

Spela Halite IV med förstärkningsinlärning

POPULÄRVETENSKAPLIG SAMMANFATTNING, Kim Haapamäki och Jesper Laurell

Agenter som kan lära sig agera i en simulerad miljö och bli bättre genom maskininlärning, kallat förstärkningsinlärning, är ett snabbt växande område inom AI. Projektet har utforskat olika aspekter av förstärkningsinlärning med hjälp av spelet Halite IV och har uppnått agenter som spelar på samma nivå som vissa traditionellt konstruerade lösningar.

Att låta datorer lära sig spela spel är något som under de senaste åren visat stor framgång och är något som har stor potential att vara till nytta även utanför spel när lösningarna är effektiva nog att även lära sig problem som människor kämpar med. Några exempel är Deepmind och OpenAI som genom sitt arbete har lyckats skapa agenter som spelar på överhuman nivå, de mest välkända är spelen Go och Dota2. Go är ett kinesiskt brädspel som ansetts vara ett av de svåraste spelen för en dator att kunna spela kompetent, detta på grund av det stora antalet drag en spelare kan göra under spelets gång. Dota2 är ett lagspel som spelas med 5 spelare i två olika lag och har ett ofantligt stort antal tillstånd och strategier.

Med hjälp av förstärkningsinlärning möjliggör vi att datorer utan förutfattade meningar försöker hitta den mest optimala metoden för ett givet problem. Detta uppnås genom att en agent styrs och testas fram i en miljö och utforskar vad som är bra och dåligt för att sedan lära sig av detta. Detta är något som vi i framtiden kan ha nytta av inom uppgifter som endast en människa klarar av idag. Att istället ha datorer som styr skulle kunna leda till att uppgifterna utförs både säkrare, snabbare och effektivare med hjälp förstärkningsinlärning. Ett bra exempel är självkörande bilar. Dessa skulle kunna kommunicera med varandra över internet samt ha ett blixtnabbt beslutsfattande i kritiska lägen, för att göra bilkörande mycket säkrare än vad människor klarar av i dagsläget.

I detta examensarbete har förstärkningsinlärning varit i fokus. Vi har utforskat hur en lösning tas fram och vilka metoder som finns tillgängliga för att uppnå resultat som presterar på en konkurrenskraftig nivå. Dessutom har vi undersökt ifall förstärkningsinlärning, som ofta kräver otroligt mycket datorkraft och tid, också kan lösa problem när det finns begränsad mängd tid och datorkraft. Genom att använda den färdiga miljön Halite IV har 8 olika tekniker och parametrar analyserats och diskuterats. Genom vårt arbete har vi funnit att förstärkningsinlärning har potentialen att med hjälp av korrekt justerade parametrar, och genomtänkta belöningar vara kapabel att lära sig komplexa miljöer med datorkraft som är tillgänglig för privatpersoner. Lösningarna som en förstärkningsinlärningsagent lär sig är inte satta i sten utan

har även potential att kunna förbättras över tid.

