

MRNA - seminar 1

Andraž Vrhovec

March 23, 2012

0.1 Hiter opis

Imamo semafor z N luckami, kjer je ob zacetku igre prizgana srednja luc. V igri sodelujeta dva igralca, ki izmenicno pritiskata tipki R1 ($\alpha = 1$) in R2 ($\alpha = 2$). Ob pritisku na tipko R1 se lucka premakne levo z verjetnostjo p in desno z verjetnostjo $p - 1$. Ob pritisku na R2 se lucka premakne levo z verjetnostjo $p - 1$ in desno z verjetnostjo p . Cilj prvega igralca je lucko spraviti na skrajni levi rob semaforja, cilj drugega igralca pa na desni rob, pri cemer nihce ne ve verjetnosti p .

0.2 Primerjava avtomatov z PCA

Najprej sem vse avtomate primerjal z PCA. V vsaki primerjavi sem izvedel 100 ponovitev z maksimalnim številom korakov 1000. Priznava se je izkazalo da vecja kot je verjetnost p , boljse je učenje avtomata in da se pri vrednostih p blizu 0.5 avtomati obnasajo podobno kot PCA. Opaziti je tudi da pri majhnem N , lahko dodatni pomnilnik pri $L_{2n,2}$ skoduje procesu učenja. Vrednosti -1 v tabeli pomenijo da se simulacija ni izvedla v doglednem casu, zato sem jo prekinil.

•	$L_{2,2}$	$L_{6,2}$	L_{r-p}	L_{r-i}
PCA	1,27	1,13	1,56	1,22

Figure 1: Razmerje $\text{zmag}(\frac{|avtomat_i|}{|PCA|})$ pri $N=7$ in $p=0,6$

•	$L_{2,2}$	$L_{6,2}$	L_{r-p}	L_{r-i}
PCA	2,33	1,32	1,94	1,32

Figure 2: Razmerje $\text{zmag}(\frac{|avtomat_i|}{|PCA|})$ pri $N=7$ in $p=0,8$

•	$L_{2,2}$	$L_{6,2}$	L_{r-p}	L_{r-i}
PCA	1,38	1,56	1,33	1,27

Figure 3: Razmerje $\text{zmag}(\frac{|avtomat_i|}{|PCA|})$ pri $N=19$ in $p=0,6$

•	$L_{2,2}$	$L_{6,2}$	L_{r-p}	L_{r-i}
PCA	11,5	32,33	6,14	6,14

Figure 4: Razmerje $\text{zmag}(\frac{|avtomat_i|}{|PCA|})$ pri $N=19$ in $p=0,8$

•	$L_{2,2}$	$L_{6,2}$	L_{r-p}	L_{r-i}
PCA	3,62	91	-1	-1

Figure 5: Razmerje $\text{zmag} \left(\frac{|avtomat_i|}{|PCA|} \right)$ pri $N=101$ in $p=0,6$

•	$L_{2,2}$	$L_{6,2}$	L_{r-p}	L_{r-i}
PCA	∞	∞	∞	∞

Figure 6: Razmerje $\text{zmag} \left(\frac{|avtomat_i|}{|PCA|} \right)$ pri $N=101$ in $p=0,8$

0.3 Primerjava avtomatov med seboj

Primerjava avtomatov med seboj je bila izvedena podobno kot primerjava z PCA, le da so se tokrat kot drugi igralec izmenjevali različni avtomati. Primerjati enak avtomat z enakim se mi ni zdelo smiselno, saj oba zavzameta podobne strategije in je izid odvisen samo od sreče.

•	$L_{2,2}$	$L_{6,2}$	L_{r-p}	L_{r-i}
$L_{2,2}$	•	0,75	0,59	0,72
$L_{6,2}$	•	•	1,13	0,78
L_{r-p}	•	•	•	1,17
L_{r-i}	•	•	•	•

Figure 7: Razmerje $\text{zmag} \left(\frac{|avtomat_{zgoraj}|}{|avtomat_{levo}|} \right)$ pri $N=7$ in $p=0,8$

•	$L_{2,2}$	$L_{6,2}$	L_{r-p}	L_{r-i}
$L_{2,2}$	•	3,55	1,63	0,82
$L_{6,2}$	•	•	0,47	0,78
L_{r-p}	•	•	•	0,19
L_{r-i}	•	•	•	•

Figure 8: Razmerje $\text{zmag} \left(\frac{|avtomat_{zgoraj}|}{|avtomat_{levo}|} \right)$ pri $N=19$ in $p=0,8$

0.4 Primerjava parametrov korekcijskih shem

V prvem primeru sem primerjal dva L_{r-p} avtomata z različnimi parametri. Pri avtomatu z večjimi parametri se vidi večje odklone v verjetnosti izbire skozi čas, vendar se proti koncu ujameta kar se vidi tudi na razmerju zmag , ki je na koncu blizu 1.

