

UNIVERZA NA PRIMORSKEM
FAKULTETA ZA MATEMATIKO, NARAVOSLOVJE IN
INFORMACIJSKE TEHNOLOGIJE

Zaključna naloga
Strojno učenje iz interakcije
Machine learning from interaction

Ime in priimek: Rok Breulj
Študijski program: Računalništvo in informatika
Mentor: doc. dr. Peter Rogelj

Koper, Marec 2014

Ključna dokumentacijska informacija

Key words documentation

Zahvala

Preface? Foreword?

Kazalo

1	Uvod	7
1.1	Okrepiteveno učenje	7
1.2	Primeri	7
1.3	Elementi okrepitevenega učenja	7
2	Problem	8
2.1	Ocenjevanje povratne informacije	8
2.2	Celoten problem okrepitevenega učenja	8
3	Tabularne rešitve	9
3.1	Dinamično programiranje	9
3.2	Predvidevanje - vrednost stanja	9
3.2.1	Monte Carlo metode	9
3.2.2	Učenje na podlagi časovne razlike - TD(0)	9
3.2.3	Združitev metod - TD(λ)	9
3.3	Krmiljenje - vrednost dejanja	9
3.3.1	Monte Carlo metode	9
3.3.2	Učenje na podlagi časovne razlike - TD(0)	9
3.3.3	Združitev metod - TD(λ)	9
4	Posploševanje in funkcijska aproksimacija	10
5	Namizna igra Hex	11
5.1	Ozadje	11
5.2	Učenje	11
6	Zaključek	12
7	Literatura	13
8	Priloge	14

Tabele

Slike

1 Uvod

Ideja učenja iz interakcije z našim okoljem je ena od prvih, ki nam pride na misel, ko razmišljamo o naravi učenja. Ko se dojenček igra, maha z rokami ali gleda naokoli nima izrecnega učitelja, ima pa neposredno senzomotorično povezavo z okoljem. Uporaba te povezave proizvede ogromno informacije o vzrokih in učinkih, o posledicah dejanj in načinih kako doseči cilje. Skozi naše življenje so takšne interakcije nedvoumno velik izvir znanja o našem okolju in samim sebi. Ko se učimo voziti avto ali pogovarjati, se zavedamo kako se okolje odziva na naša dejanja in iščemo način kako vplivati na rezultat z našim vedenjem. [1]

What's in this work.

1.1 Okrepitveno učenje

1.2 Primeri

1.3 Elementi okrepitvenega učenja

2 Problem

2.1 Ocenjevanje povratne informacije

2.2 Celoten problem okrepitvenega učenja

3 Tabularne rešitve

3.1 Dinamično programiranje

3.2 Predvidevanje - vrednost stanja

3.2.1 Monte Carlo metode

3.2.2 Učenje na podlagi časovne razlike - TD(0)

3.2.3 Združitev metod - TD(λ)

3.3 Krmiljenje - vrednost dejanja

3.3.1 Monte Carlo metode

3.3.2 Učenje na podlagi časovne razlike - TD(0)

3.3.3 Združitev metod - TD(λ)

4 Posploševanje in funkcijska aproksimacija

5 Namizna igra Hex

5.1 Ozadje

5.2 Učenje

6 Zaključek

7 Literatura

- [1] Richard S. Sutton, Andrew G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, MA, 1998

8 Priloge