möjligt. Den handlar till en del om att de regler som styr klassificeringen har sådana latenta funktioner att de i kombination med att boken i rent fysiskt hänseende måste placeras på en plats, skapar nya möjligheter att exponera innehållet i en bok. Till stor del handlar också

denna exponering om en kombination med andra regler. Till exempel en regel om att böcker placeras på hyllan efter första författarens namn eller på titeln om det inte finns någon uttalad författare eller mer än tre författare. Eller att boken är så hög att den inte ryms på en vanlig hylla utan hamnar på en hylla för folianter. Den kan också vara så aktuell att den hamnar på en hylla för nyinköpt litteratur. Detta kan se ut som tillfälligheter men som Merton visar handlar serendipity inte om slump utan om att det ändå finns vissa strukturella egenskaper som skapar denna funktion. I vårt sammanhang påstår jag att det handlar om att biblioteken också styrs av regler som är så lösa eller tillåtande att de kan skapa relationer som är av mer informell karaktär.

Det finns anledning hävda att denna serendipity funktion blir effektiviserad av DAI-system. Det beror inte så mycket på att det finns möjlighet att i datasystemen göra så att regeln i sig blir lösare. Det handlar istället om att de yttre faktorer som har gjort att serendipity funktionen uppstått, genom AI-systemens förmåga att nå enorma mängder av texter, skapar förutsättningar för att regeln som styr funktionerna kan förändras. Det handlar inte bara om att texter styr kodens möjlighet till förändring. I de söksystem på Internet som vi idag har, påverkas koderna också av människors användning av texterna. Google är ett bra exempel på ett sådant söksystem som arbetar med en blandning av AI-programmering och människornas faktiska länkvanor. Dessa länkvanor är väldigt anonymiserade och vi vet egentligen inte vad som är viktigast för dessa handlingar. Att människorna haft tillgång till ett väl fungerande söksystem eller att de själva står i förhållande till kunskapen på något annat sätt? I kombination skapar detta regler som uppenbarligen förmår expandera kunskapen på ett sätt som upplevs som mer tilltalande än resultatet av användningen av vanliga begreppsorienterade söksystem. Det finns anledning hävda att den funktionaliteten uppkommer just därför att datatekniken förmår effektivera de möjligheter som har att göra med att kunskapen är distribuerad. På så sätt skapas en artefakt som gör det möjligt att överskrida de gränser, eller regler, som människor tillsammans skapar. Genom AI kan regeln bli mycket mer föränderlig vilket gör att AI kan revolutionera den funktionalitet artefakterna har i kommunikationen. Bibliotekens tekniska förmedlingsprocess har byggt på att funktionalitet som uppkommit genom människors handlingar har blivit anonymiserad. Men vi har ändå haft kvar begränsningar som har att göra med att mänsklig funktionalitet har svårt att frigöra sig från fasta regler. Det finns anledning att tro att förmedlingsprocesserna blir mer förutsättningslösa genom att använda AI-teknik. Vi kan dock tillägga att AI teknik sannolikt också fungerar bättre i kombination med ett visst regelverk. Talar vi om elektroniska texter är det en fördel om de ordnas enligt regler som gör att de lättare kan exponeras. Kommersiella förlag använder idag troligen en XML-standard för att publicera och

lagra dokument¹¹. Det är ett exempel på att exponerade funktioner som *serendipity* effektiveras om de grundläggande strukturerna är rationellt ordnade.

7.4 Tre perspektiv på AI

Den diskussion som vi hittills har fört kan sammanfattas på följande sätt: Om den mänskliga hjärnan i kunskapsprocesserna utför komplexa funktioner som har att göra med interna mentala kapaciteter, finns det knappast möjlighet att skapa AI-system som skall fungera som artificiella hjärnor. Avhandlingen har tagit en positiv inställning till AI. Den inställningen grundas på att kunskapsutveckling handlar om social kommunikation. Och social kommunikation bygger inte på individuella handlingar. Denna sociala kommunikation har förstås kopplingar till individuella handlingar men det är inte dessa handlingar som konstituerar den sociala kommunikationen.

Om AI-system fungerar innebär det att de processer som skapar kunskap blir mindre beroende av att människorna deltar i dem. Det innebär inte att människorna inte har del i dessa processer. Ty någonstans kommer människan att vara med i processerna, eftersom det blir meningslöst med AI-system om de inte någonstans skapar en ny observatörsposition för en människa. Det har mest att göra med att vi som värderar AI-systemen är människor. Det blir helt enkelt meningslöst att tala om kunskap eller intelligens om inte människan kan uppfatta dess kvalitet. Det är emellertid en skillnad mellan detta förhållande och påståendet att denna kvalitet bara kan skapas genom denna mänskliga aktivitet. Kunskap bör kunna skapas utan någon inblandning av människan. Om så inte är fallet är kunskap bara en mänsklig konstruktion. Vi bör uppmärksamma att det också är ett mycket kontroversiellt påstående som starkt ifrågasatts åtminstone av naturvetenskapen.

Ovanstående öppnar för tre uppfattningar om vad problematiken handlar om:

1. Människan måste alltid ha tillgång till redskap för att bete sig intelligent men redskapen kan aldrig skapa något på egen hand. Det är alltid människan som måste göra så att redskapen fungerar. Även om AI erbjuder ett ytterst potentiellt redskap kan detta redskap aldrig skapa kunskap på egen hand. Det kan bara utnyttjas genom människan. Vi kan tala om svag AI men aldrig om stark AI.
2. En förutsättning för att människan skall kunna komma fram till ny kunskap är att redskapen kan skapa den nya kunskapsstruktur som tillåter

¹¹ Det finns i XML standarden en s.k. DTD (Document Type Description) för både böcker och tidskrifter/tidskriftsartiklar.

människorna att ta sig ur invanda kunskapsmönster. Det förutsätter att redskapen kan ompositionera möjligheter till observationer. Ur detta perspektiv innebär tanken på stark AI inget nytt. Kunskapen har alltid omstrukturerats av kulturella redskap som böcker och bibliotek. Den funktionalitet som behövs för intelligens kan därför finnas både hos redskap och hos människor. Det är ett förhållande som gör att vi bör se på kunskap som ett distribuerat fenomen. Att vi idag talar om stark AI som en ny företeelse handlar om att datatekniken gör att redskapen blivit så kapabla på att effektivisera dessa processer att vi upplever att de kan göra något som vi tidigare trott att bara människor klarar av.

3. Intelligens handlar om ett samarbete mellan människor och redskap där den ena eller den andra kan fungera självständigt men där samarbete alltid leder till bättre resultat. Det är en strategi som samtidigt innebär att en människa i förhållandet till en annan människa fungerar som ett redskap för den andra människan och vice versa. Sociala interaktioner mellan människor kan vara effektiva för att utveckla kunskap utan att de i just denna interaktion behöver använda kulturella redskap. Att människorna fungerar som redskap för varandra antyder att kommunikation inte alltid behöver bygga på någon själslig kapacitet som är inneboende i den mänskliga naturen. Analogt med detta perspektiv är att redskapen också kan kommunicera med varandra utan att de behöver gå genom människorna. Att vi talar om stark AI handlar om att redskapen genom datatekniken får bättre förutsättningar att innehålla den funktionalitet som möjliggör att de kan interagera med varandra. I det antagandet ligger att också den mänskliga interaktion som inte går genom artefakter är mycket enkel. Kontingensproblematiken gör att varken den interaktion som sker mellan redskapen eller den interaktion som sker mellan människor bygger på en särskilt komplicerad funktionalitet. Huvudsaken är att den kan leda till någon slags omstrukturering av kunskapen. Den komplicerade funktionalitet som behövs för att skapa ny kunskap uppkommer då människorna och redskapen interagerar med varandra och med den miljö de ingår i. På så sätt uppstår komplicerade processer genom att kunskap är ett distribuerat fenomen. Intelligens är då en egenskap som får sin komplexa funktionalitet genom de kombinationer som uppstår därför att kunskapen är distribuerad.

Det är bara den första uppfattningen som inte är förenlig med tanken på att AI kan vara stark. På den sista uppfattningen bygger den AI-strategi som vi har diskuterat i denna avhandling, nämligen den som kallas Distribuerad Artificiell Intelligens. Den grundas på att kunskapsskapande, när det är som bäst, handlar om ett samarbete mellan människor, redskap och de förutsättningar som finns i deras miljö. Det leder till en process där inte någon part har en överordnad roll. Det är en uppfattning som är väl förenlig med Luhmanns uppfattning om de förutsättningar som behöver finnas för att reflexion skall skapas.

7.5 Människors kunskapsförmedling i praktiken

Om vi överger tanken på att emotioner och medvetna handlingar ligger bakom alla de faktorer som skapar ny kunskap bör kunskapsförmedling handla om att exponera relationer så att vi kan se hur ny kunskap kan kombineras med gammal kunskap. Antagligen är det framför allt kunskapsproducenter som exponerar sådana relationer. I akademiska sammanhang handlar det främst om forskare och lärare. Det blir då intressant att jämföra denna typ av förmedling med den som intermediära förmedlare utför. Ett typexempel på en sådan intermediär förmedlare är bibliotekarier. Vi har redan talat om hur biblioteksinstitutionen förmedlar kunskap som en form av artefakt. Det kan vara lärorikt att fundera på vad den utveckling av biblioteksinstitutionens funktioner som vi får med AI-teknik innebär för mänskliga förmedlares positioner i förmedlingsprocessen.

Skillnaden mellan en bibliotekarie och en forskare handlar om att forskaren har ämneskunskap medan bibliotekarien bara kan förmedla genom att använda de redskap som biblioteken och det bibliografiska systemet erbjuder. Bibliotekarier har till idag sannolikt rätt sällan haft möjligheter att visa på nya möjligheter att kombinera kunskap. Det beror antagligen på att de inte har den ämneskompetens som krävs för att kunna gå utöver den etablerade kunskap inom ett ämne som redan beskrivits med de klassificeringar som biblioteken använder sig av. Den ämneskunnige forskaren bör ha varit bättre på att se nya mönster, vilket sannolikt beror på att ämnesexperten under en längre tid har lyckats bekanta sig med innehållet i ett relativt stort antal texter. Medan bibliotekarien med de kataloger och citeringsindex som erbjuds, bara kan se hur texterna inom ett ämnesområde brukar användas. Forskaren har genom sin erfarenhet skaffat sig en översikt av var den distribuerade kunskapen finns medan bibliotekarien måste lita till de symboler som beskriver etablerad kunskap. Bibliotekarien fungerar då enligt samma modell som den symboliska AIn. Det kan innebära att det bara är forskaren som har förutsättningar att i egentlig mening kunna skapa ny kunskap. Skall vi ifrågasätta detta perspektiv bör vi rikta in oss på det vi har sagt om medvetandets begränsningar. Det skulle innebära att forskaren är alltför fast i en uppfattning om hur kunskapen inom ämnesområdet ser ut för att kunna se nya möjligheter. Forskaren är alltför bunden av sin egen disciplins systemreferens. Det är något som understryks av att det tycks vara svårt att hitta ett disciplinöverskridande språk. Förhållanden som kan innebära att bibliotekarien med sin förmåga att bättre behärska hur redskapen skall användas, och med ett mer neutralt förhållningssätt till kunskapen, är den som har bättre förutsättningar att exponera möjliga kunskapsrelationer.

Vi har dock varit inne på att expertsystem ändå bara kan avspegla formaliserade kunskapsrelationer inom ett ämne. Även om bibliotekarierna kan använda dessa expertsystem för att nå kunskap inom andra ämnesområden som

forskaren inte sysslar med, så handlar det ändå inte om en kapacitet att nå det icke-formaliserade. Det finns också anledning att tro att den kunskap som handlar om att kunna hantera dessa expertsystem är så enkel att det knappast finns anledning att tro att inte en forskare med ganska liten ansträngning skulle kunna lära sig att hantera dessa expertsystem. Det stärks av att Wertsch menar att det sällan handlar om ett behov av att kunna internalisera kunskapen om redskapen innan vi kan använda oss av dem. Eftersom själva söksystemen fungerar likadant inom alla ämnesområden gäller det kanske mest för forskaren att skaffa sig anledning att vilja observera något utanför det egna kunskapsområdet. På så sätt bör forskaren också lätt komma åt det som är formaliserat inom andra discipliner även om han/hon inte kan observera det som inte är formaliserat. Men det finns också anledning hävda att erfarenheten av både det som är formaliserat och inte formaliserat inom det egna området skapar en observationsmöjlighet som pekar på de informella relationer som kan finnas mellan de bägge disciplinerna. Vi bör emellertid inse att det handlar om att forskaren här exponeras för kunskaper som är distribuerade på två olika håll. Det är tveksamt om det räcker till att den mänskliga kapaciteten ser det informella inom det egna området. Den kapaciteten bör snarare tendera till att man fastnar i den egna disciplinens systemreferens. Man måste bli exponerad för motsvarande informella möjligheter inom andra discipliner innan erfarenheter av informella relationer kan appliceras på en annan disciplin. Det är först när man exponerats för möjligheterna inom en annan disciplin genom att använda tekniska redskap, som denna förmåga att hantera det informella inom den egna disciplinen kan användas för att se nya informella relationer i det nya sammanhanget.

Instinktivt uppfattar vi säkert att AI är ett större hot mot bibliotekarier om dessa redskap kan skapa kunskap på egen hand. Om det är så att redskapen bara fungerar genom de människor som själva kan skapa kunskap finns det dock anledning att hävda att den svaga AI:n, som ett potent redskap för kunskapsproducenter, gör att bibliotekariernas ställning hotas. Ty om redskapen bara kan användas av en människa så har den som själv kan skapa kunskap alltid en fördel. Så beskrev jag ovan forskarens relation till expertsystemen. I enlighet med detta resonemang finns det fog för att hävda att de intermediära förmedlarnas framtid måste bygga på att de redskap de använder har funktionell kapacitet som motsvarar den som en ämneskunnig forskare har. Om redskapen kan skapa kunskap på egen hand finns det fog för att anta att ämnesexpertens försprång minskar.

Vi måste troligen inse att stark AI kan innebära att den som söker efter kunskap har lättare att klara sig helt på egen hand. Alltså att en biblioteksanvändare, eller att vem som helst som är sugen på att exponeras för nya kunskapsmöjligheter, genom AI-redskapet blir sin egen bibliotekarie. Om AI fungerar så bra att människorna direkt kan interagera med AI-systemen behövs sannolikt ingen intermediär mänsklig förmedlare. Även den som själv är bibliotekarie måste bejaka en sådan utveckling, eftersom den bör ge de bästa

förutsättningar för individerna att bli exponerade för nya möjligheter utan att behöva gå via personer som kan lägga sina egna intressen emellan.

Det är svårt att ge exempel på hur AI-system fungerar. Det är en fråga som mest handlar om att vi inte riktigt vet vad det är vi skall exemplifiera, eftersom det är så svårt att enkelt slå fast vad som över huvud taget ingår i intelligenta företeelser. Det vi bör observera är att det faktiskt är lika svårt att ge exempel på intelligens i sammanhang där inga AI-system alls är inblandade. Vi vet att människor åstadkommer saker som har krävt någon form av intelligens. Vi kan förklara det med tankeförmåga, stimulerande sällskap eller tillgång till miljöer som stimulerar kreativiteten. Men vi kan inte riktigt peka på vad det är som är själva kvaliteten i denna intelligens. Och när vi försöker operationalisera intelligens handlar det om sådant som IQ-tester som handlar om en formaliserad intelligens. Debatten kring AI handlar sällan om att den typen av intelligens **inte** skulle kunna hanteras av kvalificerade AI-program. Den intelligens, i.e. social intelligens, som man menar att AI inte kan omfatta är däremot väldigt svår att exemplifiera. Men mycket tyder på att det är just den typen av intelligens som är den starka DAIns bästa gren.

Vi har i detta kapitel sett att det finns anledning hävda att de kunskapsprocesser som är kopplade till kommunikation inte alls står i konflikt med tanken på stark AI. Vi har kunnat se att kommunikation handlar om processer i de sociala systemen och de behöver använda artefakter för att kunna fungera. Vi har också sett att den standardisering och mekanisering som gör att kunskapen inte kan utvecklas snarast är något som hör ihop med människornas tillkortakommanden. Vi har också sett att de problem som finns kring stark AI måste jämföras med alla de problem som finns kring människornas sätt att skapa ny kunskap. I det perspektivet finns det anledning att påstå att stark AI kan fungera på egen hand men att det samtidigt finns anledning att påpeka att det inte hindrar att också människorna tar del i dessa processer. Vi har påpekat att ett samarbete mellan människor och AI system sannolikt leder till det bästa resultatet.

8. Det sociala som holism – Systemteorin och det Starka Programmet

Även om det finns viss möjlighet att se att redskapen kan ha en autonom kraft även i en MI-modell handlar stark AI om att det skall finnas en funktionalitet hos en faktor som ligger utanför människan. Därför är det rimligt att en argumentation för att sociala processer är förenliga med AI bör söka stöd i holistisk sociologi. Det Starka Programmet är ett exempel på en sådan sociologi. Vi skall här använda debatten kring det programmet för att lyfta fram några frågeställningar som är kritiska för diskussionen kring stark AI. Debatten kring det Starka Programmet har visat på några svagheter kring sociologins eget resonemang kring sociala processer. Genom att se på den kritiken och försöka diskutera hur detta program kan göras bättre om det omformuleras enligt den syn på system/omvärld relationen som Luhmann har framlagt, får vi en möjlighet att göra en syntes mellan systemteorin och det Starka Programmet.

8.1 Det Starka Programmet

Det Starka Programmet menar uttryckligen att det är sociala processer som skapar kunskapen. Därför ställer SP sociologins sociala argument på sin spets. På ett sätt som framstår som riktigt men som är svårt att försvara med analytiska argument. Det är mina egna försök att hitta ett försvar för detta sociologiska program som också ledde mig fram till att inse att det inte är omöjligt att stark AI kan fungera. I sammanhanget bör vi nämna att behandlingen av begreppet social i SP, förefaller bygga på en annan uppfattning om dess innebörd än i den MI-orienterade sociologi som är kritisk till AI, eftersom begreppet här inte är direkt kopplat till interaktioner mellan individer. Det blir något oklart vad det står för och försöken

att trots allt placera in resonerandets betydelse i kunskapsprocesserna har inte varit övertygande.

Min egen reflexion under arbetet med denna avhandling har varit att sociologins negativa hållning till AI kan förklaras med att sociologin inte inser den komplexitet som finns inom sociologin själv. Jag uppfattar att den negativa hållningen till Luhmanns systemteori från MI-orienterade sociologer, har att göra med att man inte inser att Luhmann grundligt diskuterar de sociologiska grundproblem som de själva inte ser. Debatten kring det Starka Programmet visar på att vi behöver hitta en formel för att analysera förhållandet mellan kommunikation och sociala faktorer, om vi skall formulera ett rimligt argument för hur sociala faktorer påverkar kunskapsprocesserna.

David Bloor, som är det Starka Programmens grundare, menar att vetenskapssociologin måste uppfylla fyra villkor för att nå upp till de krav som är givna inom andra vetenskapliga discipliner (Bloor,1976:4-5). Tomas Brante, som är den svenske sociolog som skrivit om Bloors teser, översätter dessa så här till svenska (Brante,1984a:35). Dess förklaringar måste vara:

1.Kausala, dvs. gälla de villkor som åstadkommer trosföreställningar och kunskaper;

2.Opartiska och neutrala, dvs. inte ta ställning till föreställningars eventuella sanning eller falskhet, rationalitet eller irrationalitet, framgång eller misslyckande;

3.Symmetriska, dvs. samma typer av orsaker skall förklara både "sanna" och "falska" kunskaper;

4.Reflexiva, dvs. i princip skall förklaringsmönstret vara tillämbart även på sociologin själv. (Ett nödvändigt villkor, annars skulle sociologin motsäga sina egna teorier). Det framkommer inte explicit i detta stycke att Bloor argumenterar för att vi måste leta efter sociala faktorer för att förklara de kausala processerna bakom skapande av kunskap. Men som vi skall se är det just kring de sociala faktorerna som debatten kring Bloors program cirklar.

Bloors egen kritik av den traditionella vetenskapssynen ser ut på följande sätt

The general structure of these explanations stands out clearly. They all divide behaviour or belief into two types: right or wrong, true or false, rational or irrational. Then they invoke causes to explain the negative side of the division. Causes explain error, limitation and deviation. The positive side of the evaluative divide is quite different. Here logic, rationality and truth appear to be their own explanation. Here causes do not need to be invoked (Bloor,1976:6).

Jag vill redovisa att mitt intresse för att koppla ihop det Starka Programmet med frågan om stark AI väcktes när jag läste en bok av András Kertész som just uppmärksammade likheten mellan stark AI och detta program. Kertész jämför

något som närmast framstår som stark symbolisk AI med SP. Han menar att de har en gemensam ontologisk utgångspunkt i sin syn på att kunskap är något verkligt och oberoende av människor. Han försöker hävda att det finns någon form av modulärt förhållande mellan kognitiva och sociala faktorer. Det baserar han på att bägge dessa program ser att dessa faktorer är givna på förhand innan medvetandena kan börja processa dem. Kertész menar att denna modulära hypotes kräver att sociologin måste reduceras till kognitionsvetenskap (Kertész, 1993:162).

Jag har ändå haft svårt att acceptera en sådan hypotes just därför att sociala faktorer ändå måste vara något som knyts till den mänskliga kulturen och inte till några kognitiva egenskaper i en naturvetenskaplig världsuppfattning. Kertészs diskussion har ändå den fördelen att den visar på svårigheterna med att se på sociala faktorer som något som uttrycks genom kausala processer. Ty i ett sådant perspektiv har jag haft svårt att förklara det sociala som något annat än ett fenomen som har att göra med biologisk utveckling. Det skulle innebära att det finns fog för Kertészs påstående att det sociala måste reduceras till något kognitivt. Det har efterhand framstått för mig som att både SP och symbolisk AI har fel just därför att de utgår från kausalitet. Skall vi kunna försvara sociala faktorerers självständiga roll måste vi försöka argumentera för att kunskap är relativa storheter som dessutom kan fungera fristående från mänskliga handlingar. På så sätt kom jag att intressera mig för situationernas betydelse, vilket senare fick mig att inse att det är stor skillnad på konnektionistisk AI och symbolisk AI. Innan jag förstod det var jag nog inne på att använda SP till att ifrågasätta möjligheterna av stark AI. Istället har jag kommit att se DAI som en möjlighet att försvara de sociala faktorernas roll i kunskapsprocessen

8.2 Debatten kring Bloor's Starka Program

Kritiken mot Bloor kommer från olika håll. Vi skall börja med den som kommer från dem som förespråkar den klassiska AI:s möjligheter att skapa kunskap. Den kritiken ligger delvis vid sidan av diskussionen om konnektionistisk AI. En del av förespråkarna för klassisk AI använder dock argument som bara håller om funktionaliteten i AI-systemen inte behöver vara komplicerad. Därför kan den diskussionen hjälpa oss att diskutera relationen mellan AI och sociala faktorer.

När vi har gått igenom de argumenten skall vi fortsätta med en kritik som vetenskapsteoretiker riktar mot Bloor. Även om den kritiken liknar den kritik som kommer från klassisk AI, har den en lite annorlunda vinkling. Den betonar mer att vanligt mänskligt tänkande inte passar in i de mönster som det Starka Programmet använder. Det är alltså en kritik som så starkt betonar den mänskliga funktionaliteten att den vare sig kan ersättas med sociala lagar eller med artificiell

intelligens. Vi skall använda denna kritik för att visa på att ett program som betonar sociala faktorer roll i kunskapsprocessen åtminstone behöver en teori om det artificiellas funktionalitet för att bli trovärdig.

8.2.1 Debatten mellan datakognitivisterna och David Bloor

Peter Slezak angriper SP med ett inlägg som samtidigt är ett försvar för en stark symbolisk AI. Det gör han genom att framlägga ett exempel på en vetenskaplig upptäckt som är totalt isolerad från alla sociala eller kulturella faktorer. Han formulerar sin kritik genom att göra anspråk på att de artificiellt intelligenta programmen arbetar med en ren eller av sociala faktorer icke-kontaminerad form av induktiv inferens. Han hävdar också att AI-programmen är kapabla till att självständigt leta fram klassiska vetenskapliga lagar från råa observationer. Hans exempel är att programmen inte hade något problem att hitta den lag om gstryck och volym som Robert Boyle framlade runt år 1660. Förutsättningen var att de fick de råa eller ostrukturerade (noisy) data som Boyle använde. Detta menar Slezak är ett starkt argument emot SP (Slezak,1989:563-564). Vi bör uppmärksamma att detta innebär att lagarna, d.v.s. det som vi skall deducera ifrån, upptäcks genom induktion. Det är alltså genom induktion som världen exponerar sig för både människor och datorer. Vi bör observera att denna form av vetenskapsteori menar att kunskap inte bara är något som människor skapar. Om människan skapar kunskap så handlar det om den process där hon upptäcker den kunskap som naturlagar ger. Eftersom det här handlar om en process som styrs av naturlagar är symbolisk AI inte beroende av någon social kommunikation.