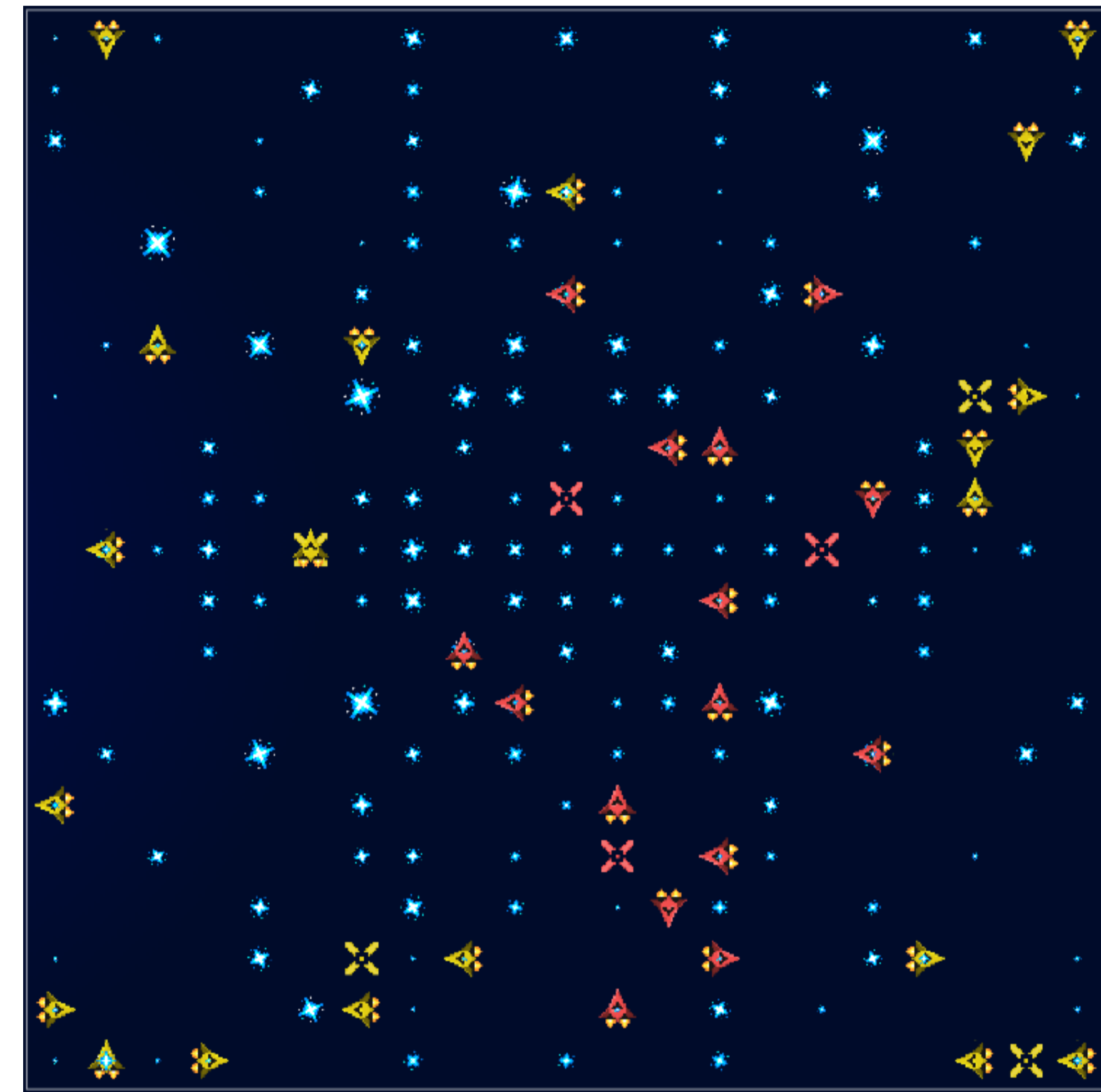


Figura 1: Spelet Halite IV, gul och röd agent tävlar om de blåa resurserna (halite) strödda över spelplanen

Under projektets gång har vi även fått insikter inom lagarbete i förstärkningsinlärning då lösningen som agenten använder sig utav endast styr en (av flera lagmedlemmar) i taget. Ett centralt dilemma för arbetet har varit hur en enskild lagmedlem bör belönas utifrån dess egna beslut och hur de enskilda belöningarna påverkar laget som helhet. Detta visade sig vara ett mycket svårt problem att lösa, men en otroligt användbar metod för att förenkla vissa implementationer.

Genom att låta våra datorer spela och träna på miljön Halite IV har vi lyckats skapa flera agenter som var och en har något unikt. Med dessa agenter samlade vi sedan in resultat för att avgöra hur varje agents särskiljande aspekter påverkar slutresultatet. Genom denna metod har vi i arbetet kommit fram till många insikter som kan vara till hjälp inom andra lösningar för förstärkningsinlärning i framtiden.