

Self-Learning Game Player

Deltagare

- Emre Berge Ergenekon 880427-1131 erge@kth.se
- Anton Jonsson 880408-0573 antonjon@kth.se

Handledare

- Mats Dam mfd@kth.se

Sammanfattning

Man kan med hjälp av belöningsbaserad inlärning få ett datorprogram att lära sig bästa tillvägagångssätt för att lösa problem. Problemet kan exempelvis vara att hitta vägen ut ur en labyrint. Vi ska undersöka hur "Q-learning"-algoritmen fungerar samt analysera hur olika parametrar påverkar inlärningsprocessen vid implementering i ett enklare brädspel. Resultat kommer att redovisas i en skriftlig rapport.

Bakgrund

Belöningsbaserad inlärning (eng. reinforcement learning; förkortat RL) är en del av det stora området maskininlärning (Alpaydin, 2004). Det finns flera andra delområden så som övervakad inlärning (eng. supervised learning) och oövervakad inlärning (eng. unsupervised learning) där belöningsbaserad inlärning kan sägas ligga ungefär emellan delområdena (Gasser 2010). Teorin bakom belöningsbaserad inlärning är att låta ett datorprogram själv lära sig vilka handlingar som är bäst i olika situationer genom att testa möjliga handlingar och observera resultatet (Sutton 1998, Kapitel 1.1).

Målet med belöningsbaserad inlärning är att agentens (programmet som ska lösa en uppgift eller spela ett spel) handlingar i en miljö ska maximera någon form av långsiktig belöningen och därmed få strategiska kunskaper. Metoden har framgångsrikt applicerats på olika problem, däribland robot kontroll, schemaläggning av hissar, telekommunikation, backgammon och schack (Sutton 1998, Kapitel 11).

Problemformulering

Undersöka hur Q-learning algoritmen fungerar, implementera den för att lära datorn att spela ett enklare brädspel (tre-i-rad, luffarschack eller liknande), samt analysera hur olika parametrar påverkar inlärningsprocessen.

Projektplan

1. Insamling av information samt inlärning av metoder och algoritmer som berör belöningsbaserad inlärning.
2. Enkel implementering av algoritmen "Q-Learning"
3. Insamling, analysering samt utvärdering av mätdata.
4. Publicering av resultat i form av en slut rapport.

Tidsplan

	Mittvägs mötet		Inlämning		
	Februari	Mars	April	Maj	
Efterforskning	[Progress bar]				
Implementering	[Progress bar]				
Analysering	[Progress bar]				
Rapportskrivning	[Progress bar]				

Referenser

- Teknomo, Kardi. 2005. Q-Learning by Examples.
<http://people.revoledu.com/kardi/tutorial/ReinforcementLearning/index.htm>
- R. Sutton and R. Barto. Reinforcement learning. The MIT Press, 1998
- Ethem Alpaydin (2004) Introduction to Machine Learning (Adaptive Computation and Machine Learning), MIT Press, ISBN 0262012111
- Kaelbling, Leslie P.; Michael L. Littman; Andrew W. Moore (1996). "Reinforcement Learning: A Survey". Journal of Artificial Intelligence Research 4: 237–285.
<http://www.cs.washington.edu/research/jair/abstracts/kaelbling96a.html>.
- Peters, Jan; Sethu Vijayakumar; Stefan Schaal (2003). "Reinforcement Learning for Humanoid Robotics". IEEE-RAS International Conference on Humanoid Robots.
<http://www-clmc.usc.edu/publications/p/peters-ICHR2003.pdf>.
- Michael Gasser (hämtad: 28 jan 2010). "Introduction to Reinforcement Learning".
<http://www.cs.indiana.edu/~gasser/Salsa/rl.html>