En svaghet i det resonemang som dessa företrädare för stark symbolisk AI för, är att de inte specifikt diskuterar vad det är som är viktigast i deras strategier. Är det att det finns lagbundenheter i världen eller handlar det om kapacitet att upptäcka denna lagbundenhet genom att AI-programmen har deduktiv kapacitet eller kapacitet att upptäcka lagbundenheterna med induktiva metoder? Troligen skall vi tolka dem som att alla tre strategierna behövs. Stark symbolisk AI bygger då på att upptäcka naturlagar och att efterlikna alla egenskaper som man tillskriver mänsklig intelligens. Det gör att programmet, som modell, blir rätt svårt att kritisera om man utgår från ett MI-perspektiv eftersom modellen antar att mänskliga handlingar kan konstrueras på ett artificiellt sätt. Searle kritiserar t.ex. inte modellen utan han kritiserar att en artificiell hjärna kan ersätta den mänskliga hjärnan i modellen. Om man vill kritisera symbolisk AI utifrån sociologiska aspekter blir det mer tacksamt att verkligen ifrågasätta om intelligens handlar om egenskaper som vi antar finns i den mänskliga hjärnan och att världen är lagbunden. De konnektionistiska programmen bygger delvis på den kritik som riktats mot att dessa symboliska AI-program innehåller för mycket. Men de skiljer sig också från de symboliska modellerna genom att de inte bygger på att

kunskapsutveckling har att göra med lagbundenheter i världen. Världen är i istället en storhet som hela tiden presenterar sig för människan genom att exponera olika relationer. Konnektionismen ser deduktiva modeller i programmen som tillfälliga uppfattningar om relationer i världen. Det som den har gemensamt med den symboliska AI:n är den induktiva strategin. Vi har tidigare sett att konnektionistiska modeller kritiserats för att ligga för nära rena induktiva strategier. Vi har också visat att Engeström menar att induktiva strategier inte kan expandera kunskap. Engeström för ett resonemang om att induktiva strategier behöver kombineras med deduktiva. Men hans diskussion implicerar samtidigt att det medvetande som står ensamt, är så begränsat att det inte kan göra deduktioner som expanderar kunskapen om det inte samtidigt finns en materialitet inblandad. Det har vi tidigare uppfattat som att det talar för konnektionistiska system. Utifrån ett sociologiskt betraktelsesätt framstår problemet med den symboliska AI:n som att den försöker argumentera för att omvärlden kan påverka de sociala systemen bara genom tänkande och analyserande inom en processande enhet.

SP liknar konnektionismen på så sätt att den tar avstånd från att det är deduktiva processer som skapar våra möjligheter att se världen. Istället menar SP att de teorier, genom vilka vi ser världen, skapas genom sociala processer. På det sättet nedtonar programmet tänkandet. Detta leder till att SP har svårt att visa på hur egenskaper i världen ändå kan påverka den sociala uppfattningen om kunskap, eftersom programmet inte ersätter tänkandet med materiell eller artificiell funktionalitet.

SP:s svårigheter kommer fram i den kritik som Slezak riktar mot SP för att vara ett behavioristiskt program. Slezak menar att kunskapssociologer nyligen har återupptäckt Wittgenstein och Skinner och därmed gett ett bifall till en behaviorism som redan funnits uttalad i deras verksamhet. Det skapar, med Slezaks ordval, en metodologisk konkurs och en teoretisk sterilitet ((Slezak, 1989:590).

Det finns en poäng med att kritisera Bloor för behaviorism. Om Bloor inte kan visa att sociala processer fungerar fristående från individuella handlingar kommer vi tillbaka till att de sociala egenskaperna hör samman med människans konstitution. Samhället har då ingen betydelse i sig självt utan den blir endast en aggregation av mänskliga handlingar. På så sätt kan man t.ex. inte påstå att en utveckling av samhällets strukturer förklarar kunskapsutveckling. Kunskapsutveckling blir istället knutet till enskilda aktörers handlingar. Om det leder till att samhällsstrukturer och kulturella strukturer utvecklas är det ändå inte på dessa strukturers axlar vi kan stå om vi vill fortsätta utveckla kunskapen. Vi måste istället alltid ha kontakt med mänskliga kunskapsproducenter. På så sätt blir det svårt att skilja Bloors position från den sociologiska konstruktivismen. Det finns emellertid anledning att hävda att Bloors kausalitetskrav skapar ett ännu större problem. Det är att kausalitet bör knytas till en faktor som grundas på en

realistisk ontologi. Då är det svårt att hitta grunden till kausaliteten i något annat än det som hör till människans biologiska konstitution. I MI-teorier förklaras människans beteenden med att de utvecklas då människor interagerar med varandra. Vi har redan varit inne på att det finns problem med att se hur en sådan utveckling kan ske och vi kommer senare att diskutera dessa problem närmare. Däremot tycks MI se på sådana egenskaper som relativa – de förändras ständigt. Om man istället ser på det sociala som kausala processer blir man tvungen att grunda det sociala i något som handlar om egenskaper som hör till människans biologiska natur. I så fall måste de ha att göra med en biologisk utvecklingsprocess, vilket skapar ett släktskap mellan SP och den brutala sociobiologin såsom den lanserades av Wilson. Då ligger det också nära att anklaga SP för behaviorism. Sådan kritik behöver inte programmet råka ut för om det omformulerades efter Luhmanns funktionella relativism. Det borde innebära att man måste förklara det sociala med hjälp av den kulturella och artefaktiska strukturen.

Nackdelen med att tala om kausalitet istället för funktionalitet är att det blir svårt att se att kausalitet i sociala processer kan skapas av någon annan faktor än människan. Det är ytterligare en orsak till att SP får svårt att förklara varifrån kausaliteten kommer om det inte handlar om egenskaper som lagts ner i människorna i en biologisk utvecklingsprocess.

Det går inte att sätta likhetstecken mellan behaviorism och induktion eftersom behaviorism handlar om att följa andra människor och induktion handlar om att världen påverkar människorna. Engeströms kritik av induktion, och behaviorismkritiken mot Bloor, bygger dock på att induktionerna respektive behaviorismen innebär ett passivt mottagande av intryck från omvärlden. Det är ett förhållande som liknar kritiken mot konnektionismen eftersom den kritiken anklagar dessa system för att inte kunna göra något annat än att passivt ta emot intrycken. Likheten mellan kritiken mot SP och mot konnektionismen, har då att göra med att bägge teorierna anklagas för att inte kunna redogöra för hur de kunskapsprocesser som teorierna anger, kan skapa nya förutsättningar. Utifrån Luhmanns synsätt kan vi säga att problemet är att de inte redogör för hur de processer ser ut som skapar förutsättningar för den reflexion som skapar en relation mellan system och omvärld.

Det som Slezak ser är SP utan en teori om förhållandet mellan de situationer som skapas av artefakter och de sociala processerna. Om de sociala processerna skapar en kontext som individerna kan tolka beroende på de situationer de själva befinner sig i, blir människan påverkad av det sociala. Men hon blir inte determinerad av det sociala eftersom kontexterna är mångtydiga. Det för med sig att det sociala är något som är otydligt och som innehåller artefaktiska inslag. Det är ett försvar för SP som kräver att programmet gör avkall på kausaliteten ty den är inte reflexiv.

Vi skall uppmärksamma att den behaviorism som SP eventuellt representerar är ett resultat av att människor av kött och blod inte är neutrala. System som inte

har tillgång till en funktion som medger en reflexiv förmedling mellan system och omvärld, blir på så sätt bundna (eller koloniserade) av de psykiska systemen ty det är bara genom omvärlden som de kan utvecklas. Det är människorna som lägger in intressen i sina kunskapsskapanden. Då skapas en socialitet som är styrd av intressen. Skall vi undvika behaviorism skall vi knappast ta avstånd från det sociala och ersätta det med mänsklig tankeförmåga. Det är rimligare att vi söker efter en socialitet som inte är styrd av mänskliga intressen. Om de konnektionistiska systemen fungerar är de knappast benägna att lägga sina egna intressen i blöt i skapandet av kunskap. De artefakter som vi länge har haft, framför allt språket, symbolerna och texterna, har säkert alltid fungerat som de artificiella inslag som de goda sociala processerna har behövt. Vi kan föreslå att sociala processer därför alltid varit beroende av kulturella artefakter. Där bör vi börja vårt sökande efter de faktorer som möjliggör att vi kan skapa en social funktionalitet som inte leder till behaviorism.

Det är en skillnad på sociala processer som utförs utifrån etablerade sociala regler och sådana som görs utifrån ett behov av att hitta nya regler. De första tenderar att bevara systemens funktioner, vilket bör betyda att de försvarar olika intressen i samhället. De processer som leder till nya regler är knutna till den expanderande process som Engeström talade om. Det är lätt att tolka det första som ett beteende som styrs av kausala processer. När vi skapar ny kunskap behöver vi dock leta efter sådant som bryter kausala mönster. Skall vi förstå hur det går till måste vi leta efter funktionella processer.

Slezak nöjer sig inte med kritiken ovan. Han hävdar också att det Starka Programmet betoning av kausala samband mellan sociala faktorer och teori-innehåll, representerar en mer extrem position än Skinners behaviorism. Han menar att Skinners orsakande faktorer endast bestod av en stimulans av sinnesorganen. SP däremot postulerar att sociala sammanhang har samma roll som Skinners stimulans. Slezak menar att detta leder till en social determinism som mer påminner om astrologi. En mildare position är att hävda att sociala intressen är något som förmedlar mellan sammanhang och innehåll. Slezak hävdar ändå att den mildare formen underminerar den ursprungliga radikalismen i SP ((Slezak,1989:594-595).

Den mildare form som Slezak talar om liknar den svagare variant av SP som en omformulering av SP till ett funktionalistiskt program skulle göra den till. Antagligen menar Slezak att också denna position är teoretiskt steril. Det antagandet hör samman med att Slezak hävdar en realistisk ontologi vilket gör att han inte kan tänka sig att kulturella kontextfaktorer kan påverka kunskapen. En uppfattning som visar sig i att han ser skapandet av kunskap som något som grundar sig på naturlagar. Om det inte finns fog för den realismen erbjuder den svagare formen av SP ett sätt att se på det sociala som uttryck för olika positioner som förmedlas genom artefakter. Om inte kunskap finns som verkliga storheter i ontologisk mening kan knappast kunskap förmedlas på ett mer pregnat sätt.

Därför bör den svagare varianten av SP vara den möjlighet vi har kvar om vi inte vill hamna i en steril position. Det är bara denna svagare variant av SP som är förenlig med den kritik av informationsbegreppet som den andra ordningens cybernetik har framlagt.

Den viktiga frågan om det socialas beroende av det kognitiva tangerar vi i det försök Ronald Giere gör för att skapa en syntes mellan Bloor och Slezak. Giere bygger sin argumentation på att det inte behöver finnas någon motsättning mellan kognitiva och sociala angreppssätt på vetenskapsstudier. Han menar istället att båda är nödvändiga för att vi skall nå en riktig förståelse av vetenskapen som en mänsklig aktivitet (Gieryn, 1989:638-639).

Gieryn hävdar att det inte finns någon teoretisk bas för att koppla samman behaviorism med SP. Han menar dock att det finns en stor motsättning mellan kognitiva angreppssätt och den så kallade sociologiska chauvinismen, den chauvinism som menar att den psykologi som studerar vetenskapsmännen är irrelevant för vetenskapens praktik och utveckling. Och samma chauvinism som menar att det som sker i vetenskapsmännens huvud är en svart låda, som studiet av vetenskapen som en mänsklig aktivitet kan låta vara stängd. Alltså menar Gieryn att vi ändå måste återvända till MI för att förklara hur de sociala processerna fungerar. Här ser vi att sociologin har svårt att försvara sina positioner, om den inte kan visa på att den funktionalitet som är inblandad i sociala processer handlar om en mycket enklare funktionalitet än den som är inblandad i mänskligt tänkande. Det innebär att Luhmanns teori är bra om vi vill förklara hur holistiska processer fungerar utan att vi behöver hemfalla till en sociologisk chauvinism. Ty genom att Luhmann ersätter tänkande med reflexivitet slipper vi ifrån behovet av att alltid kunna knyta processerna tillbaks till en människas kapacitet. Dessutom innebär tanken på stark AI att vi också kan få en modell av hur dessa processer kan se ut.

Luhmann ger oss en ny möjlighet att se hur de sociala faktorerna påverkar kunskapsprocesserna. Betoningen av observatörspositionen innebär att funktionaliteten inte är en del av en behavioristisk förklaring. Både Luhmann och Bloor ser att en social nivå påverkar kunskapsprocesserna. Men Luhmanns teori om självreferens förklarar bättre att de individuella aktörerna inte blir determinerade av den sociala nivån. Istället måste den sociala nivån påverka individerna genom det som Luhmann kallar interpenetration, vilket är den process där den sociala nivån genom strukturell koppling så småningom påverkar de psykiska systemen. Detta i sig är ett sätt att visa hur medvetandeprocesser är kopior av sociala processer.

Gieryn håller med Slezak om att denna låda måste öppnas och lämnas öppen. Han menar dock att Slezak går för långt och hamnar i samma positioner som de rationalister som hävdar att den kognitiva meningen i observationerna kan nås genom att formulera hypoteser och göra observationer. Det är en syn som betonar det kunnande subjektet. Gieryn framhåller att den synen för länge sedan övergetts av vetenskapsfilosoferna. Dessutom påpekar han att det är ironiskt att en sådan syn skall uppstå bland dem som arbetar med att utveckla AI (Gieryn, 1989:642).

Giere menar att analysobjektet istället måste vara hela vetenskapsmannen. Med det menar han att vi inte bara kan se på det som finns i vetenskapsmannens hjärna. Han menar att vi också måste ta i beaktande att han/hon har många olika intressen och sociala relationer. Dessutom kan det inte handla om en människa som tänker rationellt i alla lägen. Först när vi har insett det kan vi börja med att empiriskt beskriva de sociala och kognitiva faktorer som är inblandade i särskilda vetenskapliga perioder. Här tycks Giere ändå lämna det individuella perspektivet och börja argumentera för något som mer liknar en artificiell social positionering. När han hävdar att människan inte tänker rationellt i alla lägen öppnar han för att något annat än avsiktliga handlingar av människor också är inblandade i skapandet av kunskap. Att detta annorlunda handlar om att både sociala och kognitiva faktorer är inblandade i vissa perioder öppnar för att det sociala kan vara knutet till en viss situation. Det talar för att sociala faktorer måste påverka människan genom kontexter. Det blir därför naturligare att se på det sociala som uttryck för en artificiell positionering. På så sätt finns det heller ingen motsättning mellan sociala och kognitiva faktorer. Snarast är det så att det kognitiva och det sociala smälter samman. Det innebär också att artificiella inslag är något av normaltillståndet för att förmedla kunskap. Wertsch diskussion om internalisering gentemot externalisering pekar på att det åtminstone inte bara är vetenskapsmannens inre arbete som skapar den funktionalitet som behövs i kunskapsprocesserna.

En som försöker sammanjämka kontrahenterna är Michael Gorman. Han diskuterar vilka möjligheter det finns att låta konnektionistiska program bygga på en modell som både innehåller kognitiva och sociala faktorer. Gorman framhåller att dessa program kan vara kapabla att skapa kunskap på egen hand. Gorman hävdar att frågan om dessa program verkligen motsäger vetenskapssociologernas positioner avgörs av på vilket sätt de kommer fram till sina upptäckter. Gorman ställer några retoriska frågor som påminner om det exempel Randall Collins uppställde. Kommer programmen att lära sig på samma sätt som vetenskapsmännen gör. Och kommer de att kunna förbättra sina "papper" genom de responser som de får från referenter? Om dataprogrammen kan innehålla dessa funktioner så menar Gorman att de konnektionistiska programmen, istället för att gå emot vetenskapssociologin, kan hjälpa oss att nå en insikt om hur vetenskapsmannens sociala miljö verkligen skapar de problem han/hon måste tampas med (Gorman, 1989:646-647).

8.2.2 Debatten mellan David Bloor och vetenskapsteoretiker

Bloor har också angripits för att hans teori inte lever upp till vetenskapsteoretiska standarder. Kritiken kommer framför allt från Larry Laudan som vänder sig emot Bloors påstående att vetenskapsteoretikerna inte ser att det finns något som orsakar sann kunskap. Laudan argumenterar för sin egen position genom att hålla med om

att det är sant att filosofer menar att sanna eller rationella åsikter inte skall tillskrivas sociala orsaker. Men han framhåller att få eller inga alls av de betydelsefulla filosoferna förnekar att rationella åsikter inte påverkas av sociala orsaker. De flesta filosoferna, liksom deras sociologiska motståndare, menar att åsikter är delar av de kausala nätverk som finns i världen (Laudan,1981:178).

Laudan talar uttryckligen om ett nätverk och inte om sociala faktorer. Det stödjer uppfattningen om att vi kan rädda tanken på de sociala faktorernas betydelse om vi ser dem som det som skapar ett nätverk av relativa positioner. I så fall måste vi verkligen överge tanken på att det sociala kan vara orsakande. Ty i ett nätverk sker påverkan på så sätt att nätverket exponerar genom att olika relationer framstår som mer meningsfulla. Det blir långsökt att dessa relationer kan ha en determinerande kraft. Det är rimligare att vi tänker oss att det är tillfälligheter som skapar möjligheter att se. Om det skulle handla om orsakande krafter borde vi istället för nätverk se fastslagna och svårföränderliga relationer.

Laudan tycks också vara skeptisk till kausaliteten när han lite senare hävdar att få filosofer verkligen insisterar på en kausal relevans i både epistemologiska och rationella överväganden (Laudan,1981:179). Han framhåller att han ifrågasätter om vetenskapen har några speciella metodologiska principer eller epistemologiska preferenser som klart skiljer den från andra icke-vetenskapliga former av kognition. Han framhåller också att åtminstone han är övertygad om att ingen, varken filosof eller sociolog, ännu har framlagt acceptabla redogörelser för vilka kognitiva eller metodologiska egenskaper som skiljer vetenskaperna från icke-vetenskaperna (Laudan,1981:180-181).

Den starkaste kritiken riktar Laudan mot Bloor's tredje tes. Den som handlar om symmetri i förklaringarna. Laudan hävdar att den går emot de allra mest etablerade prejudikat inom naturvetenskaperna. Han ställer frågan om i vilka vetenskaper alla händelser kan förklaras av samma sorters kausala mekanismer (Laudan,1981:182). Laudan framhåller att denna syn leder fram till att resonerande processer inte alls har någon påverkan på produktionen av några åsikter. Laudan menar att sådana krav är alldeles för starka för att kunna tas på allvar. Laudan frågar också varför Bloor har kommit fram till denna slutsats. Han menar att det har att göra med att Bloor har en alltför snäv modell av det rationella. Laudan hävdar att man kan ha en sociologi om det rationella som skulle kunna förklara varför i vissa kulturer en del saker räknas som goda argument.

Det leder till frågan om vilka sociala faktorer som spelar en roll för det sätt som rationaliteten själv utvecklas. Laudan föreslår att även den mest ambitiösa kunskapssociologen måste ansluta sig till tesen om rationell symmetri för att kunna vara öppen för att rationella åsikter kan ha sociala ursprung. Dock föreslår han att sociologins bästa väg inte är att argumentera för att rationella och irrationella åsikter har samma sorts orsaker. Den bästa vägen är istället att erkänna att rationella och irrationella åsikter mycket väl kan ha olika ursprung. Och att man

därifrån försöker utveckla sociologiska modeller för hur rationella åsikter formeras (Laudan,1981:189).

Om vi försöker placera in Bloors position i jämförelse med Luhmann så blir det svårt att se hur sociala processer hos Bloor kan förklara både det som är sant och det som är falskt. Om kausaliteten knyts till mänskliga handlingar kan vi se att det som är intressestyrkt förklarar varför mänskligheten inte har kunnat använda alla möjligheter att observera som finns i miljön. Om vi skall kunna förklara hur människan lyckats hitta nya och mer givande positioner utan att förklara det med att kontextfaktorer gjort en sådan utveckling möjlig, måste vi mena att det är människans rationella tankeförmåga som har gjort att hon vunnit över snöda intressen.

Luhmann talar istället om att förhållandet mellan system och miljö är asymmetriskt. Komplexiteten kommer från miljön och den relationen kan inte ändras (Luhmann,1995:182). Om vi sätter det förhållandet i relation till debatten mellan Bloor och Laudan betyder det att det är ifrågasatt om medvetandet kan tänka sig fram till ny kunskap. Människor är alltför inestängda i sig själva för att vi skulle kunna säga att de kan tänka sig fram till sanningen. Det är därför som vi behöver de reflexiva processerna. Ser vi på dessa reflexiva processer som ett uttryck för de sociala processer som skapar ny kunskap får vi ett argument för att sociala faktorer förklarar all kunskap. Vi bör emellertid uppmärksamma att när det gäller att skapa ny kunskap måste de sociala systemen interagera med sin omgivning. Det gör att även sociala kunskapsprocesser påverkas av faktorer som hör till den omgivning som ofta handlar om den fysiska naturen. På så sätt blir ändå de sociala processerna påverkade av det vetenskapsteoretiker skulle kalla naturvetenskapliga fakta. Eftersom det är miljön som erbjuder komplexiteten kan vi påstå att denna miljö har större betydelse än de interna processerna inom de sociala systemen. Detta innebär att det finns anledning att påstå att kunskapsprocesser är asymmetriska. Vi kan på så sätt hävda att Bloor både har rätt och fel. Sociala processer förklarar all sorts kunskap men processerna i sig är inte symmetriska.

Det betyder att vi aldrig kan se vad som i ett visst fall lett fram till en viss kunskap. När självreferensen kräver reflexion innebär det att processerna handlar om ett komplicerat samspel mellan system och miljö. Vad som påverkar mest avgörs av den kontext som bildas i en viss tid. En och samma uppfattning av ett kunskapsinnehåll kan då vid olika tidpunkter aktualiseras av olika delar i kontextens relationer. Detta gör att sociala processer alltid måste vara asymmetriska. Samtidigt betyder det att Luhmann faktiskt visar på att den svarta box som handlar om faktorer som styr människornas tänkande, inte handlar om inre mentala processer utan handlar om möjligheterna att observera.

Att sociala processer är asymmetriska innebär ändå inte det samma som att säga att sociala förklaringar inte kan vara symmetriska. Vi behöver skilja på förklaringar och processer. Det är den asymmetriska funktionaliteten i de sociala processerna som möjliggör sociala förklaringar till både det som är rätt och det som

är fel. När vi ser att reflexiva processer är beroende av distinktionen mellan system och miljö är det mycket lättare att förklara hur det sociala hör samman med världens empiri. Det sociala är i själva verket resultat av denna distinktion. Det är ett sätt att försvara att all kunskap som Bloor hävdar skapas av sociala faktorer men det är ett försvar som gör att symmetrikravet blir ganska trivialt. Det viktiga är att Luhmanns betoning av asymmetri gör det möjligt att framhålla att sociala förklaringar inte motsägs av att det är uppenbart att någonstans har faktorer i miljön att göra med att vi ibland tycker att någon kunskap är mer sann än en annan kunskap.

AI frågan gör att vi ser att mycket av kritiken mot det Starka Programmet ställs på sin spets och framstår som relevant. Programmets uppfattning om kausala sociala processer leder till att man har fokuserat för mycket på intressestyrning. Det intressanta med AI-debatten är att den ställer så många frågor på sin spets att sociologins uppfattning om det sociala måste förändras. Det är därför viktigt att vi avstår från den traditionellt kritiska sociologiska analysen av AI för att istället rikta in oss på AIs sociala funktionalitet. Vi skapar bättre förutsättningar för den sociologiska disciplinens reflexiva förmåga om vi frilägger vad som ligger bakom de sociala processernas asymmetriska förhållanden, än om att vi letar efter förklaringar till att intressen tenderar att likrikta dessa processer.

8.3 Luhmanns funktionella relativism och ontologisk realism

De problem som den konnektionistiska programmeringen respektive det Starka Programmet har för att kunna försvara sina positioner liknar varandra. Det handlar om att bägge måste kunna förklara hur sociala processer kan fungera utan att de behöver kopplas tillbaks till mänskliga handlingar. De måste då kunna förklara hur emergens uppkommer genom system/omvärld relationen. Den starka AI:n som bygger på konnektionistisk programmering behöver visa hur detta kan ske för att inte tvingas tillbaks till att den måste bygga på en modell där man försöker efterlikna en artificiell aktör. Detta är något vi har påstått att inte är möjligt eftersom det inte går att efterlikna en människa på artificiellt sätt. SPs problem är också att det behöver visa att sociala faktorer kan påverka kunskapsprocesserna utan att de behöver gå via enskilda personers tankeverksamhet. Hur kan SP förklara detta om det skall innehålla processer som innebär mera än bara enkel behaviorism? Och hur kan konnektionistisk AI modellera sådana processer som innehåller mera än enkel associationism?

Vi har redan varit inne på att den luhmannska teorins uppfattning om sociala processer, trots att den är utpräglad relativistisk, kan förklara hur världens

konsistens ändå påverkar kunskapen trots att denna kunskap utvecklas i sociala processer. I Luhmanns teori kan världen påverka kunskapen genom att den exponerar möjliga distinktioner. Därför blir det som finns i världen inte oviktigt. Luhmanns uppfattning om världen landar i en form av naturalism. Hittills har vi hävdat att det handlar om en relativism. Men det är en relativism som ligger nära den naturalism som till exempel Willard van Ormond Quine diskuterar i *Word and Object*. Och Quine presenterar själv sin position som en form av realism.

Att hävda en relativistisk position är mycket kontroversiellt. Det vi har argumenterat för i denna avhandling är inte en relativism som innebär att allt bara flyter omkring. Den innebär inte att allt är möjligt att hävda bara för att några människor tycker på ett eller annat sätt. Det skulle vara att säga att de sociala systemens referens är godtycklig och inte något som uppstår genom system/omvärldsrelationen. Däremot kan vi säga att mycket mer blir möjligt om vi rationaliserar systemen så att deras förmåga till reflexion blir bättre. Det är samtidigt en hållning som innebär att människan inte kan förändra allt. De psykiska systemen kan inte kommunicera med de system som bygger upp vårt universum och kan därför inte ändra på dessa system. Sannolikt kan inte heller sociala system påverka naturen även om den mänskliga kulturens påverkan på naturen kan tala för att en sådan påverkan kan ske. Sådana frågor ligger dock utanför vad som är relevant för denna avhandling.