V drugem primeru sem primerjal L_{r-p} z L_{r-i} z enakim a . Visi se veliko lepše učenje L_{r-i} avtomata, kar se pozna tudi na koncnem razmerju zmag 95 : 5. Iz teh rezultatov zaključujem da je L_{r-i} avtomat boljši za okolja kjer je

verjetnost kaznovanja fiksna, vendar neznana. Dopuscam možnost da se L_{r-p} boljše obnaša v dinamičnih okoljih, kjer se verjetnost kaznovanja spreminja, saj je njegov proces učenja bolj fleksibilen v obe smeri.

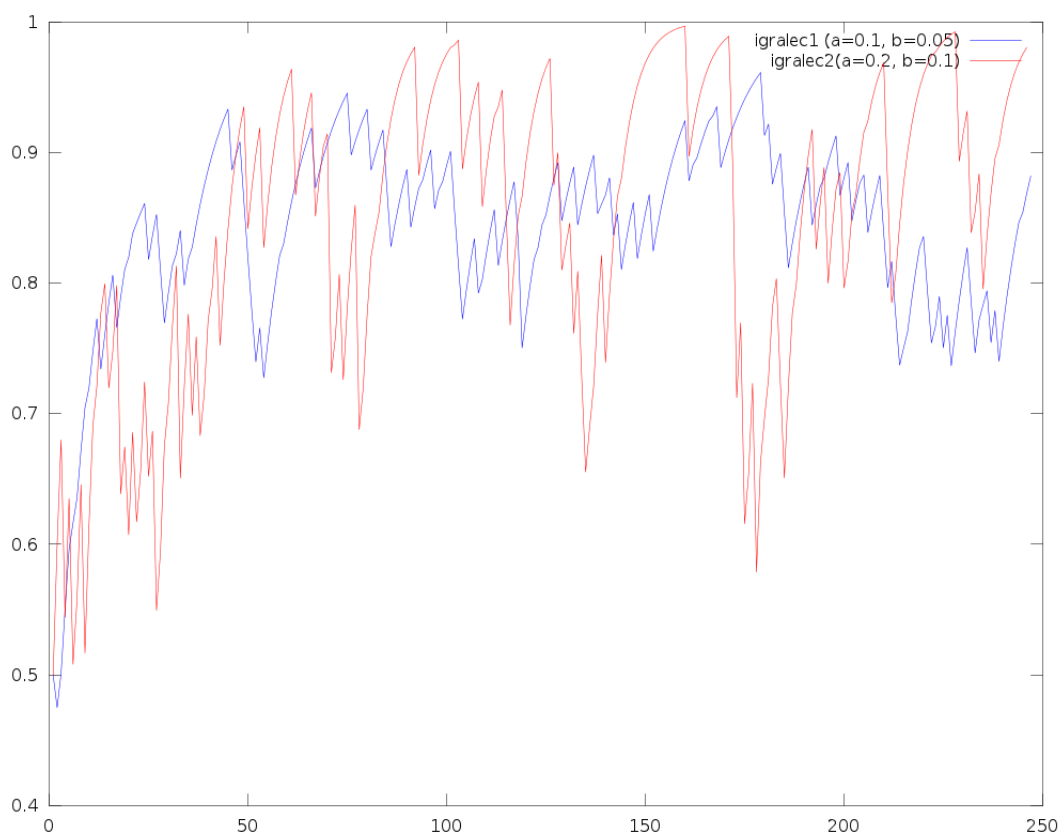


Figure 9: Primerjava dveh L_{r-p} avtomatov z različnimi parametri. Graf prikazuje verjetnost izbire za avtomat ugodne akcije (za igralca 1 je to levo, za drugega desno)

0.5 Igra avtomata z človekom

Preizkusil sem se v igri z $L_{2n,2}$ in L_{r-i} pri $N=7$. Za manjše N je igra za človeka se obvladljiva in je možno premagati avtomat v vsaj polovici primerov. Ko pa N narasca pa količina podatkov za človeka postane neobvladljiva in se avtomati izkazuje dosti bolje.