Relativismen är bättre på att förklara de sociala faktorernas roll därför att den inte behöver blanda in vare sig medvetandet eller naturlagar. Med relativism blir det naturligare att vi använder begreppet funktionalitet istället för kausalitet, eftersom en relativ värld medger att vi inte behöver tala om naturlagar eller om att människan är grunden till allt. Skillnaden mellan funktionalism och kausalitet kan vara en definitionsfråga. Det har att göra med att det finns en gräzon mellan realism och relativism. Vi bör vara öppna för att det som vi upplever som funktionalitet hos relativa värden kan bestå av verkliga värden som är så finstämda och mångbottnade att vi inte kan uppfatta deras realistiska status. Då kan vi tala om kausalitet istället för funktionalitet. Fast i så fall har människan svårt att uppfatta hur dessa kausala processer ser ut. Därför måste vårt medvetande bli uppmärksammat på dem genom någon form av funktion som har att göra med att vi hamnar i situationer som tillåter oss att se något nytt. Kausaliteten måste leda till någon funktionalitet som tillåter oss att se. Det kan handla om en kausalitet som gör att systemen omorganiserar sig. Men det är den nya funktionalitet som denna omorganisering skapar, som gör att vi ser något nytt. Då innebär kausala processer att förutsättningarna för att se nya kunskapskopplingar förändras. Det betyder inte att någon ny kunskapsstorhet måste bli följden, eftersom kausaliteten bara leder till att en del av den gamla situationen förändras. Den kausala processen behöver inte ens leda till något nytt om resultatet inte blir en tillräcklig förändring i möjligheterna att se något nytt. Eftersom vi inte kan vara säkra på att det finns en sådan kausalitet är det bättre om vi fokuserar på funktionalitet. Eftersom vi heller

inte kan veta om det finns en yttersta verklighet bakom dessa processer är det bäst om vi knyter denna funktionalitet till en relativistisk världsuppfattning.

Det framstår som oklart om SP bygger på realism eller relativism på den ontologiska nivån. Det faktum att SP tar avstånd från att det finns sann kunskap som inte är påverkad av sociala processer innebär en form av relativism. Det behöver inte innebära att programmet inte kan omfatta att det kan finnas någon form av sann kunskap. Bloor menar också själv att SP inte står i motsättning till realism:

The relativism of the Strong Program is not to be counterpoised to realism. As I have emphasised, (non-social) nature plays a central role in the formation of belief, though how nature is experienced cannot provide a sufficient causal explanation of how it is subsequently described. (Bloor,1999:102)

Det blir svårt att förstå på vilket sätt de sociala processerna kan vara orsakande eftersom kausalitet bör vara något som leder i en viss riktning till skillnad från funktionalitet som mer letar efter möjligheter. Det är ju så som Luhmann skiljer mellan kausalitet och funktionalitet. Ett sätt skulle vara att utöka SPs realism till att också omfatta de sociala processerna. Vi kommer då tillbaka till att social kausalitet skulle ha att göra med att människan har vissa biologiskt nedärvt sociala beteenden. Det skulle då finnas några inneboende genetiska egenskaper som gör så att människorna under vissa yttre omständigheter beter sig på ett visst sätt. Men genom att dessa kausala processer skulle grundas i en realistisk ontologi blir det ännu svårare att undvika att programmet ligger nära ett renodlat sociobiologiskt program. Om programmet istället använde Luhmanns relativistiska funktionalitet kan det undgå också sådan kritik. Bloors krav på symmetrisk kausalitet är väl det som skall göra programmet starkt. Men kausalitet i den sociala dimensionen ger vid handen att människors beteenden är lagbundna. Funktionalitet är däremot förenligt med relativism i både sociala och kognitiva processer. Och om det är så att världen fungerar enligt processer som har att göra med en realistisk ontologi spelar det inte så stor roll om vi där tar miste, eftersom världen bara kan exponeras för människorna genom den interaktion mellan sociala system och miljö som bygger på relativistisk funktionalism. På så sätt slipper vi den sociologiska chauvinismen. Eftersom den sociala kommunikationen inte bygger på att människor tänker, slipper vi undan kritiken mot att programmet inte kan redogöra för vilken betydelse människornas resonerande kapacitet har. Vi kan istället se att sociala processer påverkar kunskapen genom de artefakter som är det socialas bärare. Eftersom dessa artefakter inte behöver ha någon komplex funktionalitet för att kunna förändra strukturerna kan också AI-redskap vara en del av dessa sociala

processer. Snarare kan vi se att AI kan tillföra en utökad funktionalitet till de sociala processerna.¹²

Vi kan framhålla att även om världen påverkar oss via kausala processer finns det ett utrymme för att social funktionalitet kan förekomma utan att kausala processer har varit verksamma. Det kan till exempel handla om att en person av en tillfällighet råkar exponeras för något som har funnits inom räckhåll men som personen av någon anledning inte tidigare har sett, bl.a. via den serendipity funktion som Merton sociologiserade. Det är troligt att AI-system kan effektivisera en sådan funktionalitet.

Brante har argumenterat för att det finns en skillnad mellan kunskapsteoretisk och en metodologisk (eller "sociologisk") relativism. Vi får anta att kunskapsteoretisk relativism är det som vi har benämnt ontologisk relativism. Brante menar att vi måste avvisa den eftersom den leder till den extrema filosofiska skepticism som hävdar att ingen kunskap är möjlig. Och till den idealistiska position som hävdar att världen bara är vår egen mentala konstruktion.

Brante menar att den relativism vi bör ansluta oss till endast är den metodologiska relativismen. Den kunskapsteori som sociologen bör ansluta sig till är istället realismen. Brantes argumentation lyder så här:

En stark kunskapsteoretisk relativism leder till en extrem filosofisk skepticism (ingen kunskap möjlig) och idealism (världen är vår egen mentala konstruktion). Sociologen bör ansluta sig till en version av ontologisk och kunskapsteoretisk realism. Även detta begrepp har många betydelser. Den svaga (och förnuftiga) versionen av realismen kan sammanfattas i två enkla punkter: (A) Det existerar en av medvetandet oberoende yttervärld, och (B) Det mänskliga tänkandet påverkas (begränsas) av denna yttervärld (Brante, 1984b:38).

Brante menar att den svaga varianten av realismen i själva verket är en förutsättning för den metodologiska relativismen:

¹² Det skulle vara intressant att ta upp sociologins debatt om vad Wittgensteins idé om regelföljande betyder utifrån Luhmanns teori. Bloor och Harry Collins tycks vara inne på att regelföljande skapas av kausala processer. Det kan då knappast handla om att kausaliteten kan förklaras med naturvetenskaplig realism utan den måste handla om någon form av mänsklig disponibilitet. De som Bloor kallar "Right Wittgensteinians" föreslår att regler skapas genom de situationer som uppstår då människor träffar varandra. Framför allt är det etnometodologins syn på regler som används. Det är framför allt den senare positionen som borde gå att utveckla med hjälp av en diskussion om artefakternas roll. Det finns en nyare diskussion om den debatten i ett nummer av tidskriften *Social Studies of Science*, Vol. 34, No. 4 (Aug., 2004).

Det vill säga: den metodologiska relativismens huvudpunkt är att kunskaper, sanna som falska, är orsakade; det finns inget logiskt samband mellan våra teorier om verkligheten och verkligheten. Orsakerna kan vara av socialt, psykiskt, perceptuellt etc slag. För vetenskapsmannen inom vissa ämnen, t ex experimentella sådana, är givetvis den (teoriberoende) observationen en primär kausal faktor. Alltså förutsätter relativismen existensen av åtminstone några orsaksfaktorer - en ontologisk realism. Relativismen förutsätter också samband mellan dessa faktorer och våra kognitioner - en kunskapsteoretisk realism. Däremot förutsätter vare sig relativismen eller realismen i ovanstående bemärkelser att det finns en entydig korrespondens mellan verkligheten och våra teorier om densamma, eller att det finns teorineutrala observationer, eller att alla de objekt vetenskapen uttalar sig om verkligen existerar, eller att den vetenskapliga kunskapsformen är artskild från andra kunskapsformer (Brante,1984b:38-39).

Frågan är väl ändå om sociologin verkligen behöver denna ontologiska realism. Enligt Luhmanns systemteori kan världen ändå bara påverka de sociala systemen genom de ganska komplicerade processer som sker genom strukturell koppling. Det betyder att människan inte kan observera världen genom sitt medvetande varför världens tillstånd inte kan bli en mental konstruktion. Om sedan världen är ordnad enligt kausala lagar eller relativ funktionalism är inte avgörande. Det är ändå bara genom funktioner som vi kan bli påverkade av världens tillstånd. Vi kan framhålla att det faktiskt betyder att vi kan hålla fast vid Brantes två punkter: att det existerar en av medvetandet oberoende yttervärld och att det mänskliga tänkandet påverkas (begränsas) av denna yttervärld. På så sätt förutsätter inte relativismen några orsaksfaktorer men den står och faller heller inte med att det kan hända att världen styrs av kausallagar. Fördelen är däremot att människans ontologiska status inte måste förklaras med naturlagar. På så sätt är den funktionella relativism som Luhmann föreslår en bättre förutsättning för den metodologiska relativismen än Brantes svagare variant av realism.

Utifrån ett funktionellt perspektiv blir också ett fenomen som emergens naturligare. Emergens används inom sociologin för att förklara hur sociala processer kan skapa förändringar i den sociala nivån utan att de behöver kopplas till mänskliga handlingar. Keith Sawyer diskuterar hur s.k. multiagent system kan erbjuda insikter om hur det som kallas emergens kan uppstå. Enligt Sawyer handlar det både om hur mikro-fenomen skapar förändringar i makro-nivåerna, och om hur makro-nivån skapar social kausalitet på mikro-nivån (Sawyer,2003).

Sawyer frågar hur makro-strukturella fenomen kan uppstå från individuella handlingar. Att man behöver ställa den frågan menar han är ett uttryck för att man har svårt att se att emergens kan ha sin grund i något annat än mänskliga handlingar. Sawyer menar att emergens bör höra samman med holistiska teorier

(Sawyer,2001). Han framhåller emellertid att emergens också har använts av metodologiska individualister för att förklara hur individerna och deras relationer bygger upp sociala egenskaper. Sawyer menar därför att begreppet är mycket suddigt.

Sawyer framhåller att emergens också har blivit ett nyckelbegrepp inom datamodellering av komplexa sociala system både inom konnektionism och inom multiagent modellering. Detta anger att det också finns möjlighet att grunda emergens just på att de komplexa processerna skapar en funktionalitet som innebär mer än vad enskilda aktörer kan åstadkomma. Det betyder alltså att det varken är lagbundna samhällsprocesser eller individuella aktörers handlingar som förklarar emergens. Istället framkallar komplexiteten i systemen den funktionalitet som skapar emergens. Vi bör uppmärksamma att Sawyer menar att regler som skapar emergens inte behöver vara svåra, vilket stöder tanken på att vi inte behöver ställa alltför höga krav på själva programmeringen. Sawyer framhåller istället att emergenta system är komplexa och dynamiska system som uppvisar beteenden som inte kan förutses genom att man försöker göra en fullständig studie av de delar som ingår i systemet (Sawyer,2001:554).

När vi diskuterade socionik-programmet uppmärksammade vi att DAI-modellen inte behöver kopplas till komplex funktionalitet på mikro-nivån, eftersom det inte är aktörers handlingar som är avgörande. Därför blir emergens något mycket mer naturligt. Det handlar helt enkelt om den relation mellan system och omvärld som inte alls har med mänskliga aktörer att göra. Thomas Malsch nämner i artikeln *Naming the unnamable* att vi skall undvika att leta efter en dialog mellan DAI och sociologin genom att diskutera begrepp på modet. Ett sådant begrepp anser Malsch att emergens är (Malsch,2001:171). Enligt den uppfattningen kan vi hävda att AI-strategier som bygger på att skapa artificiella aktörer som kan interagera med varandra, så kallade Multi Agent systems, bara leder oss in i en AI-diskussion som blir konstig. Den blir konstig eftersom den måste förklara begrepp som egentligen är onaturliga för en strategi som bygger på att allt skall grundas på enskilda aktörers handlingar. Då blir det motsägelsefullt att det samtidigt finns något som påverkar mikro-nivåerna genom en egen kraft. Emergens är ett sådant begrepp vars mening i en DAI-strategi representerar ett så naturligt fenomen att det egentligen inte alls behöver någon egen benämning eller ytterligare förklaring. Att emergens blir ett så besvärligt begrepp i alla modeller som har sin grund i individuella handlingar, har att göra med att det blir svårt att förklara hur strukturen kan få en egen funktionalitet på annat sätt än genom att människorna är grunden till allt. I en DAI-modell, eller en systemteoretisk modell, finns det mer funktionalitet i strukturen tack vare att den i sig blir mer föränderlig. Därför är det naturligt att makro-nivåer påverkar mikro-nivåer och vi behöver inget begrepp för den funktionalitet som bara sker i makro-nivåerna. Det ger också en möjlighet att knyta emergens till att samhällsprocesser bygger på relativism.

8.4 Luhmann om kodning och programmering

Vi har använt Luhmanns teori för att hitta argument för att artefakter kan ha att göra med sociala processer och att det skapar möjligheter att tala om stark AI. Luhmann själv är emellertid på flera ställen skeptisk till dataprogrammens förmåga till självständiga aktiviteter. Bl.a. i *Art as a social system* menar han att programmering är beroende av mänskliga insatser. Baecker framhåller i artikeln *Niklas Luhmann i der Gesellschaft der Komputer* att Luhmann hör till den generation som har svårt att riktigt helt anamma eller förstå innebörden i datatekniken (Baecker,2004).

Det som borde motsvara noder i konnektionistiska modeller är den del av programmering som Luhmann kallar *coding*. Vi har redan varit inne på att koder enligt Luhmann är strukturer som gör det möjligt att identifiera vissa företeelser. De är distinktioner som är hjälpmedel för att göra observationer, men de är samtidigt rörliga strukturer som används olika beroende på situation. På grund av att koden är binär (ja/nej) och stängd har den en viktig roll för differentieringen av ett funktionellt system. Den skapar på så sätt emergens i samhället. Kodens funktion är att symbolisera och på samma gång störa den grundläggande cirkulariteten i de autopoietiska systemen. Skall koden förändras krävs att det blir möjligt att göra en ny överkorsning (Luhmann,2000:187).

Vi bör observera att kodningen i *Art as a social system* görs beroende av hur koden uppfattas utifrån olika situationer. Koder blir till skillnader som fungerar som observationella mönster. Skillnaden mellan kodning och programmering försöker Luhmann förklara med att programmering strävar efter att ge ultimata värden som läses till ett beteende (Luhmann,2000:190-193).

Luhmann visar här på den ambivalens som hans teori är behäftat med gentemot det binära kodandet. Här framlägger han en syn på kodning som är mer fristående gentemot de mänskliga värderingarna av meningen i de binära koderna. Den binära dimensionen överstämmer därför mer med intressant/tråkig som Werber och Plumpe förespråkar. Vi skulle då inte behöva någon som aktivt värderar distinktionerna i ja och nej. Om de skall värderas utifrån intressant/tråkig blir det lättare för noder att reagera på möjligheter, vilket talar för att de kan vara artificiella. Det är intressant att den omprogrammering som sker i de konnektionistiska systemen skall vara något som sker fristående från den ursprungliga programmeringen. Programmering i de konnektionistiska systemen blir då likställt med kodning. Det handlar alltså om någon slags situationskapande aktiviteter. Konnektionismen ger ett perspektiv på programmering som gör att Luhmanns uppfattning om kodning kan göras mer fristående från programmeringen.

Det kräver då förstås att de konnektionistiska modellerna verkligen tillåter att moderna kan programmera om sig själva. Men om det kan ske finns det inte något i

de konnektionistiska modellerna som i princip står i motsättning till hur Luhmann ser på hur strukturell koppling sker mellan system och omvärld.

King och Tornhill visar på hur nära kodning och programmering hos Luhmann ligger operationer som sker inom vilket som helst funktionellt socialt system. De framhåller att denna syn först framträdde i och med att Luhmanns teori kompletterades med uppfattningen om autopoiesis i verket *Social Systems*. De framhåller att vi inte skall se kodning och programmering som reduktionistiska sätt att filtrera bruset från omgivningen för att göra den komplexiteten nåbar. Koder skall ses som speciella skillnader som existerar inom systemet och som bara kan hanteras av systemet. Och programmet erbjuder de filter för utan vilka koderna framstår som brutala försök att reducera allt till en värld som består av förenklade binära påståenden. Den stora skillnaden mellan program och koder är att program kan modifieras och ersättas men koden hela tiden förblir den samma och den som identifierar systemet. Det är systemets koder som tillåter den både att bestämma vilken kommunikation som hör till systemet och som tillåter en observation av båda sidor av distinktionen. Här framstår det ändå som om King och Tornhill gör ett misstag genom att de tillskriver koderna någon form av beständighet som närmast framstår som grundat i en realistisk ontologi. Om vi ser på hur Luhmann framställer relationen i *Art as a Social System*, till skillnad från den tolkning Tornhill och King gör baserat på det andra stället där Luhmann explicit diskuterar förhållandet mellan programmering och kodning, nämligen i boken *Law as a Social System* (Luhmann, 2004:173-210), så framstår det som att kodning i *Art as a Social System* är det som skapar öppningar medan programmeringen alltid är knuten till människornas begränsningar genom att de fastställer systemreferensen. Om vi tolkar den luhmannska teorin enligt King & Tornhill är det snarast programmets komplexitet som utvecklar systemens kommunikation så långt att ett system inser att det måste överge sig själv, eller transformeras till ett annat system. Det är en uppfattning som ligger närmare den klassiska AI-designen.

Kring en tolkning enligt King & Tornhill uppstår förstås en fråga om hur ett system verkligen kan se behovet av transformering. Problemet är att det är svårt att se hur programmen kan bli komplexa bara genom att försöka svara mot komplexiteten i omgivningen. Programmen kan bara observera omgivningen genom den kod som de använder och den koden är hela tiden den samma. På så sätt kan systemet inte alls förändras vilket nog är Luhmanns åsikt. Det kan bara gå under tack vare att förändringar i omvärlden gör så att systemet inte längre klarar sig mot andra system. Förändring skapas bara genom att nya system uppstår. Om vi istället kan se på koder som något som kan förändras kan vi tänka oss att det är koderna som skapar en sådan kontakt mellan systemet och omgivningen, så att systemet i sig självt kan utvecklas. Vi bör påminna om att i konnektionismen är koderna den enhet som förändras. Och som är den första nod som gör att programmen förändras. Det är genom att kodernas relationer till varandra förändras som programmen förändras.

Använder vi konnektionismens uppfattning om koderna ger de systemet en möjlighet att kommunicera med sin omgivning, därför att koderna i olika situationer ger nya observationsmöjligheter. Det är faktiskt ett perspektiv som skapar en möjlighet att utveckla Luhmanns kommunikationsbegrepp så att det också kan omfatta relationen mellan systemen och deras omvärld. Det är också på det sättet som vi kan hävda att intelligens handlar om en distribuerad egenskap. Det är en egenskap som inte bara behöver finnas inom ett system utan som uppstår genom att systemet kommunicerar med sin omgivning och därmed förändrar kodernas möjligheter att se, genom att observationsmöjligheterna förändras då nya situationer uppstår. I det fallet uppstår intelligens just genom att systemens omvärld ger möjligheter till att observera utifrån en ny synvinkel.

Om systemen, inklusive de psykiska systemen, bara kan observera utifrån samma kod hela tiden, blir slutsatsen att en människa liksom ett datasystem tvingas att förändras och stiga upp, likt en fågel fenix, som ett annat system om det verkligen skall kunna se på världen utifrån en ny kunskap. Det framstår som lite väl tillkrånglat och visar på ett visst drag av konservatism som teorin egentligen inte skulle behöva. Denna konservatism har bland andra Dag Österberg påpekat i en kritisk artikel om Luhmanns sociologi i tidskriften *Acta Sociologica* (Österberg,2000). Också den kritiken kan nog undvikas om vi ser på koderna som något mer föränderligt. Luhmanns uppfattning om förhållandet mellan programmering och kodning är nog styrd av hur symboliska programmeringsstrategier ser ut. DAI-perspektivet talar för att den relationen kan se ut på ett annat sätt.

I detta kapitel har vi närmast oss sociologins kärnpunkt. Hur sociala processer kan förklaras som något annat än som resultat av mänskliga handlingar. Vi har tagit upp kritiken mot det Starka Programmet och visat på att det mesta av den kritik som riktats mot programmet kan undanröjas om programmet istället för att försöka grunda sina argument i en ontologisk realism och i kausalitet omfattar den luhmannska sociologins relativistiska och funktionalistiska hållning. Vi har också visat på att den omformuleringen av det Starka Programmet gör att sociologin också bör omfatta den starka AI:n. Vi har dessutom visat på att den tämligen obearbetade hållning gentemot programmering som Luhmann själv redovisar, kan utvecklas med hjälp av den konnektionistiska DAI design som vi har visat på i denna avhandling. Det ger också en anledning att ifrågasätta Luhmanns negativa uppfattning om systemens förmåga att förändra sig själva. Den uppfattningen kan vara en orsak till varför den luhmannska teorin upplevs som konservativ.

9. Ett försök till syntes – det artificiella som funktionalitet i komplexa sociala processer

9.1 Kunskapssociologins relativistiska grundsyn

Den sociologi som är kritisk mot AI för att sociala processer handlar om att bearbeta det informella, tar fasta på att egenskaper i världen är så relativa att det inte finns några lagbundenheter som starka AI-program kan använda. Människans medvetande måste vara inblandat för att kunna skapa något nytt av all den kunskapspotential som ligger i det icke-formaliserade. Det är en kritik som gäller gentemot den klassiska AI:n. Men vi har i denna avhandling visat på argument för att den AI som bygger på konnektionistiska program klarar av att hantera det icke-formaliserade. Argumentationen bygger på att också medvetandet är begränsat därför att det är bundet till det formella. Bl.a. Carley m.fl. hävdar att social intelligens är något som hör samman med att människornas kognitiva kapaciteter är begränsade. För att nå utöver det som medvetandena kan uppfatta, behöver människorna något som kan exponera dem för nya möjligheter. Det är just den funktionaliteten som den sociala intelligensen enligt AI-förespråkare förmodas ha. Eftersom den intelligensen finns utanför medvetandet finns det heller inget som talar för att stark AI behöver stå i motsättning till social intelligens. Istället kan AI vara ett redskap som utvecklar den sociala intelligensen eftersom den utvecklar en funktionalitet som alltid har funnits i de kulturella kontexterna.

Argumentationen för relativism återfinns inom den moderna kunskapssociologin och handlar om hur språket och begreppen får sina funktioner. Kunskapssociologen Michael Mulkay åberopar vetenskapsteoretikern Mary Hesse för att diskutera hur begrepp förhåller sig till varandra. Han menar då att innebörden i termen *planet* ges av hur den används inom ett nätverk av relaterade termer och påståenden istället för att bara vara etablerad genom direkta referenser till en serie empiriska storheter som kan identifieras oberoende av detta nätverk. Mulkay visar att Hesse själv menar att detta betyder att kunskap innebär en

långtgående omtolkning av skillnaden mellan teori och observation (Mulkay,1979:34). Observera här att Mulkay via Hesse anger att begreppen får sin betydelse av i vilka relationer de står till varandra och att detta påverkar distinktionen mellan teori och observation. På så sätt uppkommer en funktion genom att vissa relationer skapas. Hesse säger själv om förhållandet mellan teori och observation, att omklassificering i princip kan ske var som helst där man gör observationer beroende på vilka interna begränsningar som det teoretiska nätverket lägger på vad man kan observera. Hon menar att det i varje givet stadium av vetenskapen finns relativt befästa observationer, men vilken som helst av dem kan senare förkastas för att upprätthålla helheten i det totala systemet (Hesse,1980:107-108).

Det framgår inte klart vad det är som framkallar omklassificeringen. Hesse talar om att hon varit inne på att det som inte verbaliserats förloras, men hon förefaller ha gått ifrån den tanken. Hon har antagligen från början utgått från att vi bara kan se det som människan beskrivit med sitt eget språk. Men hon har sedan kommit fram till att omklassificeringen kan ske var som helst där observationer kan ske. Då blir frågan om det betyder att även om omklassificeringen inte behöver gå via språket, så måste det vara människor som gör de observationer som kan leda till förändring. I systemteorin har vi hämtat stöd för att omklassificeringen kan ske utan människors inblandning eftersom den teorin menar att det inte behöver vara människor som observerar. Det skulle kunna gå att tolka Hesse i den riktningen eftersom hon menar att observationer förkastas för att de inte passar in i det totala systemet.

En del kunskapssociologer framhåller hur positivt det är att vi lever i en värld som vi inte direkt kan tolka. Till exempel Karin Knorr Cetina framhåller att osäkerheten inte är till någon nackdel. Istället menar hon att osäker information är förutsättningen för en progressiv organisering av systemet mot en ökande komplexitet (Knorr,1981:10). Knorr Cetina frågar sig vad det innebär för vetenskapen att behovet av osäkerhet gör att den måste omorganisera sig själv? Hon menar att det handlar om två skärningspunkter. Å ena sidan har vetenskapen förmågan att konstruera ny information genom att producera innovationer. Å den andra sidan kan vetenskapen rekonstruera sig själv för att kunna erbjuda lösningar på de problem som dyker upp. Förutsättningen för detta är att det finns osäkerhet. Denna osäkerhet menar hon består just av de frihetsgrader som ett system använder för att kunna omorganisera sig. Det förhållandet visas av att en observatör har svårt att i detalj redovisa vilka kriterier som ligger bakom en omorganisering som har skett (Knorr,1981:11). I Knorr Cetinas argumentering finns en stark betoning av att vi behöver leva i en kaotisk värld som vi inte kan tolka direkt. Det sägs inte direkt hur dessa omorganiserande processer ser ut, men argumentationen ger utrymme för många artificiella inslag. Så här långt kan vi säga att argumentationen understryker att en central fråga är om medvetandet tolkar situationen eller om situationerna själva ändrar sig.

9.2 Det sociala som situationism eller som ”environmentalism”

I de tidigare kapitlen har vi hävdad att sociologin måste knyta uppfattningen om det sociala till sådant som hör till de artefakter som bygger upp den mänskliga kulturen, eftersom det är situationernas betydelse som är sociologins viktigaste byggsten. Den holistiska sociologin får annars problem med att förklara hur sociala processer fungerar. Och den MI-orienterade sociologin får problem med att förklara varför människor tillsammans kan göra mer av det som inte är formaliserat än vad en enskilda tänkande och kalkylerande människa kan göra.

Vi har argumenterat för att Vygotskys teori om redskapens utveckling och Luhmanns uppfattning om kommunikation, ger oss en möjlighet att se att kontextfaktorer är del i en sociologisk förklaringsmodell. I detta kapitel skall vi se vilket stöd vi kan hitta i annan sociologisk eller samhällsvetenskaplig teoribildning för att utveckla diskussionen om situationernas betydelse. Dessutom skall vi jämföra Luhmanns systemteori med Habermas talhandlingsteori, eftersom en sådan jämförelse ger en möjlighet att diskutera om dessa situationer kan skapas genom systemprocesser eller om människor måste skapa dem med sitt språk.

9.2.1 Kontextens användning inom sociologisk diskussion

Situationismen flyttar fokus från att sociala processer måste grundas i mänskliga handlingar till att de har att göra med kontexterna. Den tanken har, som vi redan nämnt, utvecklats i AI-diskussionen av Lucy Suchman (Suchman,1987). Idén om situerade handlingar har också använts för att visa på att det går att göra en syntes mellan vetenskapsteoretikers betoning av kognitiva faktorerers betydelse för kunskapsprocessen och sociologernas betoning av sociala faktorer. Rosario Conte menar i artikeln *Social Intelligence Among Autonomous Agents* (Conte,1999) att socialt placerade men självständiga aktörer utvecklar speciella kognitiva kapaciteter för att möta de krav som ställs på dem från deras sociala miljö. Det kognitiva blir då något som utvecklas genom det sociala. Det är på så sätt som situationismen kan vara något av kunskapssociologins grundposition.

I mina försök att hitta tidiga sociologiska argument för situationernas betydelse har jag stött på en artikel skriven av C Wright Mills 1940 (Mills,1940). Artikeln heter *Situated Actions and Vocabularies of Motive*. Som titeln antyder handlar det om att se situationen som motiv för olika former av handlingar. Vi skulle kunna ta detta som en argumentering för att situationerna påverkar intentionerna. Det är dock svårt att få Mills resonemang att bära för vårt syfte

eftersom analysen stannar vid att peka på situationens betydelse.¹³ Det är först med Suchmans bok som situationismen används för att ifrågasätta om de processer som skapar situationer måste vara knutna till omedelbara mänskliga handlingar. Däremot visar Mills diskussion att det inte bör vara särdeles kontroversiellt att peka på situationismens betydelse. Dessutom finns denna uppfattning om situationernas betydelse också hos symboliska interaktionister. Joseph Gusfield säger bland annat att det finns fyra aspekter på den symboliska interaktionismens sätt att observera: mening; interaktion, emergens och **situatedness** [min understrykning]; språk och symbolism; samt den mänskliga tilliten (Gusfield, 2003:121).

Vill man ifrågasätta om människan verkligen är så suverän som AI-kritiken antar, bör man leta efter belägg för att det är situationernas rikedom som gör att människan kan uppträda som en intelligent varelse. Situationerna ses då som ett resultat av vår kulturella utveckling och som det som möjliggör fortsatt utveckling. I det perspektiv vi diskuterar frågan, handlar det om att litteratur och andra artefakter skapar våra möjligheter att uppträda som intelligenta människor.

Om vi ser på kulturen som det som ger individuell variation kan vi undgå behavioristiska förklaringar på så sätt att det som framkallar variationen istället har att göra med om situationerna kan göras tillräckligt rika. Dessutom har frågan att göra med om denna rikedom kan exponeras på ett effektivt sätt.

Det är knappast kontroversiellt att hävda att situationerna fungerar som bakgrunder eller motiv till att sociala processer sätts igång. Den kontroversiella frågan är istället om de själva kan ompositionera sig så att situationerna förändras. En intressant diskussion om motivens betydelse finns i Kenneth Burkes *A Grammar of Motives*. Den diskussionen handlar nog om att se situationerna som motiv för individuella handlingar, men i framställningen finns en del som talar för att situationerna i sig har funktioner oberoende av människor.

Burke föreslår där fem analysområden som har betydelse för motivet; Act, Scene, Agent, Agency, Purpose. Burke kallar detta för en pentad och argumenterar för den genom att hävda att man måste ha ett ord som benämner det som har hänt, ett annat namn för den scen som bildar bakgrund för handlingen, vilken sorts person (eller aktör) som utförde handlingen, vilka medel eller redskap som personen använde (agency) och avsikten (purpose). Burke framhåller att vi alltid måste ha något svar på vad som gjordes, var eller när det gjordes, vem som gjorde det, hur det gjordes och varför (Burke, 1945:xv).

Burke är en något udda vetenskapsman. Det är lätt att tolka hans argumentation i sociologiska termer men ursprungligen är han litteraturvetare. Johan Asplund, som vi använde för att presentera sociologins tveksamhet mot AI,

¹³ Denna betoning av situationer går för övrigt igen i Mills mest kända bok *Den sociologiska visionen*. I den handlar det mycket om att sätta fokus på situationer som skapas av maktintressen, men den sociologiska visionen handlar lika väl om att, som det står på baksidestexten till den svenska upplagan, klargöra sambanden mellan människornas upplevda problem och den historiska processen, att väva samman det biografiska och det historiska.

har ägnat en längre essä (Asplund,1980) åt att diskutera innehållet i *Grammar of Motives*. Asplund framför många synpunkter på Burkes text som visar hur nära den symboliska interaktionismen egentligen ligger ett situationistiskt perspektiv på det sociala. Vi skall därför se på delar av Asplunds framställning. Det är i sig intressant att se hur en kritiker av AI diskuterar motiv. På så sätt kan vi närma oss möjligheter att se att många av uppfattningarna om interaktionernas betydelse för att skapa kunskap egentligen bör vara argument för DAI.

Asplund framhåller att de fem frågorna - vad?; när och var?; vem?; hur?; varför? - som enligt Burke oundgängligen måste besvaras, hör ihop. Man kan inte välja en aspekt och betona den på bekostnad av de övriga. Burke menar, enligt Asplund, också att pentaden är fullt utvecklad. Ytterligare specifikationer tillfogar inget principiellt nytt - *femkanten kan inte utvidgas till en sexkant* (Asplund,1980:134).

Asplund diskuterar om Burkes fem grundbegrepp kan vara tillräckliga. Och om de är konsistenta sinsemellan. För denna diskussion menar Asplund att vi kan ta upp sociologiska eller psykologiska begrepp och diskutera om de faller in i Burkes pentad.

Det första begrepp Asplund tar upp är *medvetandet* eller *psyket*. Asplund menar att om medvetandet tillhör agenten så faller detta begrepp in under Burkes pentad. Han framhåller emellertid att många psykologer eller sociologer gjort vad de kan för att beröva agenten dess medvetande eller psyke. Asplund redovisar några alternativ att definiera medvetandet som han har funnit (Asplund,1980:134-135). Några har hävdats att medvetandet inte finns. I så fall menar han att medvetandet inte alls kan inordnas under Burkes pentad. Om det existerar så har det hävdats att medvetandet är instrumentellt och skulle höra till Burkes begrepp *agency*. Det har också hävdats att medvetandet sammanfaller med det yttre handlandet. Vilket närmast gör det till en del av begreppet *act*. En annan möjlighet är att medvetandet utgörs av någon slags ren intentionalitet så att det hör till *purpose*. Asplund menar ändå att den vanligaste uppfattningen är att medvetandet eller psyket är återförbart på *scenen*. Asplunds slutsatser är att *medvetandet* eller *psyket* inte är ett sjätte begrepp vid sidan av Burkes pentad. Det är dock inte heller möjligt att reducera det till något av de svar han hittills har prövat. *Medvetandet har sina sceniska betingelser, men det är inte liktydigt med dessa betingelser; medvetandet sammanhänger med yttre handlingar, men det är inte identiskt med yttre handlingar* (Asplund,1980:136).

Asplund tycks mena att medvetandet måste kopplas till alla fem element i pentaden. Att som Asplund säga att medvetandet hör samman med hela pentaden men inte liktydigt med den är ett starkt argument för distribuerad kognition! Hela Asplunds argumentation låter faktiskt som ett försvar för *Second order cybernetics*. Vi kan jämföra med Spencer Brown som menar att distinktionen står för motivet, innehållet och värdet i sig självt. Att säga att pentaden har sceniska betingelser anger att medvetandet är beroende av dessa sceniska betingelser även om det inte är identiskt med dessa betingelser. Medvetandet är kanske inte beroende av bara en

scen men då skall vi minnas att distribuerad kognition anger att det finns flera scener eller att själva scenen är distribuerad. Vi bör inse att frågan handlar om medvetandets kapacitet överlag. I tänkandet är kanske medvetandet inte så beroende av scenen, men om intelligens handlar om att kunna expandera kunskapen är nog medvetandet mycket beroende av scenen.

Den kritiska frågan för oss är om det som finns utanför människan bara kan skapas av en social aktivitet som människan måste vara en del av. För att belysa den frågan skall vi gå vidare med den begreppsliga diskussion som Asplund gör av innebörden av de fem grundbegreppen i pentaden. Asplund menar själv att en diskussion av översättningsproblem kan leda oss en bit längre.

Asplund redovisar att han översatt "purpose" med syfte. Han menar att vi lätt tänker oss att syften och avsikter oupplösligt hör samman med agenten och hans akter. Om vi istället översätter "purpose" med bevekelsegrund menar Asplund att sambandet mellan agenten och dess akter blir otydligare. Detta bör innebära att vi verkligen tonar ner de mänskliga handlingarna och lägger mer tyngd på det som hör till kontexterna. Asplund menar själv att vi har både akten och dess bevekelsegrund. Han menar dock att bevekelsegrunden snarare hör till aktens bakgrund. Asplund hävdar att detta leder till att vi är på glid från agenten. Det viktiga blir drivkraften vilken har en klart scenisk prägel (Asplund,1980:137).

Om begreppet *agency* säger Asplund att det kan översättas med *medel*, *instrument* eller *verktyg*. I förstone, menar han, tycks *agency* oupplösligt hör samman med agenten. Men när vi skall förklara vad ett verktyg är så hänvisar vi till det man gör med verktyget. Han säger att vi då talar om ett verktygs funktion. Det är ett synsätt som har många likheter med Vygotskys syn på redskap. Dessutom menar han att *agency* kan översättas med *agentur*. Asplund hävdar att *agenturs* mångtydighet kan göra det begreppet till den bästa översättningen eftersom *själva ordet "agentur" bevarar ju förbindelsen med agenten och handlingssfären, men samtidigt är en "agentur" ett ombudskap, varigenom den instrumentella aspekten kommer med* (Asplund,1980:138).

Utifrån ett resonemang om distribuerad kognition kan vi hävda att det är mångtydigheten som bygger upp den relation som skapar den potential som finns i begreppet agentur. På så sätt skapas en autonom roll för situationen. Det accentueras när Asplund går vidare och försöker diskutera hur vi kan översätta "scen". Vi kan säga *bakgrund*, *omgivning* eller *inramning* men det bästa svenska uttrycket är enligt Asplund *skådeplats* (Asplund,1980:139), vilket bör vara en form av situation som är laddad med funktion.

Det blir inte helt klart vad det är som tydligt visar skillnaden mellan scen och agent. Den frågan har att göra med den diskussion om konsistensen mellan de olika delarna i pentaden som Burke för. Asplund framhåller att även begreppet konsistens är tvetydigt. Han menar att i vardagsspråket betyder konsistens beskaffenhet men det kan också betyda överensstämmelse, förenlighet och motsägelsefrihet. Asplund föreslår att Burkes konsistensteori inte handlar om några

slags formella eller "utvändiga" relationer mellan element. Han menar att den snarare handlar om elementens ontologiska beskaffenhet, eller om "invändiga" relationer mellan dem, vilket är sådana relationer som gör elementen till det de är (Asplund,1980:143).

Vi kan tillägga att vi här måste vara klara på om vi synar frågan ur ett ontologiskt relativistiskt eller realistiskt perspektiv, eftersom konsistens har olika innebörd beroende på om vi utgår från realistiska eller relativistiska utgångspunkter. Om vi talar om konsistens i termer av realism bör relationerna vara mycket homogena och oföränderliga. Om begreppet tolkas relativt handlar konsistens snarare om att vi observerar relationerna genom det sätt de fungerar tillsammans. Det innebär att vi inte upptäcker dem utan upplever dem. Vi kan alltså inte tolka dem med medvetandet, utan vi kan bara uppleva dem genom det sätt de fungerar tillsammans. Det handlar då om en funktionalitet som är avhängig av relationer. Och eftersom vi inte kan tänka oss fram till det relativa kan den funktionalitet som finns i social intelligens inte komma från det mänskliga tänkandet. På så sätt kommer ändå Asplunds argumentation tillbaka till det som har varit en av de röda trådarna i denna avhandling. Att vi måste vara klara på om social intelligens handlar om att människorna är kapabla att hantera det icke-formella eller om dessa processer kräver att självständiga redskap kan exponera möjligheter. Asplunds diskussion visar på att konsistensteorin ändå handlar om så många utvändiga relationer att det blir svårt att hävda att allt behöver gå genom huvudet.

Asplund redovisar lite längre fram i texten att Burke påstår att behaviorismen börjar först när motiven har lett fram till en handling. *Behaviorismen reducerar nämligen handlingar till rörelser* (Asplund,1980:145). Asplund (och Burke) kan tyckas argumentera för att den symboliska interaktionismen skall förskjuta sin analys mot scenen och bort från själva handlingen. Det vore logiskt eftersom vi har sett att behaviorism har knutits till att människor bara följer andra människors handlingar.

Asplund vill ändå förklara scenen med handlingar. Han framhåller att Burkes konsistensprincip inte är en reduktionsprincip. Den är snarare en dialektisk princip. Problemet blir då hur vi kan se på en handling utan att reducera den till scenen. Asplund redovisar att Burke ställer frågan hur det är möjligt för en agent att vara god i en ond situation eller att utföra bra handlingar under dåliga betingelser.

Varje handling innebär någonting nytt. Jag tror att Burke menar att det är det nya som konstituerar handlingen. Innebär en handling ingenting nytt, så är den i själva verket ingen handling (utan bara "rörelse") (Asplund,1980:153).

För mig framstår det som att det inte kan vara själva handlingen som skapar det nya. Utifrån Luhmanns synsätt måste det finnas nya situationer som tillåter oss korsa över på nytt sätt. Finns inte dessa förutsättningar kan vi bara korsa över på samma sätt, i detta fall göra samma handlingar gång på gång. Därför kan det inte vara själva handlingen som konstituerar det nya. Vi skulle istället kunna säga att det avgörande är scenens mångfald.

Det kan förefalla som att tvista om petitesseer att fråga om det är handlingen eller situationen som skapar det nya. Men en diskussion om möjligheterna att utföra bra handlingar under dåliga betingelser ger en möjlighet att belysa problematiken. Om det är handlingarna som är avgörande så bör dåliga betingelser handla om att människor är mycket beroende av andra människors handlingar. Man är då så inestängd i vad andra människor tillåter och man får svårt att se möjligheter att bryta dåliga mönster. Om det istället är situationen som är dålig har människorna större möjligheter att aktivt leta efter andra situationer, t.ex. via litteratur, som kan göra att de kan jämföra den dåliga situationen med goda situationer. Helt enkelt hitta förebilder utanför den egna sociala gruppen. Då kan dessa andra situationer peka på att den aktuella situationen är så ond att människan måste göra en god handling för att motverka situationen. På så sätt leder den onda situationen henne till goda handlingar.

Det är alltså scenens mångfaldighet som gör att agenten faktiskt kan handla på många sätt i en situation. Om vi koncentrerar oss på situationen och ser på den utifrån hela pentad-perspektivet, bör det vara så att AI kan skapa en situation som faktiskt ger människan många fler möjligheter att handla. Den konnektionistiska varianten av AI bör kunna leda till att scenens rum blir så oändligt mycket större. Ytterst är det scenens möjligheter som avgör vilka handlingar människorna kan utföra. Det innebär att funktionaliteten i situationen kommer från pentadens mångfald, vilket också innebär att det är egenskaper i kontexten som skapar mer mångfald. Den individuella handlingen påverkas av denna mångfald och kan inte användas för att förklara samma mångfald.

9.2.2 Från embodiment till environmentalism

Som vi har sett handlar en del av debatten kring AI om att vi måste skapa artificiella aktörer som motsvarar de egenskaper som finns i mänskliga kroppar. Om man håller fast vid detta är det också svårt att hävda att situerade artificiella agenter kan innehålla de egenskaper som vi lägger i begreppet embodiment. Tom Ziemke, som är professor i datavetenskap vid Skövde högskola, har tillsammans med Tarja Susi, skrivit en artikel där de argumenterar för att tanken på embodiment bättre kan analyseras genom idéer om distribuerad kognition (Ziemke&Susi,2001).

Ziemke och Susi studerar hur artefakterna själva kan skapa situationer. För att komma fram till ett svar jämför Ziemke och Susi den av Vygotsky inspirerade aktivitetsteorin, teorin om situerad handling och distribuerad kognition. Ziemke och Susi menar att i distribuerad kognition ser man att både agenterna och artefakterna är delar av ett komplext kognitivt system. Det bör innebära att det som vi ser som ett komplext kognitivt system i själva verket är ett uttryck för den komplexa funktionalitet som finns i ett socialt system. Ziemke och Susi argumenterar också för att en del av de kognitiva processer som sker blir mer observerbara om vi flyttar fokus i analysen från att studera det som sker i huvudena på individerna till det system som består av agenter och artefakter. På så vis blir både agenten och artefakterna synliga så att vi kan se hela den struktur genom vilka representationer transformeras och blir uttalade. Distribuerad kognition kommer på så sätt att understryka den betydelse kognitiva artefakter har (Ziemke&Susi,2001). På så sätt kan vi också hålla med Ronald Giere när han menar att det inte behöver finnas någon motsättning mellan kognitiva och sociala faktorer i kunskapsprocesserna.

I en artikel som Ziemke skrivit tillsammans med Jessica Lindblom föreslår de att begreppet situerad kan förflyttas till den sociala världen. De menar att det bör ske genom att utvidga den fysiska miljön till den sociala miljön (Lindblom&Ziemke,2003).

Ett begrepp som i sammanhanget är intressant är *tyst kunskap*, som brukar härledas till Michael Polanyi. Han skiljer mellan *tyst kunskap* (tacit knowledge) och *uttalad kunskap* (explicit knowledge). Det finns många fördelar med att se att den tysta kunskapen finns utanför människorna. En viktig fördel är att *tyst kunskap* blir mycket effektivare om den inte måste vara inestängd i människorna. Det bör framhållas att människor ofta inte kan redogöra för den tysta kunskap de har. Det beror på att den finns lagrad i medvetandet utan att ha blivit internaliserad. Ofta framträder denna tysta kunskap inte innan någon yttre faktor gör att man blir påmind om denna tysta kunskap. Det ger vid handen att det handlar om en strukturell egenskap i hjärnan och att det finns anledning tro att det även i kulturella kontexter finns mycket kunskap som är *tyst* därför att det som bygger upp den tysta kunskapen finns lagrat på ett distribuerat sätt. Det handlar om en kunskap som inte är uttalad som en gemensam beskrivning eller som ett allmänt vedertaget begrepp men som ändå förmår skapa en kontextuell egenskap. Programvaror för *kunskapsmanagement* riktar bland annat in sig på att underlätta att sådan distribuerad kunskap kan föras samman. Sven-Eric Liedman framhåller i boken *Ett oändligt äventyr*, att *tyst kunskap* är förenad med sinnena och eftersom den inte kan uttryckas blir den ofta till en *tystnande kunskap* (Liedman,2001:117ff.). Liedmans perspektiv är annorlunda än det som vi har fört i denna avhandling, och i hans diskussion innebär snarast datoriseringen att kunskapen *tystnar*. Vi kan uppmärksamma att om *tyst kunskap* bara finns inom personer så ger det också en grund för att de personer som skaffat sig

maktpositioner inom samhället kan använda den tysta kunskapen för att värna om sin ställning. Den kunskap som är tyst behöver man ju inte dela med sig av vilket faktiskt är ett problem i större organisationer. Det är då en fördel om den tysta kunskapen istället finns lagrad i de kulturella kontexterna. Även om den är svår att tolka finns det åtminstone ingen som har kontroll över den. I det perspektivet kommer stark AI snarast att se till så att den tysta kunskapen kan nå fler personer och på så sätt kan den förhindra att den tysta kunskapen blir till en kunskap som tystnar.

Att AI-tekniken kan användas till att bättre exponera denna tysta kunskap har också att göra med att det i ett relativistiskt perspektiv finns mindre anledning att skilja mellan tyst och uttalad kunskap. Vi kan säkert påstå att uppfattningen om att det finns en skillnad mellan dessa båda kunskapsformer har att göra med att man ser en skillnad mellan objektiv och subjektiv kunskap. Om all kunskap är relativ så finns det naturligtvis ingen substantiell skillnad mellan den tysta och uttalade kunskap som en människa använder. Bägge uttrycker någon slags tillfälliga strukturella relationer. Skillnaden är att det i det ena fallet bara är en människa som har vant sig vid vissa kunskapsrelationer och i det andra fallet är hela samhället eller åtminstone ett kollektiv med människor som har vant sig vid att uppfatta vissa strukturella relationer som sanna. Den stora svårigheten handlar om att människan behöver exponeras för nya kunskapsrelationer, och i det sammanhanget är det en fördel om den kunskap som är "embodied" finns i den humanistiska kulturella kroppen istället för i den mänskliga biologiska kroppen. Om den finns i den kulturella kroppen kan exponering ske med hjälp av datasystemen, men om det förkroppsligade finns i mänskliga kroppar måste vi vänta på att träffa någon som genom personligt exempel kan visa oss på ny kunskap. Då kan vi bara lära av den som själv är kunskapsproducent. Det scenariot blir i sig tämligen icke-humanistiskt eftersom det förminskar texternas betydelse.

Den som tydligast har formulerat att *embodiement* handlar om en kontextuell egenskap är Andy Clark. I boken *Being There* talar Clark om hur begreppet *Embodied Agents* utvecklades till att bli *Embedded Agents*. Clark menar att kognition handlar om inbäddade relationer därför att verklig "embodied" intelligens måste handla om att koppla ihop sig med världen. Det är en strategi som tillåter att man kan lösa problem på ett robust och flexibelt sätt (Clark, 1996:98). Clark ligger också nära den luhmannska systemteorin när han talar om att vi måste överge tanken på att ett system representerar ett annat. Och att vi istället måste se på två system som är kopplade till varandra.

I den sociologiska debatten har Alan Wolfe, som vi tidigare såg att var kritisk både till stark AI och till den luhmannska systemteorin, försökt koppla ihop frågan om stark AI med debatten om sociobiologin, på ett sätt som anknyter till diskussionen om *embodiement* gentemot *environmentalism*.

I boken *The Human Difference* framhåller Wolfe att även Edward Wilson argumenterar för att kulturen har en betydelse. Wolfe argumenterar för att

kontexterna i form av den mänskliga kulturen ändå har att göra med medvetandenas kapacitet. Wolfe menar att Vygotsky framhåller att den kulturella utvecklingen är snabbare än den genetiska utvecklingen. Wolfe anger att också Wilsons idéer innehåller en sådan tanke (Wolfe,1993:45). Wolfe hävdar själv att den kulturella utvecklingen är så snabb därför att de mänskliga agenterna har de speciella mentala kapaciteter som möjliggör för dem att förstå (och därmed kunna ändra) den fart med vilken den kulturella variationen sker (Wolfe,1993:47). Han menar också att det betyder att vi måste betona medvetandet istället för hjärnor, vilket leder till att vi måste förändra synen på vad vi skall försöka förklara. Med hänvisning till George Herbert Mead menar han att det är begreppet medvetande som leder till att vi behöver ett begrepp för jaget. Medvetanden är inte samma sak som hjärnor. Medvetandets kvaliteter har visserligen något att göra med den datoriserande kraft som finns i kognition och logik. Mead underströk emellertid att det sociala jaget kompletterar den biologiska hjärnan genom att detta jag kan uppfatta de sociala fakta som ligger utanför hjärnan¹⁴. När vi talar om medvetandet behöver vi en teori som inte förklarar hur hela arter utvecklas. Den skall istället förklara hur individer gör val som varken är bestämda av sina gener eller av deras kultur (Wolfe,1993:49-50).

Därför menar Wolfe att sociobiologin ger ett motvilligt bidrag till en teori om medvetandet. Han menar att den inte på ett övertygande sätt förklarar varför människor i dagens USA klär sig olika, talar olika och handlar olika än de som levde i Mellanöstern för tusen år sedan. Skall vi förklara det måste vi sluta se på arternas utveckling och istället se på hur medvetandena har utvecklats hos de enskilda individerna

Such an explanation will take us away from the study of whole species and toward the analysis of individual minds. It will use the tools of a different kind of science, one that, instead of copying the rules of biology, will be cognizant of the specific and unique human powers of mind. What we gain by moving in that direction is an explanation of human evolution more powerful than such passive ones as time in itself, the sheer speed of change, or a minimal theory of mind. What we have to give up is the precision gained by models that operate by fixed mathematical laws borrowed from the sciences concerned with other than human species (Wolfe,1993:50-51).

Det vi bör observera är att Wolfe argumenterar enligt en tankegång som innebär att det ändå är något speciellt med det mänskliga medvetandet som inte handlar om att den mänskliga hjärnan är en oöverträffad maskin (vilken vi har sett att är

¹⁴ Vi kan observera att designen påminner om den som Carley m.fl. anger som argument för datoriserad social AI fastän slutsatserna ang. AI åtminstone för Wolfes del blir annorlunda.

Searles position). I överensstämmelse med den kritik vi tidigare har riktat mot AI-kritiska sociologiska positioner har vi visat på att det blir svårt att förklara detta mänskliga medvetande med annat än genetiska egenskaper om det inte är så att det finns en funktionalitet i de kulturella artefakterna i sig. Situationismen förklarar detta därför att vi med den kan se att kulturen har förändrats och därmed människornas förutsättningar att tänka. Artefakterna får en funktion därför att de skapar situationer som kombineras med andra situationer. Wolfe, och de som hävdar *embodiement*, får problem med att förklara hur det medvetande uppstått som han grundar sin argumentation på. Om vi inte fokuserar på egenskaper som hör till *environment* blir det svårt att förklara att en resonerande kapacitet hos människan på en så kort tid som några tusen år skulle kunna ha utvecklats så långt att människan skulle ha skaffat sig en så mycket större kulturell kapacitet än kring vår tideräknings början. Medvetanden som inte utvecklas på grund av att kulturen har utvecklats bör bara utvecklas genom biologisk utveckling.