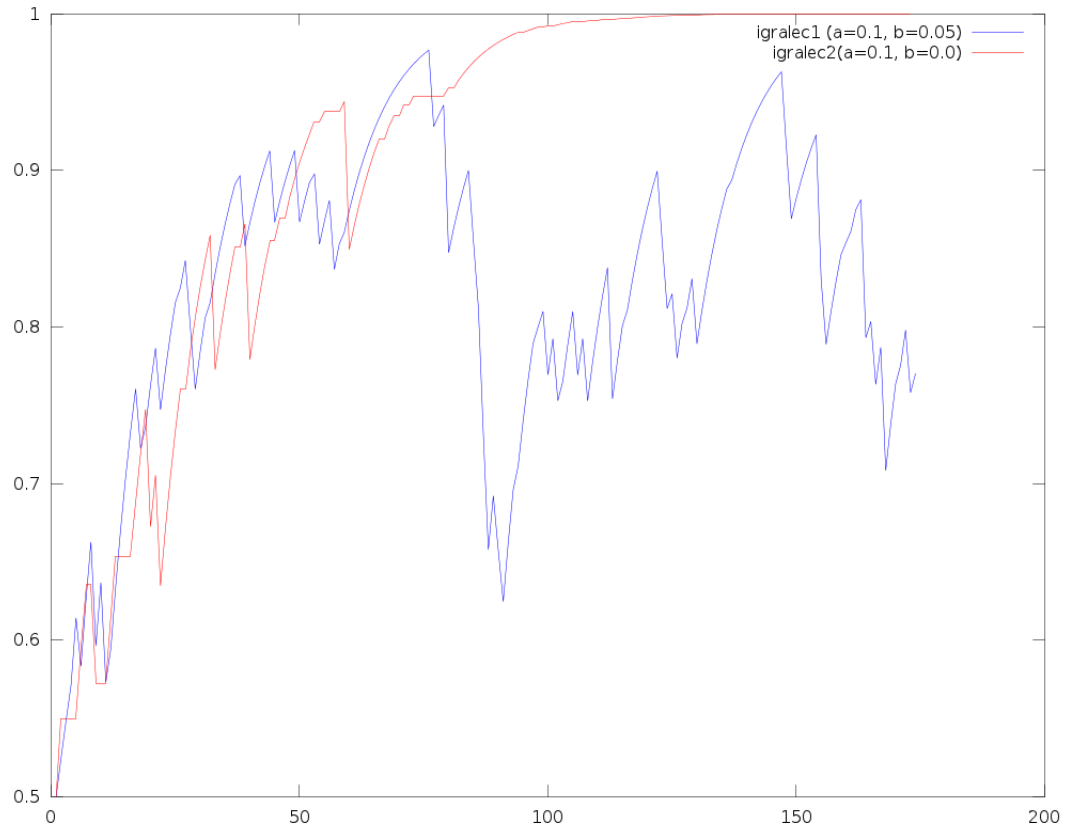


Figure 10: Primerjava L_{r-p} in L_{r-i} avtomatov z enakim a . Graf prikazuje verjetnost izbire za avtomat ugodne akcije (za igralca 1 je to levo, za drugega desno)

0.6 Zaključek

Predvsem se bil presenecen nad uspešnostjo avtomata $L_{2n,2}$, ki kljub svoji preprostosti kaze veliko zmognost učenja v tem primeru in premaguje druge, po zasnovi naprednejše avtomate. Kot drugi najboljši se izkaze L_{r-i}