Sociologin bör istället leta efter förklaringar till det sociala som har att göra med den kulturella fernissan. Att den har betydelse för utvecklingen av den mänskliga kulturen har i svensk debatt framhållits av före chefredaktören på DN, Hans Bergström. I artikeln *Biologi för humanister* hävdar han att föreställningen att biologism vinner terräng genom den nya biologin är felaktig. Istället påstår han att det i centrala avseenden förhåller sig alldeles tvärtom. Han menar att dagens biologiska forskning slår sönder fördomar och lyfter fram betydelsen av kulturen som det som ger människan hennes nutida särart. Han menar att den nya biologin slår sönder rasismen i dess linda eftersom den bland annat visar att rasblandning är nyttig. Han menar också att denna nya biologi förstärker den tendens som finns i samhället att minska betydelsen av de naturgivna krafterna, snarare än tvärtom. Genforsknigen ökar betydelsen av kulturen som definierande för människan. Bergström understryker att det blivit tydligt att kulturen blir allt viktigare för att definiera människans särart. Inte minst påpekar han insikten om hur mycket språk och text betyder. Språket har gjort det möjligt för människan att lära av många tidigare generationer, att ständigt ackumulera kunskap i stället för att behöva uppfinna det mesta på nytt under varje århundrade. Via språket kan abstrakta tankegångar formuleras och värdesystem diskuteras. Han refererar till Luigi Luca Cavalli-Sforza som i boken *Genes, peoples, and languages* framhåller att språket tar bort begränsningen till ett livsspann för att samla information. På så sätt accelererar språket också spridningsprocesser. Medan det tar tusentals år innan en genetisk mutation hos en individ hunnit sprida sig till många individer, så kan nya idéer vinna fäste snabbt. Vi bör kanske uppmärksamma att språket här handlar om det artefaktiska redskapet och inte det beteende som handlar om att människor "språkar" samman med varandra.

För att kunna förklara att det i datateknisk mening kan finnas en självständig funktionalitet i kontexterna bör vi återvända till de genetiska algoritmerna. De algoritmer som Carley m.fl. försökte knyta till den sociala intelligensen. Frågan

gäller inte bara om de kan vara modell för den funktionalitet som finns i kontexterna. Frågan gäller också om de kan representera den funktionalitet som vi har argumenterat för att borde ersätta det Starka Programets kausala processer.

I *Darwin's Dangerous Idea* tar Dennett upp frågan om huruvida intelligens kan handla om processer som kan formuleras i algoritmiska modeller. Dennett argumenterar emot matematikern Roger Penrose som i boken *The Emperors New Mind* påstår att intelligens inte kan skapas genom algoritmiska processer.

Dennett argumenterar för sin sak genom att nedtona betydelsen av algoritmer som styr processerna. Han menar att algoritmer bara är en del av den heuristik vi behöver för att kunna se möjligheter (Dennett, 1995:210). I och med det har de inte den funktionen att de leder fram till en förutbestämd kunskap. De hjälper inte till att leta efter kausala regler som leder oss fram till den nya kunskapen. Penroses argument bygger på det så kallade Gödel-teoremet som säger att det är omöjligt att en viss kunskap kan följa av en annan kunskap, eftersom all kunskap bygger på annan kunskap och att det därför inte finns någon grund att bygga på. Dennett menar att Penroses kritik av AI, och kritiken mot neo-darwinismen, bygger på att de tror att både AI-strategierna och neo-darwinismen är beroende av lagbundna och algoritmiska processer. Istället skall AI-programmen fungera som det heuristiska hjälpmedel som skall tillåta att den kunskapsmiljö vi lever i har en större chans att påverka vårt eget tänkande. På samma sätt fungerar de processer som enligt evolutionsteorin gör att vi anpassar oss till de möjligheter och de hot som kommer från den natur vi lever i. Här anknyter Dennett till den fråga som dök upp när vi diskuterade den holistiska sociologins förhållande till sociala processer. Det är frågan om vad som skiljer kausala processer från funktionella processer.

I alla avseenden bör det vara lättare att argumentera för att denna algoritmiska heuristik räcker om vi lämnar Dawkins realistiska utvecklingsbiologi och istället lutar till Luhmanns relativistiska systemteori. Det framstår ändå som om Dawkins syn på evolutionsprocessen som ändamålsenlig, gör att han måste se på evolutionen som en kausal process. Visserligen menar Dawkins att det inte finns något mål för evolutionen. Ändamålsenligheten handlar bara om att svara upp mot de villkor som ges av miljön vid en viss tidpunkt. Men ändå menar Dawkins tydligt att dessa processer skall grundas i en realistisk ontologi.¹⁵ Då måste det vara naturens kausallagar som styr evolutionens processer. Vi har sett att Gould istället vill betona att mycket handlar om faktorer som vi närmast kan hänföra till att slumpmässiga omständigheter har påverkat. Om anpassning sker till en värld som hela tiden förändras genom att aktivitet leder till förändrade relationer, blir det som vi kallar slump mer naturligt. Och då behöver vi heller inte förklara ändamålsenlighet med

¹⁵ Dawkins säger bland annat i boken *River Out of Eden* att den som säger sig vara relativist och sitter i ett flygplan på 30.000 tusen fots höjd är en hypokrat eftersom flygplan är byggda enligt vetenskapliga principer. s. 32

hänvisning till en skyhook. Vi har visserligen redovisat att Luhmann explicit tar avstånd från att funktionella processer är ändamålsenliga. Ändå bör vi observera att en ändamålsenlighet som har att göra med relativa processer bör kunna omfattas av Luhmann eftersom rationalitet har att göra med en utökad förmåga att kunna skapa reflexion. Ändamålsenlighet uppfattat på detta sätt blir fri från den determinism som kritikerna av utvecklingsläran i den nydarwinistiska tappningen ofta tar fasta på.

Det borde vara naturligare att argumentera för att olika rytmer i människans kulturella utveckling beror på att vid vissa tider har redskap utvecklats som påskyndat den kulturella utvecklingen, än att det som Wolfe m.fl hävdar, handlar om att medvetandena utvecklas. Det kan handla om tryckeriteknikens utveckling och utveckling av redskap för produktion av massmarknadsvaror. Eller att Internet i kombination med AI-teknik, d.v.s. DAI, har skapat redskap som starkt påverkar människornas förmåga till intelligenta beteenden.

Om vi vill understryka att detta resonemang har en koppling till de kulturella förhållanden som den nya datatekniken för med sig, erbjuder Dennett ett exempel. Han talar om att den traditionella synen på hur AI-modeller fungerar, utgår från den modell av biblioteket i Babel som den argentinske skönlitterära författaren Jorge Luis Borges (1899-1986) formulerade i essän *Biblioteket i Babel* (Borges,1998). En uppfattning som anger att det finns givna relationer mellan kunskapsstorheter i form av all världens litteratur. Likaså ser man ofta på neo-darwinismen som ett argument för att det är naturgivna regler i det mendelianska biblioteket (d.v.s. genbiblioteket)¹⁶ som säger vilken väg utvecklingen skall gå. Vilket felaktigt skulle innebära att den biologiska utvecklingen har ett mål. Dennett talar istället om det bibliotek som han kan nå med sin dator. Det handlar om den förlängning av hans intuitiva förmåga som datorn medger. Dennett påpekar att Penrose påstår att matematiker använder sin matematiska insikt för att se vilka påståenden som följer av sundheten i ett bestämt system. Och att det inte kan finnas någon algoritm som beskriver denna matematiska insikt. Dennett menar däremot att Penrose bortser från att det kan vara så att det är många matematiska algoritmer som skapar denna insikt, trots att de inte ursprungligen var avsedda för detta (Dennett,1995:439). Ur denna avhandlings perspektiv skulle då

¹⁶ Mendel, Gregor 1822–84, österrikisk ärftlighetsforskare, från 1868 abbot i ett kloster i Brunn (nuv. Brno i Tjeckien). I klosterträdgården genomförde M. försök med korsningar mellan ärtsorter, och han framlade en teori och en modell för hur egenskaper nedärvs. Föräldraplantorna i M:s korsningar skilde sig i olika egenskaper och kunde t.ex. bestå av plantor med runda, gula frön och plantor med skrynkliga, gröna frön. M. fick då i första generationen enbart runda, gula frön och i den andra alla fyra kombinationerna i bestämda proportioner, vilket visade att arvsanlagen kombinerats slumpvis. M. förutsåg härav förekomsten av arvsfaktorer som i olika kombinationer bestämmer avkommans egenskaper, och han införde termer och beteckningar som senare blev vanliga inom genetiken. M. betraktas som genetikens grundläggare, men hans banbrytande uppsats, *Versuche über Pflanzenhybriden* (1865), uppmärksammades först efter hans död. Ur: Nationalencyklopedin <http://www.ne.se.ludwig.lub.lu.se/artikel/254381>.

Penrose missa den distribuerade karaktären av dessa algoritmiska processer. Dennetts egen slutsats är att den algoritm som eventuellt kan förklara hur medvetandet fungerar, inte är synlig eller igenkännbar för dem vars medvetande den skapat.

As a product of biological design processes (both genetic and individual), it is almost certainly one of those algorithms that are somewhere or other in the vast space of interesting algorithms, full of typographical errors or “bugs,” but good enough to bet your life on—so far. Penrose sees this as a “far-fetched” possibility, but if that is all he can say against it, he has not yet come to grips with the best version of “strong AI” (Dennett,1995:443).

Det går förstås att hävda att den kulturella påverkan på medvetandena skulle ha skett i en process som närmast påminner om den process där nya egenskaper tas fram genom avel. Det är ju en process som bryter med det naturliga urvalet på så sätt att det är människan som gör ett urval på basen av egenskaper som anses önskvärda. Det lär inte ta många generationer innan man på så sätt kan odla fram en ny egenskap. Då kan vi tänka oss att intelligens utvecklas genom att de personer som kunnat använda de kulturella redskapen till att skaffa sig bra positioner är de som förökar sig bäst och som ser till så att medvetandenas kapacitet i biologiskt hänseende utvecklas. Även om en sådan utveckling luktar rasbiologisk politik behöver det inte vara så. Det kan helt enkelt vara så att det stora flertalet människor under de senaste generationerna har haft förutsättningar att utveckla sina medvetanden på detta sätt. Det skulle då vara en medveten politik att utveckla ett mer jämlikt samhälle som avlar fram folk med mer kapabla medvetanden. Det bör vara sannolikt att sådana medvetna politiska processer bryter mot det naturliga urvalet på ett sätt som innebär att det inte behöver gå så många generationer innan en hjärna som bättre kan hantera AI-tekniken utvecklas. Det skulle kunna vara ett stöd för Wolfes uppfattning om att medvetandenas kapacitet utvecklats sedan Jesu tid till våra dagar. Jag är inte tillräckligt skolad i genetisk avelsteknik att avgöra om den relativt korta tid vi har levt i ett någorlunda jämlikt samhälle räcker till för att skapa en sådan utveckling. Men det är väl inte riktigt så att det bara är egenskaper som har med intelligens att göra som premierats under vår tideräkning. Olika mer kroppsliga egenskaper som idrottsprestationer har också premierats. Går vi några århundraden tillbaks har väl den fysiskt välväxte krigaren också haft en stark position. Och under den tidsrymd som Wolfe hänvisar till finns det många belägg för att den dogmatiskt troende människan har haft det väl förspänt. Det har väl inte alltid under kristendomens historia varit så att den rationellt tänkande människan har premierats.

Det må vara hur som helst med möjligheterna att förklara biologisk utveckling av medvetandena. Den utvecklingen bör inte ha kunnat ske om inte den

första noden i den utvecklingen handlar om att de kulturella strukturerna har utvecklats. Om det inte handlar om biologi så finns det förstås också den möjlighet som Wolfe antagligen tar fasta på och som handlar om att medvetandena utvecklas i dialog med varandra. Jag hoppas att denna avhandling har visat på hur svår en sådan dialog är om man inte har tillgång till kulturella redskap och lever i en kulturell miljö. Det är som om Wolfe argumenterar för att tabula rasa människan kan utveckla intelligens bara genom att de samtalar med varandra. Det låter rimligare att den första noden i denna utveckling istället kommer just från kulturen. Det kan vara så att människan kom att använda en vass sten av en slump när hon försökte döda ett djur och märkte att det gick lättare än med en rund sten. Det borde betyda att den första noden finns i redskapen. Att människan sedan säkert har kunnat utveckla redskapen kan också ha att göra med att dessa redskap fått dem att kunna observera fler möjligheter. Stark AI kräver inte att det är den tekniken i sig som har all funktionalitet. Frågan kräver bara att dessa redskap har en funktionell kapacitet som gör att de kan påverka människornas förhållande. Ett litet starkare krav är att de också har en funktionalitet som människorna själva inte klarar av. Om Wolfe hävdar att det bara handlar om att medvetandena utvecklas oberoende av redskapen så kan det heller inte handla om att kulturen har en funktion som gör att det finns anledning förvänta att det uppstår den slags avelsfunktion som påskyndar medvetandenas biologiska kapaciteter. I så fall kan det bara handla om det naturliga urvalets långa process.

Om vi ser på hur individer förändras genom datoranvändning så handlar det så gott som helt om en utveckling som skett inom en generation. Konnektionistisk AI lanserades så sent som 1986. Det talar för att kulturell utveckling, i.e. utvecklingen av AI-system, på ett självständigt sätt påverkar människornas kapaciteter. Det bör alltså vara så att vår förmåga att hantera datatekniken helt hör till ontogensen.

Det är viktigt att observera att frågan om stark AIs kapacitet och frågan om AIs förmåga att interagera med människan egentligen handlar om två olika aspekter på utvecklingsbiologin. De konnektionistiska AI-programmens förmåga att omprogrammera sig är en fråga som bör motsvara den rena biologiska utvecklingen av arterna. Genen som den minsta enheten byts bara ut mot koden. Vi bör observera att bägge utvecklingarna styrs av en annan funktionalitet än den som kommer från den mänskliga hjärnan. Vi bör också observera att det i bägge fall handlar om processer som kan beräknas med matematiska modeller. På så sätt är den starka AIs förmåga något som måste förklaras med samma modell som Dawkins använder för att försvara Darwins utvecklingsteori. Alltså genom att användning och anpassning gör att den starka AIs förmåga utvecklas, vilket naturligtvis i sig innebär att människors användning av tekniken påverkar den starka AIs utveckling men då som en del av många kontextfaktorer.

Frågan om hur människan interagerar med tekniken är däremot en fråga som kanske har närmare beröring med om det finns anledning att tro att utvecklingen

av mänsklig intelligens inte varit kumulativ. Frågan visar då på att det kan finnas en del i AI-tekniken som talar för att den tekniken i sig kan vara något som gör att mänskligheten kan ta ett nytt stort språng. Alltså något som ligger nära det som Gould hävdar. Skall däremot Wolfes tes hålla så betyder det att kulturen i sig bara är en passiv faktor som förstärker positionerna för de människor som är bra på att använda sina medvetanden. Vi kan observera att det är en uppfattning som går stick i stäv mot uppfattningar om att t.ex. en utveckling av skolsystem påverkar alla människor att bli mer kapabla. Vi är då tillbaka till en del av de mer brutala aspekterna på sociobiologin. Om vi istället skall hävda att det har en emanciperande betydelse att kulturen utvecklas handlar det om att den i sig utvecklar människornas kapaciteter. Det påstås att samhället inte längre utvecklas i jämlikare riktning. Är det så kan det bero på att vi underskattar den möjlighet till utveckling som har att göra med att artefakter som AI kan användas för att vidga förmågan till intelligens. Det betyder i så fall att den kulturella strukturen ändå har en egen kraft. Det är åtminstone det som är innebörden i att stå på sådana axlar som består av det som någon annan skrivit.

Diskussionen om arvsanlag påminner om att sociologins divergerande åsikter gentemot AI liknar den konflikt som fanns mellan olika företrädare för utvecklingsbiologin i slutet på 1800-talet. Thomas Brante visar i boken *Vetenskapens sociala grunder* (Brante, 1984b:53-61), att det fanns en konflikt mellan mendelianer och andra utvecklingsbiologer om hur arvsanlagen utvecklades. Brante menar att bägge sidorna stred mot varandra utan att någon sida egentligen tog del av den andra sidans argument. Brante menar att det är ett exempel på att sociala faktorer kring kunskapsproduktionen skapar konflikter som inte har grund i verkligheten. Det var först när en forskare i början av 1900-talet med hjälp av ny empirisk forskning visade att de bägge förhållningssätten inte behövde stå i opposition till varandra som konflikten kunde biläggas. Denna konflikt framstår som ett bra exempel på en process där sociala faktorer inte bidrar till kunskapsutveckling utan håller utvecklingen nere. Det är ett exempel på en brutal social funktionalitet till skillnad från den mer finstämda som ligger bakom den sociala intelligensen.

Sociologins övervägande negativa hållning kring frågan om stark AI tycks representera en modern kopia av debatten kring arvsanlagen för mer än hundra år sedan. Vi ser exempel på att man sällan tar till sig den andres argument. Jag hoppas att mitt bidrag kan hjälpa sociologin att se att företrädare för stark DAI bör vara sociologins vänner. För sociologin är det viktigt att vi för frågor om kulturell utveckling vidare. Diskussionen om stark AI kan vara ett bra sätt för sociologin att förklara hur kulturella och samhälleliga faktorer kan kopplas till den utvecklingsbiologiska diskussionen

Innan vi lämnar denna fråga om utvecklingen av människans kapaciteter bör vi också nämna det som kallas *singularity tesen* (Vinge, 1993). Den går ut på att AI-systemen blir så starka att de kan få en övermänsklig intelligens och på så sätt bli ett

hot mot den mänskliga rasen. Det är i och för sig ett argument för att AI verkligen kan vara stark. Och frågan om detta kan vara ett hot mot mänskligheten är i sig inte relevant för denna avhandlings syfte. Perspektivet kan dock användas för att ifrågasätta om det är sunt att vi överhuvudtaget skall försöka utveckla så potenta redskap som stark AI.

Jag kan tänka mig dessa system kan kontrollera människorna på två sätt:

Vi kan skapa robotar som vi inte bara tilldelar intelligens utan de tilldelas också sådan teknik att de fysiskt kan hota oss. Det kan handla om att dessa robotar sätts att kontrollera kärnkraftsvapen. Men då skall vi observera att det inte är intelligensen i sig utan kapaciteten till våldsutövande som är avgörande. Vi kanske bör observera att vi brukar hävda att det är vissa samhällsstrukturer som gör så att människor blint följer dåliga ledare. Det är sannolikt att användning av redskap som stark AI inom kunskapsprocesserna gör att många fler får kapacitet att motsätta sig dåliga ledare.

De positiva egenskaperna hos stark AI kan naturligtvis förbytas i negativa om dessa AI-system börjar producera kunskapslänkar som inte leder till att ny kunskap kan observeras. På så vis skapas en kulturell struktur som är utarmande istället för expanderande. Frågan är väl om dessa AI-system verkligen förmår etablera sig om vi märker att de bara producerar tomhet. Att vi idag är så intresserade av dem handlar också om att vi faktiskt har märkt att vi har användning av dem. Det är förstås möjligt att de så sakta börjar leda oss in på banor som gör att våra kapaciteter förminskas utan att vi är medvetna om det. I så fall handlar det om att stark AI i sig inte fungerar. Det finns anledning tro att det ändå finns tillräckligt många som inser att dessa system inte fungerar och som slår larm. Att saker kan gå snett måste alltid människan vara medveten om och det finns redan skäl till att varna för sådan utveckling som människan styr utan att några AI-redskap har varit inblandade.

9.3 Luhmanns systemteori eller talhandlingsteori

Hans Bergström visade på språkets betydelse som en kulturell kontextfaktor. Det är en bra utgångspunkt för att diskutera om den kulturella strukturen, bl.a. i form av språket, kan positionera om sig själv utan hjälp av människan. Vi kommer tillbaks till om vi bör se på sociala processer som talhandlingar eller som systemkommunikation. Den frågan kan vi analysera utifrån den diskussion som förts mellan Habermas och Luhmann angående kommunikationens väsen.

Handlingsteori och systemteori har oftast uppfattats som konträra sociologier. Hävdar vi att kognitioner handlar om distribuerade funktioner så har vi argument för att handling inte är det centrala för att förstå intelligens. Luhmanns betoning av att det är kommunikationen som skapar de sociala system inom vilka de psykiska systemen handlar, talar också för att systemteori snarare är en förutsättning för handlingsteori (Stichweh,2000:5).

Rudolf Stichweh framhåller att Luhmann i polemik mot andra hävdar att utvecklingen efter Durkheim, Weber och Parsons innebär att det inte finns någon motsättning mellan handling och system. Istället menar Stichweh att handling och system måste gå hand i hand i sociologiska teorier. Även i Luhmanns teori är handling en del av de sociala systemen, men den kritiska frågan handlar om vem som är ansvarig för en handling (Stichweh,2000:8).

Stichweh spårar Luhmanns syn på komplementariteten till hans syn på att systemteorin egentligen är en kommunikationsteori. Speciellt har det att göra med den utveckling av kommunikationsteorier som skedde efter Shannon och Weavers bok *The Mathematical Theory of Communication*. I den boken relateras information till oförutsägbarhet. Det menar Stichweh så småningom ledde Ruesch och Bateson fram till att i boken *Communication: The Social Matrix of Psychiatry* göra en analogi mellan information och entropi. En analogi som så småningom formulerades i den formel som vi tidigare bekantat oss med - *information is a difference which makes a difference* (Stichweh,2000:7). Det är helt i överensstämmelse med den idé som vi har förfäktat i denna avhandling att informationens relativa värde implicerar att människorna inte tänker sig fram till kunskap. Vi behöver kommunikation för att komma fram till det nya samtidigt som kommunikationen föregår tänkandet. Vi bör också understryka att fastän detta sätt att uppfatta information som osäkerhet är hämtat från matematisk informationsteori, så skiljer sig Bateson, Ruesch och Luhmann från Shannon och Weaver på så sätt att det inte är något problem med att information representerar osäkerhet. Istället är det osäkerheten som skapar möjligheter. På så sätt blir entropi inte till något som är negativt och som måste omarbetas till redundans.

Att Jürgen Habermas ändå vill argumentera för att kommunikation måste handla om att människor samtalar med varandra bygger på att han menar att Luhmann skär ner språk och kommunikation till en så liten nivå att kulturell reproduktion, social integrering och socialisering försvinner från vår åsyn. Därför kan Luhmann inte omfatta att en intersubjektiv och ömsesidig förståelse uppnås genom uttryck som har gemensam mening och kritiserbara validitetskrav. Habermas menar att hos Luhmann existerar språket bara som bärare av symboliskt generaliserade tidigare förekomster av mening.

Luhmann's strategy is clear: If the achievement of linguistic symbols is exhausted by articulating, abstracting, and generalizing prelinguistic conscious processes and meaning connections, communication carried on by linguistic means cannot be explained in terms of specifically linguistic conditions of possibility. And if language can no longer be taken into account as a structure that makes possible the intrinsic connection among understanding, meaning, identity of signification, and intersubjective validity, the path of language analysis cannot lead to an explanation of understanding expressions with identical meaning, of consensus (or dissensus) about the validity of linguistic utterances, of the commonality of any intersubjectively shared context of meaning and reference - that is, to an explanation of communicative participation in a lifeworld that is represented in a linguistic world view. The aspects of linguistically generated intersubjectivity must be derived, rather, as self-generated artifacts from the reciprocal reactions among meaning-processing systems. Here Luhmann makes use of figures of thought familiar from empiricism (Habermas,1987:380).

Habermas framhåller att Luhmann menar att språket används för kommunikationen. Men det erbjuder inga uttryck via identiska signifikanter. Tecken kan bara vara substitut för mening (Habermas,1987:381).

Luhmann kan försvara sig med att hänvisa till Spencer Browns formlogik och hävda att ett tecken som en grund för kommunikation inte behöver vara något komplicerat:

“Sign” is itself a sign. Thus semiotics, since it contains itself as one of its own objects, is an autologically constituted science. Furthermore, a sign is not a thing, but rather a distinction. Accordingly, being autologically founded amounts to being based on a difference. This has consequences for determining which conceptualization of the world at large semiotics might reasonably find acceptable. The world can no longer be considered the entirety of signifying and signified objects and the relations between them. Concepts of substance, as well as concepts of relation, are dissolved by the interpretation of the sign as form, or they occur - to the extent that they are still used - only as signs. Hence, it is possible to signify the world only as that which is bifurcated by the form of the sign. That is to say, one must presuppose the “unmarked state” in order to begin positing signs as a specific way of making distinctions. This notion of the “unmarked state” again refers to Spencer-Brown's calculus of form (Luhmann,1999:56).

I en konversation mellan Luhmann och Katherine Hayles ställer Hayles en fråga till Luhmann som handlar om huruvida man kan tala om en observation utan att någon använder ett medvetande för att göra observationen. Hon framhåller att Luhmann i en artikel har framhållit att en observatör kan vara sådant som en cell, en person o.s.v. Hon säger sig ha svårt att förstå att en cell kan observera på det sätt som hon uppfattar att en människa observerar (Luhmann et.al,2000:115-116).

Luhmanns svar understryker att det är distinktionen som avgör om en observation görs. En cell skapar skillnader med det som den tar in och vad den vägrar att ta in. Eller så uppstår skillnaden genom att cellen tar in något på ett sätt eller på ett annat sätt som är olik det första. Luhmann framhåller att det gäller att se på distinktioner, eller på observationer, utan att det handlar om att uppfatta något med ögonen. Vi måste istället tänka oss att det handlar om att vi väljer att inte se något, att förneka något eller att ge något ett negativt värde. Det här är ett perspektiv på kommunikation som överensstämmer med den kritik av talhandlingar som vi tidigare såg att Suchman levererade, och ett argument mot Habermas krav att språket måste bygga på gemensam mening. Katherine Hayles har för övrigt utvecklat resonemanget om språk och koder i boken *My Mother was a Computer* (Hayles,2005). Där framhåller hon att språk och kod har växt samman och att det som skapat en skillnad mellan människor och maskiner har blivit otydligt. Hon menar att vi alltmer lever i en förmedlande värld som gör att gamla uppfattningar om språk, subjektivitet, litterära objekt och textualitet blir ifrågasatta.

Det finns luhmannianer som idag försöker föra in en större roll för individen i systemteorin med hjälp av Peirces semiotiska teori. Søren Brier menar att Luhmann är dålig på att redogöra för vad det innebär att observera, medan semiosis och abduktion hos Peirce är teoretiska begrepp som är knutna till hans tre grundläggande kategorier; förstahet, andrahet och tredjehet (Brier,2007:35). Brier framhåller att det är människan som skapar mening ur de tre kategorierna. Hos Luhmann har vi inte en människa med en fri vilja och inget som kontrollerar kommunikationer och handlingar. Det är istället kommunikationen som skapar de villkor som avgör vad som människorna kan observera (Brier,2007:36).

Brier redovisar så den kritik som kommit från bl.a. Habermas om att mening hos Luhmann skapas utan intentionalitet och bara som ett slags vetenskapligt registrerande. Brier menar att han håller med om denna kritik och hävdar att han inte förstår hur kommunikativa system i sig kan ha kapacitet att göra distinktioner. Människorna utgör bara en omgivning för det kommunikativa system som de bara kan påverkas av genom strukturell koppling och irritation (Brier,2007:40). Brier menar därför att vi måste hitta på något sätt att förkroppsliga kommunikationen. Det vill han göra genom att skapa en syntes mellan Peirce och Luhmann. Brier menar då att vi måste hitta ett sätt att göra observationer kommunicerbara med ett språk. Då krävs det: 1)en förkroppsligad observatör; 2)några objekt eller skillnader utanför en individs medvetande som kan refereras till eller representera; 3)ett

system av tecken för kommunikation med andra människor; 4) man måste acceptera att dessa människor är verkliga och 5) att de lever både i naturen samt 6) i ett samhälle och i en kultur (Brier, 2007:42-43).

Vi kan inte gå in på Peirce i detalj och jag får erkänna att jag tycker Peirce begreppsvärld känns föråldrad. Hans begrepp *Förstahet* som något som existerar oberoende av allt annat förefaller vara knutet till en realistisk ontologi. Utifrån Luhmanns relativism kan det inte finnas något sådant vilket gör att de distinktioner som skapas i social kommunikation blir mycket viktigare. *Andraheten* som varat reagerande på något annat tycks förutsätta att det finns ett aktivt subjekt, men samtidigt kunde det handla om de relativa positioner som i relation till varandra skapar en observation. *Tredjeheten* som förmedling mellan ett första och ett andra borde inte behöva gå via en kropp. Det kunde istället handla om just den förmedlande processen.

Viktigare är nog att vi observerar att abduktion¹⁷ som den process som skall ge ny kunskap, väldigt mycket påminner om Engeströms argumentation för expanderande processer. Vi bör hålla i minnet att frågan om stark AI just handlar om möjligheter att kunna skapa något nytt. Vi har visat på att sådana processer inte behöver bygga på en komplex funktionalitet i en aktör. Om vi hävdar att social kommunikation också bör handla om interaktionen mellan systemen och omvärlden bör nog en utveckling av Luhmanns teori ta fasta på detta hellre än att utvidga individens roll. En sådan utveckling bygger faktiskt på att artefakter som ett system av tecken blir till skillnader utanför en individs medvetande och att dessa skillnader kan finnas i naturen, i ett samhälle och i en kultur.

Om vi istället ser på språket som något som förmedlar med hjälp av texter finns det fog för att anta att språket kan vara relativt självständigt från författarnas intentioner. Det finns en annan ganska ny artikel som utifrån systemteorin också diskuterar omgivningens betydelse i form av det litterära systemet (Miles, 2006). Det intressanta med den artikeln är att den använder Wolfgang Iser's teorier om den litterära textens relativa självständighet från författarens intentioner, som utgångspunkt för att argumentera för att det litterära systemet fungerar som en kulturell omgivning med egen kapacitet att skapa nya kulturella kontexter. Den litterära texten blir på så vis en fiktion som inte är en exakt representation av vad som finns i världen men ändå något som säger mer om världens tillstånd än en massa fakta. Det är en diskussion som understryker vikten av det kulturella systemet och som ansluter till den diskussion vi redovisade i avsnittet om det sociala som text/artefakt.

¹⁷ abduktion, den benämning C.S. Peirce använde om den kreativa formuleringen av nya hypoteser vid förklaring av fakta som i initialskedet ter sig egendomliga (t.ex. det faktum att det finns fossila rester av fiskar långt in på land), men som sedan kan härledas deduktivt ur den nyformulerade hypotesen (att havet en gång gick så långt upp på land). Ur *Nationalencyklopedin* <http://www.ne.se.ludwig.lub.lu.se/artikel/107103>.

Det finns också en motsvarighet till debatten mellan Habermas och Luhmann i en debatt mellan Searle och den postmoderna filosofen Jacques Derrida som handlar om huruvida talhandlingar måste kopplas till direkt mänsklig handling. Den direkta debatten mellan dem började med Derridas angrepp i artikeln *Signature Event Context* i tidskriften *Glyph* (Derrida,1977) och Searles svar *Reiterating the Differences: A Reply to Derrida* (Searle,1977) samt Derridas långa replik *Limited Inc abc* (Derrida,1978). Derrida argumenterar för att det han kallar parasitära talhandlingar, bland annat olika former av texter, är väl så viktiga för kommunikation som mänskliga talhandlingar. Det vi bör framhålla är att Derrida genom detta faktiskt bryter loss från att kunskapen alltid måste vara en skapelse av människor och istället framhåller det han kallar "La difference". En position som liknar väldigt mycket den som Luhmann, Spencer Brown och den andra ordningens cybernetiker står för.

Detta ger oss också en anledning att knyta diskussionen om AI till den pågående debatten mellan sociologer och naturvetare om kunskapens status. Den debatten har kallats det nya vetenskapskriget eller Sokal-debatten efter fysikern Alan Sokal. Han var den som redan 1996 startade debatten genom att till tidskriften *Social Text* skicka in en artikel med titeln *Transgressing the Boundaries: Toward a Transformative Hermeneutics of Quantum Gravity* (Sokal,1996b). I den argumenterar han för att vetenskaplig objektivitet är en myt genom att referera till ett antal postmoderna filosofer och till konstruktivistiska kunskapssociologer. Lite senare avslöjade han i tidskriften *Lingua Franca* (Sokal,1996a) att artikeln var en bluff och en drift. Den kritik som berör sociologin är inriktad på den konstruktivistiska varianten, men eftersom Det Starka Programmet har svårt visa hur sociala processer blir påverkade av egenskaper i världen, träffar kritiken också SP. Däremot träffar den inte den relativism som finns i den luhmannska systemteorin eftersom den bygger på att egenskaper i omvärlden också påverkar de sociala processerna. Den uppfattningen bör därför innebära den möjlighet sociologin har för att försvara sig mot den rationalistiska naturvetenskapens argument i detta nya vetenskapskrig. Det vi då bör koncentrera oss på är hur vi kan omformulera kunskapssociologin i denna relativistiska och funktionella riktning. Kan vi visa på att intelligens finns i miljön får vi också belägg för att det är genom reflexiva processer som miljöns möjligheter att exponera nya relationer frigörs, och som därmed gör det möjligt för vetenskapen att befria sig från de dogmer som både sociologin och naturvetenskapen bär på.

Om vi återvänder till diskussionen om talhandlingar gentemot systemkommunikation framhåller Baecker att det är människorna som inför de begränsningar som hindrar kommunikation. (Vilket liknar Susan Leigh Stars position) Baecker framhåller bland annat att Bateson hävdar att det går att avgöra vad som är möjligt för människorna att göra, om man ber dem att berätta vilka begränsningar de själva och människorna i deras omgivning uppställer (Baecker,2001:61). Baecker framhåller att Luhmann istället försökte upptäcka de

människor som försöker korsa gränserna. Han framhåller att Luhmann tar sats i en uppfattning om att det alltid finns för få orsaker och samtidigt för få och för många effekter för att vi skall kunna tillskriva orsaker och effekter till olika fenomen (Baecker,2001:63). Baecker framhåller att detta leder till att den enhet som systemteorin talar om, inte är enhet hos en helhet utan en enhet som bildas av system som är ekologiskt länkade till varandra (Baecker,2001:64). Detta bör tolkas som att människan inte bara skall bedömas som ett system. Istället bör vi alltid bedöma människorna utifrån de system de ingår i. Intelligensen hos en människa bör sålunda inte bara bedömas av den intelligens hon kan uppvisa i sig själv. Istället måste hon bedömas enligt den kapacitet hon får genom de system hon är länkad till. Värdet med AI-redskap är att de skapar en funktionalitet som medger att människan kan länka till ett system som är mycket större till omfånget än vad hon själv kan hantera. Vi bör nog avvisa Habermas kritik av Luhmann med att identiska signifikanter bara kan finnas inom en given systemreferens. Och att processer som syftar till kunskapsexpandering först måste förändra denna systemreferens. Utanför den givna systemreferensen finns det inget givet vilket påvisar situationernas betydelse för kommunikationen. Baecker framhåller lite längre fram i sin text att det är just då vi börjar leta efter situationernas möjligheter som vi börjar bete oss som sociologer (Baecker,2001:66).

Flera tänkare har också uppmärksammat att det kan finnas en möjlighet att göra en syntes mellan de bägge varianter av kommunikationsteori som Habermas och Luhmann har levererat. Ett exempel på en artikel som tar upp möjligheterna av en sådan syntes är skriven av den tyske sociologen Wolfgang Schneider med titeln *Die Komplementarität von Sprechakttheorie und Systemtheoretischer Kommunikationstheorie*. Det är för övrigt inte så märkligt att syntesmöjligheter börjar upptäckas eftersom Luhmann och Habermas länge och med ömsesidig respekt diskuterat med varandra och justerat sina teorier för att de bättre skall klara sig mot den kritik de bägge framlagt mot varandra. För vår del är förstås det viktigaste om språket kan fungera som en självständig artefakt. I så fall finns det en möjlighet att medla mellan Luhmann och Habermas. Den frågan har nog att göra med om kommunikation sker under villkor som karaktäriseras av Luhmanns relativism eller Habermas intersubjektiva realism.

Av de moderna Luhmannefterföljarna är det kanske Colin Grant som mest pekar på betydelsen av att människans kapacitet att värdera måste bedömas utifrån att kunskap skapas i en värld som är relativ. Grant för diskussionen ett snäpp längre genom att tala om att kommunikation karaktäriseras av porositet, snarare än av gemensamma överenskommelser.

Grant argumenterar för sin relativistiska position genom att hävda att interaktionsbegreppet kan omformuleras med en annan teoretisk modell. En modell som inte behöver grundas i en realistisk filosofi, i en rationell kommunikationsteori eller grundas med hjälp av teorier om intersubjektivitet. Denna alternativa modell kan man komma fram till genom att se att samhället

hänger samman genom fiktionernas funktionalitet. Han menar, med hänvisning till Siegfried Schmidt, att sociala fiktioner också fungerar som pragmatiska fiktioner genom att använda rekursivt länkade kommunikationer. Dessa bygger stabila sociala ordningar genom att kulturen blir en socialt nödvändig semantisk avbildning av modeller av världen (Grant,2003:100).

Argumentationen för porositet kopplar Grant till den skillnad mellan entropi och redundans som Shannon och Weaver arbetar med. Grant uppmärksammar att desto mer organiserad kommunikationen är, desto lägre blir entropin. Ur det synsättet innebär mening tillförlitlighet och är därför nära förknippat med redundans i kommunikationen. Grant menar istället att vaghet i social kommunikation kan användas för att utveckla en teori om komplex kommunikation (Grant,2003:101). Vi har också redan flera gånger i denna avhandling sett att osäkerhet snarast är en förutsättning för att ny kunskap skall kunna skapas.

Det är för att utveckla denna idé som Grant försöker ersätta begreppet entropi med begreppet porositet. Begreppet porositet kommer från geofysik och biologi. I geofysik betyder ett poröst medium en volym som delvis består av ett solitt skikt och delvis av ett flyktigt utrymme. Det går inte direkt att använda denna definition för att diskutera social kommunikation. Men vi bör enligt Grant observera att i ett poröst medium interagerar det solida och det flyktiga (Grant,2003:103). I kommunikationsteoretiska termer menar Grant att begreppet betecknar ett stadium som är öppet och dynamiskt mot omgivningen, vilken gör den operationella stängningen av systemen mycket instabil. Han ger några exempel på en sådan porositet: Intertextualitet, sammanblandningen mellan privat och allmän diskussion, hacking och virtuell kommunikation...(Grant,2003:104)

Grant visar, med stöd av filosofen Hilary Putnam, att referensfrågan måste bedömas utifrån att kommunikation sker under porositet. Vi kan därför inte tala om liktydiga referenser. Därför innebär realism ett omöjligt försök att se världen från ingenstans (Grant,2003:106-107). Verkligheten är därför lika instabil som kommunikationer.

Också Grant tar upp situationernas betydelse. Det gör han genom att hävda att Wittgensteins språkregler inte konstrueras med hänvisning till en extern kunskap utan de följer relationella kopplingar mellan talarna (Grant,2003:111). Detta använder Grant för att ifrågasätta Habermas uppfattning om att det går att undvika entropi genom en rationell kommunikation som tar allt med i beräkningen. Grant frågar hur entropi kan reduceras eller stabiliseras som funktionell kommunikation utan att förneka frihet eller utan att patologisera synbar abnormal diskussion som parasitär? (Grant,2003:113) Grant menar att gränsen mellan det faktiska och det som Habermas kallar kontrafaktiskt måste suddas ut eftersom bägge är observatörsberoende (Grant,2003:114).

Också Leydesdorff formulerar om Luhmanns modell så att den mer bygger på att se på den osäkerhet som koderna representerar istället för att se dem som binära

koder (Leydesdorff,2001:64ff). Det gör naturligtvis att betoningen flyttas bort från de psykiska systemens tankeverksamhet till själva kommunikationsprocesserna i de sociala systemen. Det är svårt att ge Leydesdorffs framställning rättvisa på några rader. Men framställningen bygger på vad denna osäkerhet betyder för observatörsproblematiken. Vi har redan sett att observatörsbegreppet arbetar med en tredje ordningens observatör. Det handlar om att fler observerar samma fenomen. Exemplet är när massmedia diskuterar fenomenen. Leydesdorff kallar denna form av reflexivitet för hyperreflexivitet. Aktörerna ges tillfälle till en utvidgad reflexivitet genom att de reflekterar över sådant som andra redan reflekterat över. Han räknar också med att det finns ett fjärde ordningens observationsbehov vilket han kallar ett behov av infrareflexivitet. Leydesdorff framhåller att det är svårt att analytiskt beskriva denna fjärde dimension (Leydesdorff,2001:66).

Leydesdorffs text erbjuder ingen konkret beskrivning av vad infrareflexivitet är för något. Det blir inte klart vilken av de tre formerna av självreferens han hänför infrareflexiviteten till. Eftersom han talar om reflexivitet bör han avse den andra formen. I enlighet med vår diskussion bör istället problemen kring social intelligens handla om den tredje formen. Det problemet handlar om att vi människor inte kan se alla möjliga kombinationer därför att vårt medvetandes reflexivitetsförmåga är begränsat. Det klarar inte av att se sådant som inte kan kopplas till något som vi redan känner till, eftersom medvetandet alltid befinner sig inom en systemreferens. Vi kan därför inte fullt ut ta vara på alla förutsättningar för en fjärde ordningens observation. Det förklarar varför vi behöver social kommunikation, men det är fortfarande ett problem att förklara hur denna sociala kommunikation fungerar, eftersom den inte kan bygga på individernas medvetanden.

Leydesdorff diskuterar om det går att skapa datorprogram som kan hantera den fjärde ordningens problematik. Om det är möjligt skulle det vara möjligt att tala om en stark AI. Det problem som han ser är att ett fjärde ordningens system bara kan analysera en ordning under sig. Ett fjärde ordningens system, dit Leydesdorff räknar människorna, kan bara se den tredje ordningens problematik. Det kritiska för den starka AI:n blir då om vi kan konstruera dem som femte ordningens system, som är kapabla att göra fjärde ordningens observationer (Leydesdorff,2001:113ff).¹⁸

Leydesdorffs diskussion av möjligheterna för ett femte ordningens system är väldigt formaliserad. Han menar att det blir svårt att bygga datorer med en högre dimension eftersom det inte går att bygga datorer i en dimension högre än vad människor förmår eftersom det måste vara människor som konstruerar dem (Leydesdorff,2001:115). Leydesdorff ger inget slutgiltigt svar men vi kan hävda att

¹⁸ Frågan om stark AI handlar då paradoxalt nog om huruvida vi kan konstruera system som är mer kapabla än människans medvetande, min anm.

det visar på att AI-systemen i så fall på egen hand måste kunna konstruera denna dimension.

Det framstår som att Leydesdorff här bara diskuterar vilken roll det artificiella har för själva funktionaliteten inom de autopoietiska systemen. Då kan systemen inte expandera kunskap utan stannar vid att försöka vidmakthålla den kunskap som redan skapats. När vi diskuterar frågan utifrån de förutsättningar som elektronisk kommunikation och elektronisk publicering ger för att kommunicera med andra system och med en mer omfattande omvärld, bör vi också tänka oss att AI-redskapen har en funktion som handlar om att vi kan nå denna vidare värld. Det är svårt att säga om detta underlättar förmågan till femte ordningens system, eller om det helt enkelt bara gör så att omvärldens förmåga att exponera komplexitet blir större. Vi bör inse att det här handlar om en funktionalitet som om den fanns tidigare, åtminstone blivit mycket effektivare genom AI-tekniken. Det bör vara rimligt att den sociala kommunikationen lika väl bör innehålla en horisontell dimension som en vertikal dimension. Vi bör uppmärksamma att det är en dimension som pekar mot att kommunikation sker mellan system och omvärld istället för inom systemen som den horisontella dimensionen mer är knuten till.

Frågan bör handla om AI-systemen kan omfatta Luhmanns tredje form av självreferens. Den som han kallar reflexion. Leydesdorff tycks inte använda den skillnad som finns mellan reflexivitet och reflexion för sin analys. Det som han diskuterar är reflexivitet som har att göra med de sociala systemens kommunikation inom den egna systemreferensen. Det som bör vara viktigare är den interaktion som kan utnyttja differensen mellan system och omvärld. Leydesdorffs analys tar också sikte på diskussionen kring den symboliska AIns förmåga. I Luhmanns diskussion kring referens handlar självreferens och reflexivitet fortfarande om systemens förmåga att reflektera över sig själva och sina villkor. Det blir rimligt att det är psykologiska redskap som används i den kommunikationen. Det är först med reflexionen som vi talar om de sociala systemens förmåga att reflektera över gränsen mellan systemen och deras omvärld. Det är rimligt att en sådan relation behöver bli mer materialiserad.

När vi diskuterar de förutsättningar som ges av materialiserade redskap, i.e. elektronisk publicering och elektronisk kommunikationen, handlar det om en horisontell dimension. Det som är centralt i reflexion är att kunna reflektera över skillnaden mellan system och miljö. Därför bör AI ge större möjligheter för att denna skillnad skall kunna komma till uttryck. Återigen vet vi inte om AI-systemen kan ha den funktionaliteten att de lyfter fram denna skillnad. Dock bör vi inse att vi inte ens vet hur människan kan reflektera över denna skillnad.

Det går att tolka Luhmann som att han menar att någon form av kommunikation uppstår om villkoren för någon av de tre formerna av självreferens är uppfyllda. Han tycks emellertid mena att den relation mellan systemen och omvärlden som förmår förändra systemen, bara uppstår om förutsättningar för reflexion finns. I enlighet med Leydesdorffs argumentation kan vi vara tveksamma

till om konnektionistisk AI kan uppfylla villkoren för reflexivitet. Däremot bör den horisontella rationaliseringen innebära att möjligheterna för reflexion ökar. Den process som inte psykiska system kan hantera. Om det handlar om social kommunikation kanske inte är den viktigaste frågan. För min del har jag lättare att förstå Luhmanns uppfattning om att kommunikation har att göra med villkoren för att kunna observera något nytt, om det är så att det handlar om en relation mellan system och omvärld. Om den sociala kommunikationen bara kan ske inom ett socialt system som är operativt stängt är det svårt att förstå hur någon kommunikation kan uppstå. Den kommunikation som sker inom en sådan given systemreferens bör inte kunna observera något nytt. Det kan vara så att interaktioner inom sociala system är funktionellt stängda men det blir ändå svårt att se hur de kan vara kognitivt öppna om inte detta baserar sig på en öppenhet hos de artefakter som bör vara de kognitiva förmedlarna. Det är möjligt att vi istället skall se på artefakterna som något som står utanför systemen. Vi har uppmärksammat att det är lite svårt att förstå tingens placering i Luhmanns teori. Hör de till systemen eller till omgivningen. Om de hör till omgivningen och om det är på dem som kommunikationens funktionalitet bygger så bör Luhmanns kommunikationsbegrepp framför allt handla om relationen mellan systemen och deras omvärld. Det är först när omvärlden genom systemens kognitiva öppenhet lyckas irritera systemen så mycket att de ser något nytt som en verklig kommunikation kan uppstå. I den processen bör artefakterna stå för funktionaliteten.

Vi bör inse att ett fullständigt krav på stark AI är en fråga om att alla tre nivåer av självreferens skall fungera. Vi bör då inse att inte ens människor kan hantera alla tre nivåer. Frågan om stark AI blir då en fråga om huruvida datorerna skall klara av mer än människan. Ett fullt tillräckligt krav på att de skall kunna kallas intelligenta är att de bättre än människan kan hantera den nivå som handlar om reflexion.

Andrea Pronzini har i artikeln *First-Order Semantics and Artificial Intelligence* visat på att diskussionen om AI ofta fastnar i antaganden som hon kallar den första ordningens semantik. Pronzini visar på att studierna av kognition som sker med hjälp av datorer är bunden av de strukturella begränsningar som finns i denna semantik. Hon kritiserar några element som AIs kognitiva och konnektionistiska paradigm delar med den forskning kring distribuerad artificiell intelligens som använder sig av multiagent modeller. Hon menar att den strategin innebär att man inte till fullo kan inse varför det är meningsfullt att koppla ihop datorn med intelligens. Den samhällsvetenskapliga systemteorin erbjuder en bättre utgångspunkt för att definiera intelligens och framför allt för att definiera datorns roll. Orsaken till det är att den inte talar om datorn som något som skall jämföras med psykiska system utan som något som är ett medium för kommunikation (Pronzini, 2002:2).

Pronzini framhåller att den första ordningens semantik har att göra med de möjligheter som öppnades av skrivandets medium. Dessutom dubblar skrivandet kommunikationens funktioner och möjliggör den andra ordningens observationer. Detta innebär att en skriven text kan observeras och jämföras med andra texter. Hon framhåller att diskussionen kring AI är fast i den europeiska traditionen av att diskutera kognition som något som tillhör den första ordningens observation, fastän det moderna samhället karaktäriseras av den operationella normaliseringen av den andra ordningens operationer (Pronzini,2002:3).

Pronzini framhåller till exempel att den fråga som handlar om huruvida datorer kan tänka, som AI-forskningen har varit så upptagen med, har att göra med den första ordningens cybernetik (Pronzini,2002:5).

Hon framhåller också att det för tillfället är omöjligt att förutse allt vad denna nya syn på intelligens innebär. Hon menar att den enda möjliga typ av rationalitet för det moderna samhället och dess olika system ligger i hur man sköter och exploaterar skillnader. Det är för att utnyttja den potentialen, som datorn kan öppna de vägar till kommunikativa hjälpmedel som behövs för att vi skall kunna närma oss den rationalitet det moderna samhället behöver (Pronzini,2002:16).

I detta kapitel har vi argumenterat för att situationerna måste ses som en helhet som skapar förutsättningar för att observera något nytt. Vi har också argumenterat för att språket bör ses som en artefakt som skapar olika situationer och att talhandlingar är beroende av vad dessa situationer erbjuder för perspektiv. På så sätt har vi visat att den luhmannska systemteorin snarast är en förutsättning för talhandlingsteorin.

Vi har också visat att betonandet av situationernas betydelse och betydelsen av Luhmanns uppfattning av att de sociala systemen måste kunna interagera med sin omgivning för att kunna skapa ny kunskap, gör att vi kan se att sociala processer inte behöver stå i motsättning till att naturvetenskapliga fakta påverkar kunskapen trots att vi framhåller att det alltid är sociala processer som avgör vad som kommer att uppfattas som etablerad kunskap. Vi har också visat på att det mest väsentliga för att sådana processer skall kunna uppstå är att förutsättningar för reflexion finns. Eftersom reflexion handlar om systemens relationer till omgivningen finns det anledning att hävda att sådana processer underlättas av AI system. På så sätt finns det också fog för att hävda att AI verkligen kan vara stark.

10. Uppsummering av argumentationen

Vi har uppmärksammat att frågan om stark AI har mycket att göra med att kritiker och försvarare menar att kunskap handlar om att göra något av det icke-formaliserade. Kritiken handlar om att stark AI inte fungerar därför att världen inte är lagbunden. Mina egna argument för stark DAI handlar om att det redskapet bättre än människor kan se möjligheterna att hitta ny mening i en relativistisk värld. Vi skall därför avsluta med en uppsummering av argumenten i form av ett index i 10 punkter. Indexet är inspirerat av det index som finns i Stephan Fuchs bok *Against Essentialism*, som är den bästa argumentation för relativism jag har läst (Fuchs,2001:337-339).

1. Frågan om artificiell intelligens är en fråga som mer handlar om hur människan använder sig av den intelligens som finns nerlagd i den omgivning eller den kultur som hon lever i, än en fråga som handlar om egenskaper i människans biologiska natur.