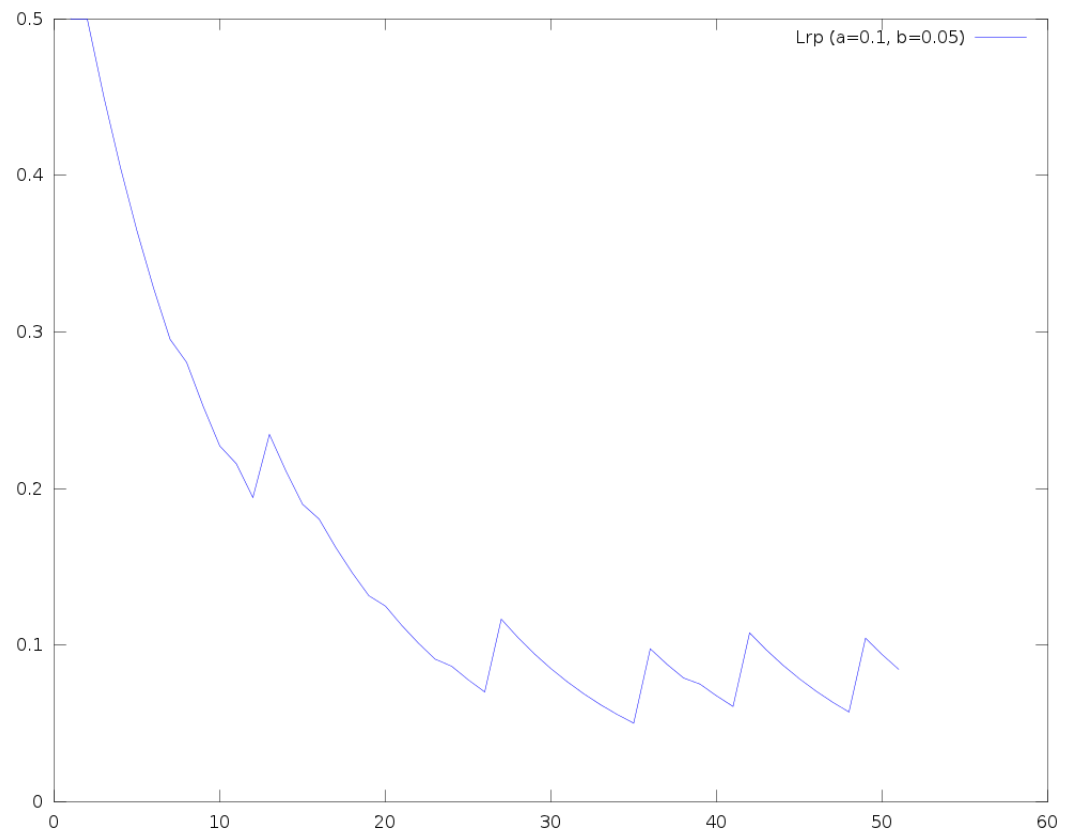


Figure 11: Casovna odvisnost $M(n)$ pri avtomatu L_{r-p}

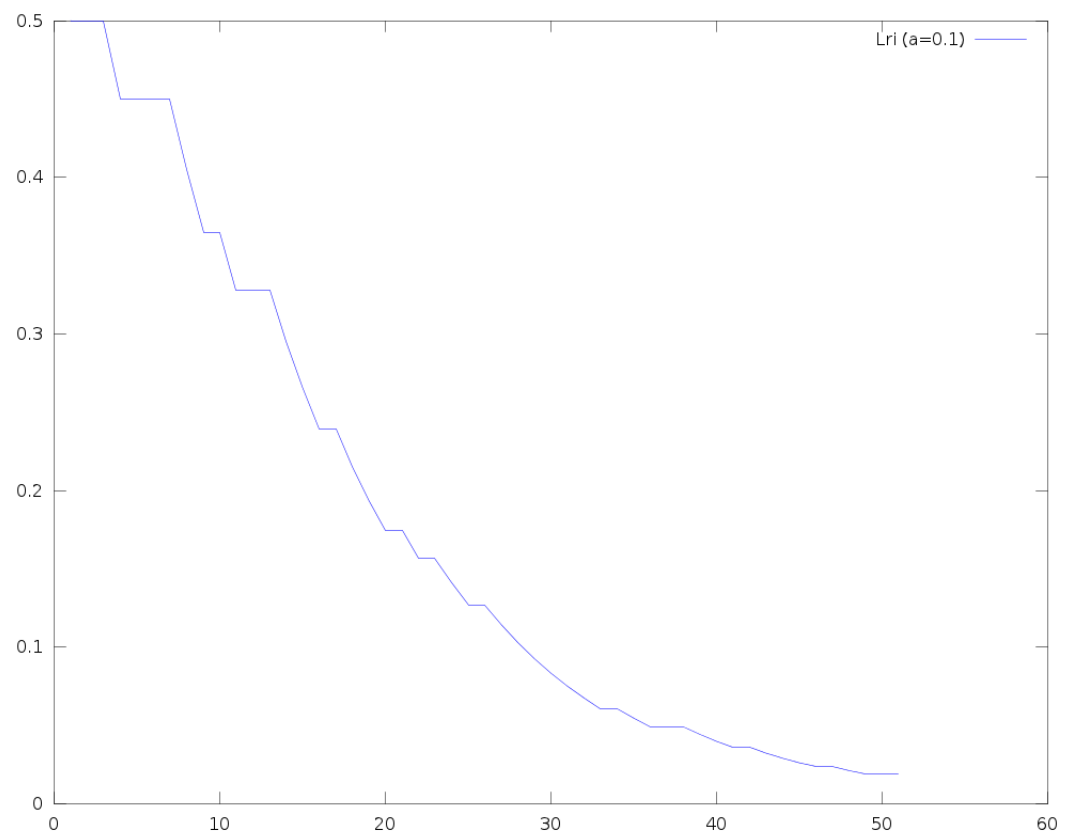


Figure 12: Casovna odvisnost $M(n)$ pri avtomatu L_{r-i}