2. Frågan om huruvida stark AI kan skapa kunskap på egen hand, i.e. om artificiellt intelligenta AI-system kan skapa kunskap oberoende av mänskliga handlingar, är en fråga som sociologin som en samhällsvetenskap bör ha lätt att bejaka. De vetenskaper som sysslar med den mänskliga hjärnans biologiska konstruktion eller hennes psykologiska tillstånd, i.e. neurovetenskapen eller den psykologiska vetenskapen bör däremot ha svårare att bejaka frågan. Ändå är det sociologin som har förhållit sig mest kritisk till möjligheterna av stark AI.

3. Intelligens handlar inte om problemlösning utan om processer som är inblandade i en kreativ process. Kreativa processer sker genom kommunikation i sociala system och genom interaktion mellan dessa system och deras omgivning, och inte genom att enskilda individer tänker med sina hjärnor. Att människor kan bete sig kreativt har att göra med att de kan använda sig av de redskap som skapats under den mänskliga kulturens utveckling. Den intelligenta människan är därför

en så kallat posthumanistisk människa som kan använda den mänskliga kulturens komplexitet för att starta kreativa processer.

4. Så kallade neurala dataprogram eller konnektionistiska datamodeller utgör framför allt modeller av sociala processer i samhället och skall inte i första hand uppfattas som modeller av hur mänskliga hjärnor fungerar. Den funktionalitet som behöver finnas i dessa datamodeller, skall därför jämföras med den funktionalitet som finns i samhällsprocesser, och inte med den komplexa funktionalitet som karakteriserar den mänskliga hjärnan. Den debatt kring AI som handlar om jämförelser med människans sätt att förstå, bör därför inte vara utgångspunkten för att värdera möjligheterna av stark AI.

5. Ett förhållande som gör att människan inte kan skapa ny kunskap bara genom att använda sitt eget tänkande, är att kunskap är en storhet som antingen inte alls har något givet värde, eller ett värde som åtminstone inte är direkt åtkomligt för människans medvetande. Därmed kan människan inte skapa ny kunskap genom att rationellt bearbeta den kunskap som hon exponeras för. Det är samtidigt ett förhållande som förklarar varför sociala processer i stor utsträckning bestämmer vilken kunskap som kan skapas. Dessa sociala processer fungerar genom en kommunikation där funktionaliteten beror på den hänvisningsrikedom som finns i de kulturella artefakterna eller i omgivningen. Den funktionaliteten kan bygga på artificiella processer eftersom social kommunikation sker utanför människornas hjärnor. Det finns anledning att tro att det finns en tröghet i denna kommunikation som beror på praktiska svårigheter. Vi har talat om bibliotekens förmedling och att den har haft begränsningar tack vare att den gamla tekniken varit trög. Om vi talar om akademiska texter har denna tröghet berott på att de varit bundna till tryckta medier. Artefakter som exponeras i elektronisk form är lättare att nå och kan behandlas av AI-program. Det finns därför anledning att tro att stark AI innebär en naturlig utveckling och förbättring av den kommunikativa funktionalitet som hör samman med kreativa processer. Det är således inget som naturligt behöver stå i motsättning till en rätt användning av begreppet intelligens.

6. Sociologisk teori bör konstrueras som en teori som förklarar hur samhällsprocesser fungerar utan att de behöver grundas på mänskliga handlingar. Stark AI, som betonar funktionaliteten i en kommunikation som sker mellan kulturella artefakter och mellan dessa artefakter och omgivningen, kan förklara hur sociala processer påverkar kunskapsskapande utan att de behöver reduceras till biologiska egenskaper i den mänskliga naturen. Den starka AI:n kan därför användas som en modell för hur sociologin kan konstrueras som en makro-teori samtidigt som den kan förklara hur mikro-processer kan fungera utan att de behöver grundas på mänskligt handlande.

7. Det sociala bör ses som en form av situationer som skapas av de sociala systemens kommunikation, både inom sig själva och med den miljö som omger dem. Det sociala är därmed något som är förknippat med det som människan inte kan förstå. Eftersom en viktig egenskap hos intelligens är en förmåga att kunna gå utöver det som människan inte kan tänka sig fram till, kan intelligens handla om en social process. Det som vi tror är resultatet av tänkande är ofta resultat av förutsättningar som finns i den situation som kommunikationen har lett till. Det gör att artificiell intelligens bör vara förenlig med en naturlig social process.

8. När vi skall värdera möjligheterna av stark AI skall vi framför allt diskutera om de kan bidra till den reflexion som utgör grunden för systemens interaktion med omvärlden. Också människor som interagerar med varandra använder en reflexiv funktionalitet som är mycket enklare än den som är inblandad när de tänker med sina hjärnor.

9. Eftersom intelligens, som en förmåga att skapa något nytt, handlar om reflexiva processer mellan sociala system och deras miljö, består intelligenta processer inte bara om kommunikation inom sociala system. De skapas också i den reflexiva process där verkligheten deltar i form av den natur som också är en del av situationerna. Diskussionen om stark AI visar på en möjlighet att förklara hur sociala processer även är inblandade i de kunskapsprocesser som bygger på en tolkning av naturvetenskapliga fenomen.

10. Frågan om AI-systemen kan skapa den funktionalitet som behövs för intelligens handlar om huruvida de kan ha all den funktionalitet som finns i tre olika former av reflexivitet; självreferens, reflexivitet och reflexion. Vi bör då inse att inte ens människor kan klara av alla tre nivåer. En svaghet i människans förmåga att skapa kunskap är att hon har svårt att observera omvärlden eftersom den relationen handlar om processer som hör till system/omvärld relationen. Det innebär att ett fullständigt krav på starka AI-system skulle innebära att de måste klara av mer än vad människor kan. Ett fullt tillräckligt krav på att de skall kunna kallas intelligenta är att de bättre än människan kan hantera den nivå som handlar om reflexion. Det finns anledning tro att DAI klarar av den processen bättre än människor.¹⁹

¹⁹ Fuchs eget index är i 25 punkter:

1. In the beginning, there were networks. Networks are fields of forces. They do not consist of nodes. Nodes are outcomes of networks. Nodes without relationships are nodes without qualities.

2. Society occurs in encounters, groups, organizations, and networks. Networks link these, and themselves, into larger networks. Nothing controls society or even holds it together. It

has no "order," or only that temporary and fragile order which it makes and remakes for itself.

3. The boundaries of networks shift together with their own velocities and those of related networks.

4. Society is a massively plural and parallel occurrence or happening. There is no master- or mega-encounter that could regulate all the other encounters.

5. Everything that happens happens locally, at a certain place and time. All that exists exists empirically, and only until further notice.

6. Transcendence and universality are innerworldly, temporary, and variable outcomes of expanding networks. These outcomes are improbable and reversible. They require the painstaking work of network.

7. A core rule of method is to allow for variation. There are no natural kinds, essential properties, or things-in-themselves. But a network can condense and converge into kinds and properties that appear natural and essential to it.

8. Natural kinds and stable objects appear when an increasingly self-similar network hums to itself. They both take time.

9. Explaining is not action, but interaction at a distance. As the distance closes, explanation turns into understanding and finally into love.

10. There is no logical relationship between scientific explanation and technological control. Science augments complexity; engineering cages it, but only for the time being. Sometimes, complexity bursts out of its cage.

11. That science is instrumental reason is an invention of philosophy, not science.

12. The unity of science extends only as far as actual mergers and unions between sciences. There is no "science as such." No scientist does "science in general."

13. The subject or actor is part of the explanandum, not the explanans, of sociological science. Agency is a variable attribute to personhood, not an essential fact about natural persons. Consciousness is overrated.

14. Persons do not act, much less act rationally. "Action" is how some observers make sense of some events. Personhood is an institution.

15. Free will is the residual from failed attempts at explanation and domestication.

16. What persons say is not the cause, but the outcome, of communication. Nothing communicates like communication.

17. A culture or science has nothing to work with but the results of its previous operations. No science or culture goes to work on "reality" or "nature."

18. What makes science special is not truth, but time out from ideological suspicion.

19. Constructivism is adequate only if it can explain why some constructions appear unconstructed. Otherwise it is an arrogance and imposition.

20. Sociology is not about social facts, but about how and when such facts emerge, and when not. All facts are constructs, but some are weaker, others stronger. All cultures are constructed, but only some of them are constructivist.

11. English Summary

This dissertation is concerned with Sociology's stance in the debate on Artificial Intelligence – AI. There is a need for sociologists to realize the difference between two approaches to constructing AI systems: Symbolic AI (or Classic AI) and Distributed AI – DAI. Sociological literature shows a largely critical attitude towards Symbolic AI, an attitude that we think is justified. Symbolic AI occupies a rough cognitive position based on ontological realism.

The main theme of this dissertation is that DAI is not only compatible with Sociology's approach to what is social, but also constitutes an apt model of how a social system functions. This is consolidated with help from Luhmann's social systems theory, modern Luhmannians such as Dirk Beacker, Loet Leydesdorff and Colin Grant and Luhmannian-inspired German research on the relation between social systems and technology. This research is called *Socionics* (Sozionik).

Chapter 1 describes discussions on AI in sociological literature and shows that the attitude is mainly negative, but that there are some sociologists who do have a positive attitude to AI. Some methodological problems arise when we analyze a phenomenon that is far beyond what is commonly thought possible. Humans have difficulties understanding a phenomenon that cannot be judged through the human mind, which can only observe what is comprehended through the functions of the same human brain. This phenomenon is referred to as the Frame Problem. The chapter also points out that the sociological argument for DAI has to be connected to sociology's comprehension of what is called social intelligence.

21. A culture is not in the mind. It does not consist of ideas, although ideas summarize and simplify some cultural results.

22. Realism and relativism are not opposites, but are linked by a continuum of cultural closure and settlement.

23. The limits of an observer are where he stops (for now) asking further questions. An observer is always a cultured observer, or no observer at all.

24. Everything is what it is in relation to what it is not, not yet, or not anymore.

25. A culture or science cares - for itself. It is interested - in itself. It values - itself. A disinterested observer is careless.

As AI critical sociologists find it rather difficult to accept the suggestion that AI may be a model for social processes, Chapter 2 provides reasons why AI can be incorporated into modern sociology. It starts with a discussion of possible definitions of knowledge, information and intelligence, and the implications these have for social scientists who have considered the functions of artefacts and nonhuman agents in social processes. The writings of Karin Knorr-Cetina and Bruno Latour can be turned around and translated into arguments for AI. Vygotsky-oriented education scientists claim that processes leading to new knowledge are about expansion and not about problem solving. Further, AI is not about robots thinking in the same way as humans, but about holistic social processes where the actor does not need to have a complex functionality. This provides a good opportunity to explain what sociologists call emergency, i.e. how social processes shape new knowledge independent of human actors. Given that the Luhmannian opinion on functions of social systems is about the capacity to communicate, this chapter takes up the question of what social communication is. It ends with a discussion about how complex processes may be modelled in a social process.

Chapter 3 begins with a well-known AI test – the Turing Test, and various criticisms of it, followed by a connectionist programming method on which DAI strategies are built. The chapter next describes one of the most important arguments against AI – the Chinese Room Experiment by philosopher John Searle. However, this only holds against symbolic AI, and Hubert Dreyfus, one of the opponents of symbolic AI, suggests that the argument against Symbolic AI does not hold against connectionist AI models. The anthropologist, Lucy Suchman, claims that shaping new knowledge is not about capacities for rational actions. More important are situations that lead to opportunities to shape new knowledge, and what we believe is the result of thinking, i.e. purposeful action, is in fact dependent on the situation in which it is produced.

Chapter 4 makes a case for *Social* not being compatible with strong AI. We start with a discussion of strong AI by the sociology professor, Randall Collins, who takes a positive stance against the possibility of strong AI and claims that the critical point is whether AI can handle emotions and construct conversations with new and creative ideas. Our own contention, based on Luhmann's theory, is that emotions have a tendency to strengthen existing knowledge. Therefore, the question is more one of opportunities to see new possibilities without being blinded by emotions. This chapter describes the AI discussion, from its origins up to the Socionics research program. It also takes up a critical assessment of Dreyfus' position vis-à-vis connectionist AI by sociologist Harry Collins. The latter's position may be summarized thus; shaping new knowledge is dependent on the social capacity to do something about the circumstance that the world is chaotic and difficult to understand. Collins' belief is that computers can only process knowledge in an ordered world. Nonetheless, humans also have difficulties coping

with an unordered world. The reason why Collins thinks that computers are so capable is that much more is ordered than we think. This dissertation takes an opposite view, which is that computers in a DAI architecture have, in some respects, a strong capacity to handle complexity that belongs to the environment because they can process information from a wider area. Because the criticism of strong AI is based on assumptions of human nature, the discussion is connected to Darwinian evolution theory. Strong AI is a question of the development of human cultural evolution. Humans have always used technological artefacts to increase their capacity to shape human culture.

Chapter 5 shows how *Second Order Cybernetics* has developed from a mechanistic theory into one that tries to explain how the complex and relativistic world can best be modelled by cybernetics. Besides, the posthuman being, who has to use several types of artefacts and tools to produce culture and knowledge, does not only refer to the human who uses computer technology. Amongst others, Katherine Hayles shows that humans have always been posthuman. An example is Gregory Bateson's famous question: "Suppose I am a blind man, and I use a stick. Where do I start? Is my mental system bounded at the handle of the stick? Is it bounded by my skin? Does it start halfway up the stick? Does it start at the tip of the stick"? Bateson's answer is that these are nonsense questions. "The stick is a pathway along which transforms of difference are being transmitted. The way to delineate the system is to draw the limiting line in such a way that you do not cut any of these pathways in ways which leave things inexplicable. If what you are trying to explain is a given piece of behaviour, such as the locomotion of the blind man, then, for this purpose, you will need the street, the stick, the man; the street, the stick, and so on, round and round". This shows that humans are dependent on their environment and that knowledge is not only something that has to do with their brain, but also the rest of their bodies and their environments. The chapter continues with a summarized introduction to certain important concepts in Luhmannian theory that are relevant to the discussion of AI. Luhmann's treatment of meaning is described and the Luhmannian form of functionalism is shown to have a construction that gives good reason to state that it is a radical theory of society, in contrast to older forms of sociological functionalism. The important factor in this form of functionalism is differentiation and not mechanisation.

Chapter 6 explains how social processes may depend on artefacts and artificiality, starting with an analysis of Vygotsky-oriented theories. The activity theoretician, Yrjö Engeström, maintains that problem solving only leads to elaborating existing knowledge. Processes that shape new knowledge are dependent on expansion, and human beings are dependent on artefacts to be able to shape conditions that make expansion possible. The possibility of AI has to be evaluated in terms of human beings' capacities to cope with reflexive processes. Luhmann teaches us that we have to see the difference between three different forms of reflexivity: self-reference, reflexivity and reflection. We contend that, in order to be

able to shape new knowledge in expanding processes, there must be circumstances that allow reflection. This is strengthened by the assertion that communication is not something that people do. It is only social systems that communicate. The chapter continues with a discussion of this way of defining communication and ends with another Vygotsky-inspired theoretician, James Wertsch, teaching us how connectionist systems fit into Vygotskian theories of artefacts.

Chapter 7 discusses two arguments for and against AI from methodological individualist oriented sociologists. Johan Asplund argues against AI on the grounds that although thinking is a social process, it has to depend on the last node that must be a human brain. Our belief is that Asplund does not understand the real implications of his view that problem solving is a social process. It is not a process where several people communicating with each other solve problems and shape new knowledge. What Asplund sees as problem solving is instead something that is dependent on communication in the social system. Because the social system interacts with its environment, it gives new opportunities to make observations, which provides individuals in the system with new possibilities of expanding their knowledge. Several people together will have difficulties in solving problems if they are unable to observe new possibilities. Susan Leigh Star gives a different perspective; that we are always dependent on technology and that new technology brings us chances to break up conventions and old structures that make it difficult to shape conditions to fit new observations. Criticizing the fact that we judge intelligence in terms of what a single human is capable of, she proposes that we replace the Turing Test with a Durkheim Test, because we have to test intelligence according to how it may be used for society's opportunities to change and adapt to new circumstances.

Arguments for the possibility of a strong AI rest on sociological holism, as shown in Chapter 8 by a critical discussion of one of the best known holistic theories of the growth of knowledge – the *Strong Program in Sociology of Knowledge* by David Bloor. The chapter takes a stance with regard to the opinions of philosophers who criticize Bloor for not being able to show how knowledge may be shaped without being connected to reasoning processes originating in human thinking. The Luhmannian form of defining social communication could give adherents of the *Strong Program* an opportunity to defend the capacity of social factors to function without being immediately connected to human thinking. Hence, these sociologists, who often base their arguments on a critical attitude to technology, will have to see that technology may also have a deliberating function. Sociologists should not explain social processes as causal processes, since this gives the theories a conservative character. Only if they see that the world is dependent on functional and ever changing processes, is it possible to claim that a social environment can shape new knowledge independent of the thinking processes of its humans. The chapter ends by looking at the difference between coding and programming in Luhmannian theory. Luhmann's rather static attitude to coding

represents a weakness in the theory. This is not elaborated further here, but is taken up in the next chapter showing that a more open view of coding is advocated by several prominent Luhmannian students and followers. This part of Luhmann's theory may be the reason why many still think he was a conservative sociologist. From a connectionist point of view, it is ever changing codes that make up conditions for changing the programs.

Chapter 9 tries to synthesize the earlier discussions. It starts by enlarging the argument for a relativistic ontology of the social processes that lead to new knowledge, and continues with theories that allow us to broaden the concept of the situation to incorporate the whole environment. Sociology's use of the term situation has always meant more than can be contained in a special and concrete situation. This argument is developed in a discussion of some aspects that Johan Asplund has drawn attention to in an analysis of Kenneth Burke's book *A Grammar of Motives*. According to Asplund, Burke means that motives have to do with five questions; what, when and where, who, how and why. Asplund's position is that we cannot understand a motive only by asking one of the questions. We have to take them all into account without holding any as more important than the others. This is why we have a pentad. In contrast to Asplund, we show that this may be used to suggest that situations, because of their distributed character, are independent of human actions. The environmentalist perspective is further developed by connecting it to theories that explain how the concept of the posthuman being may be expanded by showing how computer programs help humans to increase their capacities. We also get back to Darwinism and discuss whether it implies that humans cannot change their capacities to use the new technologies before their biological capacities have changed over thousands of years. The theory of cultural evolution implies that humans do not have to wait until their biological genes have adapted to the new cultural environment. It is easier to justify this stance if we abandon the realistic ontology of Richard Dawkins and replace it with Luhmann's relativism.

The argument for relativism is strengthened by showing that sociology can defend its position in what is called the New Science War, starting with the accusation from Alan Sokal that sociological constructivism is a fraud, if we can incorporate into social processes the relation between social systems and the environment - the process that Luhmann talks about as reflection. We point out that this should imply that Luhmann's concept of communication also holds for relations between systems and their environments.

We therefore take the discussion on Luhmann's communication theory a bit further. First, we compare Habermas' view on communicative action with Luhmann's view on social communication and find that it is possible to see a convergence of Habermas' and Luhmann's theories.

Second, we present Colin Grant's reasons for why we should reformulate Luhmann's static view on coding and the concept of communication in a direction

that gives opportunities to also include interactions between the system and the environment in what may be seen as communication. We have stated that this opportunity is what is most important for the possibilities of making a case for the capacity of DAI systems to shape new knowledge on their own. Still, the most important question is not whether this is part of communication. What is most important is that reflection between the system and the environment is significant for knowledge-producing processes. Strong AI has more capacity to handle these processes than humans have, while the strongest argument for strong DAI is that it can reach a far wider area than human beings' brains can. This capacity for reflection is a better way of judging the capacity of strong AI, than questions about causal capacity and capacity for action.

Referenslista

- Anderson, Bo (1989) *On artificial intelligence and theory construction in sociology*. I: *Journal of Mathematical Sociology*, Vol. 14, ss. 209-216.
- Anderson, Michael L. (2003) *Embodied Cognition: A field guide*, I: *Artificial Intelligence* 149, ss. 91-130.
- Asplund, Johan (2002a) *Avhandlingens språkdräkt*, Göteborg, Korpen, 137 s.
- Asplund, Johan (2002b) *Genom huvudet*, Göteborg, Korpen, 163 s.
- Asplund, Johan (1980) *Noter till Kenneth Burkes motivationsgrammatik*, I: Asplund, Johan, *Socialpsykologiska studier*, Almqvist & Wiksell Förlag AB, Stockholm, 200 s., ss. 127-169.
- [Bacon] *The Works of Francis Bacon* [1605] (1863) collected and edited by James Spedding, Robert Leslie Ellis, and Douglas Demon Heath, Cambridge, Riverside Press, vol. 6.
- Baecker, Dirk (1999) *Introduction*, I: *Problems of form I* edited by Dirk Baecker, Stanford, Stanford University Press, (Writing science), 287 s., 1-14.
- Baecker, Dirk (2004) *Niklas Luhmann i der Gesellschaft der Komputer*, I: Baecker, Dirk, *Wozu Soziologie?*, Berlin, Kulturverlag Kadmos, 351 s., ss. 125-149. Artikeln är det tyska originalet till *Niklas Luhmann in the Society of the Computer*.
- Baecker, Dirk (2006) *Niklas Luhmann in the Society of the Computer*, I: *Cybernetics & Human Knowing*, Volume 13, 2, ss. 25-40.
- Baecker, Dirk (2008) *The network Synthesis of Social Action II: Understanding Catjects*, I: *Cybernetics and Human Knowing*, Vol. 15, 1, ss. 45-66
- Baecker, Dirk (2001) *Why Systems?*, I: *Theory, Culture and Society*, Vol. 18,1, ss. 59-74, s. 59-60.

- Bainbridge, William Sims m.fl. (1994) *Artificial Social Intelligence*, I: *Annual Review of Sociology*, Vol. 20, ss. 407-436.
- Bateson, Gregory (1973) *Steps to an ecology of mind: collected essays in anthropology, psychiatry, evolution and epistemology*, New ed., repr., London, Granada, (Paladin books. Anthropology), 510 s.
- Bechtel, W., & A. Abrahamsen (1991) *Connectionism and the mind: An introduction to parallel processing in networks*, Oxford, Blackwell, 349 s.
- Becker, Howard Saul (1986) *Writing for social scientists: how to start and finish your thesis, book, or article*, Chicago, Univ. of Chicago Press, 180 s.
- Belkin, N. J., Oddy, R. N., & Brooks, H. M. (1982) *ASK for information retrieval: Part I. Background and theory*, I: *Journal of Documentation*, 38, 2, ss. 61-71.
- Bengtsson, Mats, & Hjern, Benny (1973) *Sociologisk uppslagsbok*, Stockholm, Rabén & Sjögren, 327s.
- Berg, Henk de & Prangel, Matthias (Ed.) (1993) *Kommunikation und Differenz: systemtheoretische Ansätze in der Literatur- und Kunstwissenschaft*, Opladen, Westdeutscher Verlag, 253 s.
- Bergström, Hans (2001) *Biologi för humanister*, I: *Dagens Nyheter*, 2001-06-13.
- Bloom, Paul (1996) *Intention, history, and artifact concepts*, I: *Cognition*, Vol. 60, ss. 1-29.
- Bloor, David (1976) *Knowledge and social imagery*, London, Routledge, 156 s.
- Bloor, David (1992) *Left and Right Wittgensteinians*, I: *Science as practice and culture*, Pickering, Andrew, red, Chicago, Univ. of Chicago Press, 474 s., ss. 266 – 282.
- Bloor, D. (1999) *Anti-Latour*, I: *Studies in History and Philosophy of Science*, Vol. 30, ss. 81-112.
- Borges, Jorge Luis (1998) *Biblioteket i Babel*, I: *Biblioteket i Babel: en antologi sammanställd ur novellsamlingarna Ficciones och El Aleph*, Stockholm, Bonnier, 154 s., ss. 31-38.
- Brante, Thomas (1984a) *Det rationella och det sociala*, I: *Att förstå världen: vetenskapsteoretiska essäer*, Stellan Welin (red.), Lund, Doxa, 166 s., ss. 23-50.

- Brante, Thomas (1984b) *Vetenskapens sociala grunder: en studie av konflikter i forskarvärlden*, 222 s.
- Brante, Thomas (2001) *Sociologiskt lexikon*, Stockholm, NoK, 363 s.
- Brier, Søren (2007) *Applying Luhmann's Systems Theory as Part of a Transdisciplinary Frame For Communication Science*, I: *Cybernetics and Human Knowing*, Vol. 14, 2-3, ss. 29-65.
- Burke, Kenneth (1945) *A Grammar of Motives*, New York, Prentice-Hall, 530 s.
- Carley, Kathleen M. (1996) *Artificial Intelligence within Sociology*, I: *Sociological Methods and Research*, Vol. 25, 1, Aug, ss. 3-30.
- Carley, Kathleen M. & Newell, Allen (1994) *The Nature of the Social Agent*, I: *Journal of Mathematical Sociology*, Vol. 19, 4, ss. 221-262.
- Cavalli-Sforza, Luigi Luca (2000) *Genes, peoples, and languages*, New York, North Point Press, 227 s.
- Clark, Andy (1993) *Associative engines: Connectionism, concepts, and representational change*. Cambridge, Mass.: MIT Press, 252 s.
- Clark, Andy (1996) *Being there: putting brain, body, and world together again*, Cambridge, Massachusetts, MIT Press, 269 s.
- Clark, Andy (2003) *Natural-born cyborgs: Minds, technologies, and the future of human intelligence*, Oxford University Press: Oxford, 229 s.
- Cole, George Douglas Howard (1920a) *Guild socialism re-stated*, London, Leonard Parsons, 224 s.
- Cole, George Douglas Howard (1920b) *Social theory*, London, Methuen, 219 s.
- Collins, Harry. M. (1996) *Embedded or Embodied: Hubert Dreyfus's What Computers Still Can't Do*, I: *Artificial Intelligence*, Vol. 80, 1, ss. 99-117.
- Collins, Harry. M & Kusch, Martin (1998) *The Shape of Actions: What Humans and Machines Can Do*, Cambridge, Mass. & London, MIT Press, 212 s.
- Collins, Randall (1992) *Can sociology create an artificial intelligence?*, I: *Sociological insight: an introduction to non-obvious sociology*, 2. ed., New York, Oxford University Press, 206 s., ss. 155-184.

- Connectionism* (1995) I: *The Oxford Companion to Philosophy*, Edited by Ted Honderich, Oxford, Oxford University Press, 1009 s., s. 151.
- Conte, Rosaria (1999) *Social Intelligence Among Autonomous Agents*, I: *Computational & Mathematical Organization Theory*, Vol. 5, 3, ss. 203–228.
- Dautenhahn, Kerstin, Ogden, Bernard, Quick, Tom (2002) *From embodied to socially embedded agents: Implications for interaction-aware robots*, I: *Cognitive Systems Research*, Vol. 3, 3, ss. 397–428.
- Dawkins, Richard (1988) *Den Blinde Urmakaren*, Stockholm, Wahlström & Widstrand, 328 s.
- Dawkins, Richard (1992) *Den Själviska Genen: en sociobiologisk studie*, 2. bearb. uppl, Stockholm, Tidens förlag, 383 s.
- Dawkins, Richard (1995) *River Out of Eden*, London, Weidenfeld and Nicolson, 172 s.
- Dennett, Daniel (1995) *Darwin's Dangerous Idea*, New York, Simon & Schuster, 586 s.
- Derrida, Jacques (1977) *Signature Event Context*, I: *Glyph: Johns Hopkins Textual Studies*. Baltimore, MD., Vol. 1, ss. 172-197
- Derrida, Jacques (1978) *Limited Inc abc....*, I: *Glyph: Johns Hopkins Textual Studies*. Baltimore, MD., Vol. 2, ss. 162-254.
- Dourish, Paul (2001) *Where the Action Is: The Foundations of Embodied Interaction*, Cambridge, Massachusetts, The MIT Press, 233 s.
- Dreyfus, Hubert. L & Dreyfus, Stuart.E. (1988) *Making a mind versus modelling the brain: artificial intelligence back at a branchpoint*, I: *Daedalus*, Vol. 117, ss. 15-43.
- Dreyfus, Hubert. L. (1972) *What Computers Can't Do*. New York, Harper & Row, 259 s.
- Dreyfus, Hubert. L. (1992) *Introduction to the MIT Press Edition*, I: *What Computers (Still) Can't Do*, 3rd ed, Cambridge, MA, MIT Press, 354 s., ss. ix-lII.

- Engels, Friedrich (1965) *Arbetets andel i apans förvandling till människan*, I: Gustafsson, Bo (red.), *Karl Marx, Friedrich Engels i urval*, Stockholm: Wahlström & Widstrand, 269 s.
- Engeström, Yrjö (1987) *Learning by Expanding: An Activity Theoretical Approach to Developmental Work Research*, Helsinki, Orienta Konsultit, 368 s.
- Filosoflexikonet* (1988) red. Poul Lübcke, Stockholm, Forum, 606 s.
- Flyman, Simon (2001) *Integration av artificiella agenter i mänskliga organisationer*, Uppsats Soc 446 41-80 p, Sociologiska institutionen, Lunds Universitet, vårterminen 2001, 72 s.
- Fuchs, Stephan (2001) *Against Essentialism: A Theory of Culture and Society*, Cambridge, Massachusetts and London, Harvard University Press, 380 s.
- Functionalism in Sociology I: International Encyclopedia of the Social & Behavioral Sciences*, ss. 5847 – 5852.
- Garson, James (2007) *Connectionism*, I: *The Stanford Encyclopedia of Philosophy* (Spring 2007 Edition), Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/spr2007/entries/connectionism>
- Geyer, Felix (1996) *Virtual Communities in Cyberspace*, I: *Kybernetes*, Vol. 25, 4, ss. 60-66.
- Gibson, K.R. & Ingold, T. (Eds.) (1993) *Tools, language and cognition in human evolution*, Cambridge: Cambridge University Press, 481 s.
- Giere, Ronald N. (1989) *Computer Discovery and Human Interest*, I: *Social Studies of Science: An International Review of Research in the Social Dimensions of Science and Technology*, Vol. 19, 4, ss. 638-643.
- Gilbert, G. Nigel, Heath, Christian (ed.) (1985) *Social Action and Artificial Intelligence*, (Surrey conferences on sociological theory and method, 3), Aldershot, England, Gower, 198 s.
- Gladwin, Thomas (1964) *Culture and logical process*, I: *Explorations in Cultural Anthropology: Essays Presented to George Peter Murdock*, ed. by W. Goodenough, New York, NY, McGraw-Hill, 635 s., ss. 167-177.
- Gorman, Michael E. (1989) *Beyond Strong Programmes: How Cognitive Approaches Can Complement SSK*, I: *Social Studies of Science: An International Review of Research in the Social Dimensions of Science and Technology*, Vol. 19, 4, ss. 643-653.

- Gould, Stephen Jay (1983) *Den felmätta människan*, Stockholm, Alba, 358 s.
- Gould, Stephen Jay & Eldredge, Nils (1977) *Punctuated equilibria*, I: *Paleobiology*, Vol. 3, 2, ss. 115-151.
- Gould, Stephen Jay (1989) *Livet är underbart: om femögda djur och naturens slumpmässiga historia*, Stockholm, Ordfronts förlag, 355 s.
- Grant, Colin B. (2003) *Destabilizing Social Communication Theory*, I: *Theory, Culture & Society*, Vol. 20, 6, ss. 95-119.
- Gusfield, Joseph R. (2003) *A Journey with Symbolic Interaction*, I: *Symbolic Interaction*, Vol. 26, 1, ss. 119-139.
- Habermas, Jürgen (1987) *Excursus on Luhmann's Appropriation of the Philosophy of the Subject through Systems Theory*, I: *Habermas, Jürgen, The philosophical discourse of modernity: twelve lectures*, Cambridge, Polity Press, 432 s.
- Hansson, Joacim (1999) *Klassifikation, bibliotek och samhälle: en kritisk hermeneutisk studie av "Klassifikationssystem för svenska bibliotek"*, Borås, 329 s.
- Haraway, Donna (1991) *Simians, Cyborgs, and Women: The Reinvention of Nature*. London, Free Association Books, 287 s.
- Harter, S. P. (1992) *Psychological relevance and information science*, I: *Journal of the American Society for Information Science*, Vol. 43, 9, ss. 602–615.
- Haselager, W.F.G. & van Rappard, J.F.H. (1998) *Connectionism, Systematicity, and the Frame Problem*, I: *Minds and Machines* Vol. 8, 2, ss. 161–179.
- Hayles, N. Katherine (1999) *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*, Chicago, University of Chicago Press, 350 s.
- Hayles, N. Katherine (2005) *My Mother was a Computer: Digital Subjects and Literary Texts*, Chicago, University of Chicago Press, 290 s.
- Heckathorn, Douglas D. (1989) *Cognitive science, sociology, and the theoretic analysis of complex systems*, I: *Journal of Mathematical Sociology*, Vol. 14, 2-3, ss. 97-110.
- Helfer, Joe (1998) *Order out of chaos: a practitioner's guide to Knowledge Management*, I: *Searcher*, Vol. 6, Issue 7, Jul/Aug, ss. 44-52.

- Hesse, Mary (1980) *Revolutions and reconstructions in the philosophy of science*, Brighton, Harvester, 271 s.
- Hutchins, E. (1995) *Cognition in the wild*, Cambridge, Mass., MIT Press, 381 s.
- Ingvarsson, Jonas (2003) *En besynnerlig gemenskap: teknologins gestalter i svensk prosa 1965-70*, Göteborg, Daidalos, 256 s.
- International Encyclopedia of the Social & Behavioral Sciences*, / editors-in-chief: Neil J. Smelser, Paul B. Baltes, Elsevier Science, Amsterdam: Elsevier, <http://www.sciencedirect.com.ludwig.lub.lu.se/science/referenceworks/9780080430768> (090301)
- Jönhill, Jan Inge (1997) *Samhället som system och dess ekologiska omvärld: en studie i Niklas Luhmanns sociologiska systemteori*, Lund, Sociologiska institutionen, 521 s.
- Kertész, András (1993) *Artificial intelligence and the sociology of knowledge: prolegomena to an integrated philosophy of science*, Frankfurt am Main: Lang, 176 s.
- King, Michael & Thornhill, Chris (2003) *Niklas Luhmann's Theory of Politics and Law*, Basingstoke, Palgrave Macmillan, 266 s.
- Knorr Cetina, Karin (1997) *Sociality with Objects: Social Relations in Postsocial Knowledge Societies*, I: *Theory, Culture and Society*, Vol. 14, 4, ss. 1-30.
- Knorr Cetina, Karin (1981) *The manufacture of knowledge: an essay on the constructivist and contextual nature of science*, Oxford, Pergamon press, 189 s.
- Korb, Kevin B. (1998) *The Frame Problem: An AI Fairy Tale*, I: *Minds and Machines* Vol. 8, 3, ss. 317 – 351.
- Laermans, Rudi (2007) *Theorizing Culture, or Reading Luhmann Against Luhmann*, I: *Cybernetics and Human Knowing*, Vol. 14, 2-3, ss. 67-83.
- Lakoff, George & Johnson, Mark (1980) *Metaphors we live by*, Chicago, University of Chicago Press, 242 s.
- Lakoff, George & Johnson, Mark (1999) *Philosophy in the flesh: the embodied mind and its challenge to western thought*, New York, Basic Books, 624 s.
- Latour, Bruno (1993) *We Have Never Been Modern*. Cambridge, Mass., Harvard University Press, 157 s.

- Latour, Bruno (2005) *Reassembling the Social: an Introduction to Actor-Network-Theory*, Oxford, Oxford University Press, 301 s.
- Laudan, Larry (1981) *The Pseudo-Science of Science?*, I: *Philosophy of the Social Sciences*, Vol. 11, 2, ss. 173-198.
- Leydesdorff, Loet (1996) *Luhmann's sociological theory: its operationalization and future perspectives: Collected papers on Niklas Luhmann*, I: *Social Science Information*, vol. 35, 2, ss. 283-306.
- Leydesdorff, Loet (2001) *A sociological theory of communication: the self-organisation of the knowledge-based society*, Parkland, FL, Universal publ., 351 s.
- Liedman, Sven-Eric (2001) *Ett oändligt äventyr: om människans kunskaper*, Stockholm, Albert Bonniers förlag, 411 s.
- Lindblom, J. & Ziemke, T. (2003) *Social Situatedness of Natural and Artificial Intelligence: Vygotsky and Beyond*, I: *Adaptive Behavior*, Vol. 11, 2, ss. 79-96.
- Luhmann, Niklas (1987) *The Evolutionary Differentiation between Society and Interaction*, I: *The Micro-Macro Link*, Alexander, Jeffrey, m.fl., Berkeley, University of California Press, 291 s., ss. 112-131.
- Luhmann, Niklas (1990) *Die Wissenschaft der Gesellschaft*, Frankfurt/M., Suhrkamp, 732 s.
- Luhmann, Niklas (1990a) *Communication and Society*, I: Luhmann, Niklas, *Essays on Self-Reference*, New York, Columbia University Press, 244 s.
- Luhmann, Niklas (1990b) *Essays on Self-Reference*, New York, Columbia University Press, 244 s.
- Luhmann, Niklas (1990c) *Meaning as Sociology's basic concept*, I: Luhmann, Niklas, *Essays on Self-Reference*, New York, Columbia University Press, 244 s., ss. 21-79.
- Luhmann, Niklas (1990d) *The individuality of the individual*, I: Luhmann, Niklas, *Essays on Self-Reference*, New York, Columbia University Press, 244 s., ss. 107-122.
- Luhmann, Niklas (1992a) *"Gibt es ein System der Intelligenz?"*, I: *Intellektuellendämmerung? Beiträge zur neuesten Zeit des Zeitgeistes*, Meyer, M. (ed), München, Hanser, s. 57-73.

- Luhmann, Niklas (1992b) *The Concept of Society*, I: *Thesis Eleven*, Vol. 31, ss. 67-80.
- Luhmann, Niklas (1995) *Social Systems*, Stanford, CA, Stanford University Press, 627 s.
- Luhmann, Niklas (1997) *Die Gesellschaft der Gesellschaft*, Frankfurt/M., Suhrkamp, vol. 1 & vol. 2, 1164 s.
- Luhmann, Niklas (1999) *Sign as Form*, I: *Problems of form I* edited by Dirk Baecker, Stanford: Stanford University Press, (Writing science), 287 s., ss. 46-63.
- Luhmann, Niklas (2000) *Art as a social system*, Stanford, Calif.: Stanford Univ. Press, 422 s.
- Luhmann, Niklas (2002) *How Can the Mind Participate in Communication?*, I: *Theories of Distinction: Redescribing the Descriptions of Modernity*, edited and introduced by William Rasch, Stanford Univ. Press, 226 s., s. 169 - 184.
- Luhmann, Niklas (2004) *Law as a Social System*. Oxford, Oxford Univ. Press, 498 s.
- Luhmann, Niklas, Hayles, N. Katherine, Rasch, William, Knodt, Eva and Wolfe, Cary (2000) *Theory of a Different Order: A Conversation with Katherine Hayles and Niklas Luhmann*, I: *Observing complexity: systems theory and postmodernity*, William Rasch och Cary Wolfe (ed), Minneapolis, MN, University of Minnesota Press, 308 s., ss. 111-136.
- Maasen, Sabine & Weingart, Peter (1995), *Metaphors - Messengers of Meaning*, I: *Science Communication*, Vol. 17 1, September, ss. 9-31.
- Malsch, Thomas (1997) *Die Provokation der "Artificial Societies": Warum die Soziologie sich mit den Sozialmetaphern der Verteilten Künstlichen Intelligenz beschäftigen sollte*, I: *Zeitschrift für Soziologie*, Vol. 26, 1, Feb, ss. 3-21.
- Malsch, Thomas (ed.) (1998) *Sozionik: Soziologische Ansichten über künstliche Sozialität*, Berlin, Edition Sigma, 393 s.
- Malsch, Thomas (2001) *Naming the Unnamable: Socionics or the Sociological Turn ofto Distributed Artificial Intelligence*. I: *Autonomous Agents and Multi-Agent Systems*, 4, ss. 155-186.

- Martinsson, Harry (1997) *Aniara: en revy om människan i tid och rum*, efterord av Johan Wrede, Stockholm, Bonnier i samarbete med Harry Martinson-sällsk., Originaluppl. 1956, 243 s.
- Maturana, Humberto R. & Varela, Francisco J. (1980) *Autopoiesis and Cognition: The Realization of the Living*, Dordrecht, D. Reidel, 141 s.
- Maturana, Humberto R. & Varela, Francisco J. (1987) *The Tree of Knowledge: The Biological Roots of Human Understanding*, Boston, New Science Library, 263 s.
- Merton, Robert K., & Barber, Elinor (2004) *The Travels and Adventures of Serendipity*, Princeton, Princeton Univ. Press, 313 s.
- Merton, Robert K. (1993) *On the Shoulders of Giants: The Post-Italianate Edition*, Chicago, University of Chicago Press, 320 s. (första uppl. utg. 1965 på The Free Press)
- Methodological Individualism in Sociology*, I: *International Encyclopedia of the Social & Behavioral Sciences*, s. 9751-9755. [<http://www.sciencedirect.com/science/referenceworks/0080430767>], (03-03-15)
- Miles, Chris (2006) *The Excluded Environment: Preliminary remarks towards a systems theory of literature*, I: *Cybernetics and Human Knowing*, Vol. 13, 1, ss. 19-33.
- Mills, C. Wright (1940) *Situated Actions and Vocabularies of Motive*. I: *American Sociological Review*, Vol. 5, 6, ss. 904-913.
- Mills, C. Wright (1985) *Den sociologiska visionen*, Lund, Arkiv, 259 s.
- Misheva, Vessela (2005) *Luhmann's Systems Theory and the Question of the Mass Media*, I: *Cybernetics and Human Knowing*, Vol. 12, 4, ss. 23-50.
- Moe, Sverre G. (2003) *Tid for Luhmann: en studie av tidens betydning i Niklas Luhmanns systemteori og samfunnsanalyse*, Tromsø, Institutt for sosiologi, 436 s.
- Mulkay, Michael Joseph (1979) *Science and the sociology of knowledge*, London, Allen & Unwin, 132 s.
- Nassehi, Armin, & Kneer, Georg (1997) *Niklas Luhmann: introduktion till sociale systemer*, Köpenhamn, Hans Reitzels Forlag, 212 s.

- [*Nationalencyklopedin*], NE.se [Elektronisk resurs]. (2000-). Malmö: Nationalencyklopedin.
- Nowotny, Helga, Scott, Peter, Gibbons, Michael (2001) *Re-Thinking Social Science*, Cambridge, Polity Press, 278 s., s. 98-120.
- Penrose, Roger (1989) *The Emperors New Mind: Concerning Computers, Minds, and the Laws of Physics*, Oxford, Oxford Univ. Press, 466 s.
- Plumpe, Gerhard & Werber, Niels (1993) *Literatur ist codierbar: Aspekte einer systemtheoretischen Literaturwissenschaft*, I: Schmidt, Siegfried J. (Ed.), *Literaturwissenschaft und Systemtheorie: Positionen, Kontroversen, Perspektiven*, Opladen, Westdeutscher Verlag, 296 s., ss. 9-43.
- Preda, Alex (1999) *The Turn to Things: Arguments for a Sociological Theory of Things*, I: *The Sociological Quarterly*, Vol. 40, 2, ss. 347-366.
- Pronzini, Andrea (2002) *First-Order Semantics and Artificial intelligence*, I: *Journal of Sociocybernetics*, Vol. 3, Number 1, ss. 1-20.
- Quine, Willard van Ormand (1960) *Word and Object*, Cambridge, Mass., MIT Press, 294 s.
- Radford, Gary P. (1992) *Positivism, Foucault, and the Fantasia of the Library: Conceptions of Knowledge and the Modern Library*, I: *Library Quarterly*, Vol 62, 4, ss. 408-424.
- Roberts, David (1999) *Self-reference in literature*, I: *Problems of form* / edited by Dirk Baecker, Stanford, Stanford University Press,, 287 p. (Writing science), s. 27-45.
- Ruesch, J. & Bateson, G. (1951) *Communication: The Social Matrix of Psychiatry*, New York, Norton, 314 s.
- Sawyer, Keith R. (2001) *Emergence in Sociology: Contemporary Philosophy of Mind and Some Implications for Sociological Theory*, I: *American Journal of sociology*, Vol. 107 (November 2001), 3, ss. 551-585.
- Sawyer, Keith R. (2003) *Artificial Societies: Multiagent Systems and the Micro-Macro Link in Sociological Theory*, I: *Sociological Methods and Research*, Vol. 31, 3, Feb, ss. 325-363.

- Schmidt, Benjamin Marius & Jahraus, Oliver (1999) *Systems Theory and Literary Studies in the 1990s: From the Difference between Symbolic and Social System to the Difference between Consciousness and Communication*, I: *Germanic Review (GR)*, Vol. 74, 3, Summer, ss. 242-54.
- Schmidt, Siegfried (1989) *Die Selbstorganisation des Sozialsystems Literatur im 18. Jahrhundert*, Frankfurt am Main, Suhrkamp, 489 s.
- Schneider, Wolfgang Ludwig (1996) *Die Komplementarität von Sprechakttheorie und Systemtheoretischer Kommunikationstheorie. Ein hermeneutischer Beitrag zur Methodologie von Theorievergleichen*, I: *Zeitschrift für Soziologie*, Vol. 25, 4, Aug, ss. 263-277.
- Schwanitz, Dietrich (1990) *Systemtheorie und Literatur: ein neues Paradigma*, Opladen, Westdeutscher Vlg, cop., 284 s.
- Schütz, Alfred (1967) *The phenomenology of the social world; with an introduction by George Walsh*; translated by George Walsh and Frederick Lehnert, Evanston: Northwestern Univ. Press, 255 s.
- Schütz, Alfred [1899–1959] I: *International Encyclopedia of the Social & Behavioral Sciences*, ss. 13603-13607.
- Schütz, Alfred (1970) *Reflections on the problem of relevance*, New Haven and London, Yale University Press, 186 s.
- Searle, John (1977) *Reiterating the Differences: A Reply to Derrida*, I: *Glyph: Johns Hopkins Textual Studies*. Baltimore, MD., Vol. 1, ss. 198-208.
- Searle, John (1980) *Minds, Brains and Programs*, I: *The Behavioral and Brain Sciences*, Vol. 3, ss. 417-424.
- Segerstråle, Ullica (2000) *Defenders of the Truth: the Sociobiology Debate*, Oxford, Oxford University Press, 494 s.
- Seidl, David & Becker, Kai Helge (2006) *Organizations as Distinction Generating and Processing Systems: Niklas Luhmann's Contribution to Organization Studies*, I: *Organization*, Vol. 13, 1, ss. 9-35.
- Shannon, C. E. & Weaver, W. (1949) *The Mathematical Theory of Communication*, Urbana, IL: University of Illinois press, 117 s.
- Simon, Willian & Gagnon, John H. (1967) *Recension I: American Journal of Sociology*, Vol 72, 4, ss. 424-425.

- Slezak, Peter (1989) *Scientific Discovery by Computer as Empirical Refutation of the Strong Programme*, I: *Social Studies of Science: An International Review of Research in the Social Dimensions of Science and Technology*, Vol. 19, 4, ss. 563-600.
- Sokal, Alan (1996a) *A Physicist Experiments with Cultural Studies*, I: *Lingua Franca*, Vol. 6, 4, May/June, ss. 62-64.
- Sokal, Alan (1996b) *Transgressing the Boundaries: Toward a Transformative Hermeneutics of Quantum Gravity*, I: *Social Text*, 46/47, Vol. 14, Nos. 1 and 2 (Spring/Summer 1996), ss. 217-252.
- Spencer-Brown, George (1979) *Laws of form*, New York, Dutton, 141 s.
- Star, Susan Leigh (1989) *The Structure of Ill-Structured Solutions: Boundary Objects and Heterogenous Distributed Problem Solving*, I: *Distributed Artificial Intelligence*, Vol. II, ed. by Les Gasser and Michael N. Huhns, London, Pitman, 519 s., ss. 37-54.
- Star, Susan Leigh (1991) *Power, technology and the phenomenology of conventions: on being allergic to onions*. I: *Sociology of monsters: Essays on Power, Technology and Domination*, John Law (ed), London, Routledge, 273 s., ss. 26-56.
- Stichweh, Rudolf (2000) *Systems Theory as an Alternative to Action Theory? The Rise of 'Communication' as a Theoretical Option*, I: *Acta Sociologica*, Vol. 43, 1, ss. 5-13.
- Suchman, Lucy A. (1987) *Plans and situated actions: The problem of human-machine communication*, Cambridge, Cambridge University Press, 203 s.
- Susi, Tarja (2006) *The Puzzle of Social Activity: the significance of tools in cognition and cooperation*, Department of Computer and Information Science: Linköpings universitet, (Linköping Studies in Science and Technology, Dissertation No. 1019), 211 s.
- Svenska akademiens ordlista över svenska språket* (1998) Stockholm, Norstedts ordbok, 12. Uppl., 1066 s.
- Tuomi, Ilkka (1999a) *Corporate Knowledge: theory and Practice of Intelligent Organizations*, Helsingfors, Metaxis, 453 s.

- Tuomi, Ilkka (1999b) *Data is more than knowledge: implications of the reversed knowledge hierarchy to knowledge management and organizational memory*, Institute of Electrical and Electronics Engineers, Inc. (EEE), Proceedings of the 32nd Hawaii International Conference on System Sciences (HICSS-32), 12 s.
- Turing, Alan M. (1950) *Computing Machinery and Intelligence*, I: *Mind*, Vol. 59, No. 236, ss. 433-460 [Reprinted in *Minds and machines*, A. Anderson (ed.), Engelwood Cliffs NJ, Prentice Hall, 1964.]
- Turkle, Sherry (1997) *Leva Online*, Stockholm, Norstedts, 401 s.
- Udén, Lars (2002) *The changing face of methodological individualism*, I: *Annual Review of Sociology*, Vol. 28, ss. 479-507.
- Vanderstraeten, Raf (2000) *Autopoiesis and socialization: on Luhmann's reconceptualization of communication and socialization*, I: *British Journal of Sociology*, Vol. No. 51, Issue No. 3, (September), ss. 581-598.
- Varela, Francisco J., Thompson, Evan & Rosch, Eleanor (1991) *The Embodied Mind: Cognitive Science and Human Experience*, Cambridge, MIT Press, 308 s.
- Warwick, Kevin (2002) *I, Cyborg*, London, Century, 312 s.
- Wertsch, James W. (1985) *Vygotsky and the Social Formation of Mind*, Cambridge Mass., Harvard University Press, 262 s.
- Wertsch, James W. (1998) *Mind as Action*, Oxford, Oxford University Press, 203 s.
- Wilson, Edward O. (1975) *Sociobiology: the new synthesis*, Cambridge, Mass., Belknap P. of Harvard U.P., 697 s.
- Vinge, Vernor (1993) *The Coming Technological Singularity: How to Survive in the Post-Human Era*, <http://www-rohan.sdsu.edu/faculty/vinge/misc/singularity.html> 2008-05-11.
- Wolfe, Alan (1991) *Mind, Self, Society, and Computer: Artificial Intelligence and the Sociology of Mind*, I: *American Journal of Sociology*, Vol. 96, 5, Mar, ss. 1073-1096.
- Wolfe, Alan (1993) *The Human difference: animals, computers, and the necessity of social science*, Berkeley and Los Angeles, California, University of California Press, 243 s.

- Woolgar, Steve (1985) *Why not a sociology of machines? The case of sociology and artificial intelligence*. I: *Sociology*, vol. 19, 4, ss. 557–572.
- Ziemke, Tom & Susi, Tarja (2001) *Social cognition, artefacts, and stigmergy: A comparative analysis of theoretical frameworks for the understanding of artefact-mediated collaborative activity*, I: *Journal of Cognitive Systems Research*, Vol. 2, 4, ss. 273–290.
- Österberg, Dag (2000) *Luhmann's General Sociology*, I: *Acta Sociologica*, Vol. 43, No. 1., ss. 15-25.