

RESEARCH SCHOOL OF COMPUTER SCIENCE
COLLEGE OF ENGINEERING AND
COMPUTER SCIENCE

An Implementation of MC-AIXI-CTW

Jarrah Bloomfield* Luke English[†] Andrew Haigh[‡] Joshua Nelson[§] Anthony Voutas[¶]

COMP4620 - ADVANCED ARTIFICIAL INTELLIGENCE
ASSIGNMENT 2

October 31, 2012

u4669875* u4667010[†] u4669875[‡] u4850020[§] u4519169[¶]

Chapter 1

Description of the MC-AIXI-CTW implementation

The AIXI agent is a formal, mathematical agent which represents a solution to the general reinforcement learning problem. Principally, AIXI consists of an expectimax search over a Bayesian mixture of Turing machines in order to choose optimal actions by predicting future observations and rewards based on past experience. Given infinite computational resources, AIXI represents the optimal reinforcement agent: maximising future expected reward in any unknown environment.

In the far more limited world of what is tractable, we require an approximation to AIXI. Here, we approximate AIXI via a combination of UCT search (Monte-Carlo Tree Search with the Upper Confidence Bound) (Kocsis and Szepesvári, 2006) and Context Tree Weighting (Willems et al., 1995), yielding MC-AIXI-CTW.

Chapter 2

User Manual

2.1 Arguments

The agent can be compile with the make command. The agent then can then be run using

```
./main <environment> <logfile>
```

<environment> is a compulsory argument, which specifies the environment configuration file the agent is to use. In this implementation, it is one of the following

- `coinflip.conf`: Biased coin flip enviroment
- `grid.conf`: Gridworld environment
- `kuhnpoker.conf`: Kuhn poker environment
- `pacman.conf`: Pacman environment
- `rps.conf`: Biased rock paper scissors environment
- `tiger.conf`: "Tiger" Environment
- `composite.conf`: A combination of the above environments

<logfile> is an optional argument, which specifies the name of a log file to output results to.

2.2 Configuration files

.conf files are *configuration files*, specifying which environment is to be used, and relevant parameters for each environment. Each configuration file has the following parameters

- `environment`: The name of the environment to use. One of {4x4-grid, kuhn-poker, tiger, biased-rock-paper-scissor, pacman, composite}.
- `exploration`: The rate at which the agent explores, by making random decisions.
- `explore-decay` : The rate at the exploration rate decreases

In addition to this, some configurations have parameters that are specific to their environments.

- Coinflip
 - `coin-flip-p`: The probability of a flipping heads ($0 \leq \text{coin-flip-p} \leq 1$).
- Kuhn poker

- `gamma`: A constant that determines the environment's Nash equilibrium strategy. ($0 \leq \text{gamma} \leq 1$)
- `Pacman`
 - `mapfile`: The location of the map file for the pacman board.
- `tiger.conf`
 - `left-door-p`: The probability that the gold is behind the left door
 - `listen-p`: The probability that a listening observation is correct.
- `composite.conf`
 - `environmentN`: Specifies the N^{th} environment, where $0 \leq N \leq 10$. The value of this parameter is an integer ≤ 10 , and indicates which environment `environmentN` represents.
 - `startN`: Specifies the time step that at which the N^{th} environment starts, where $0 \leq N \leq 10$.
 - Parameters required for the environments 1..N specified in `environment.cpp`.

Chapter 3

MC-AIXI-CTW Implementation

3.1 Context Tree Weighting

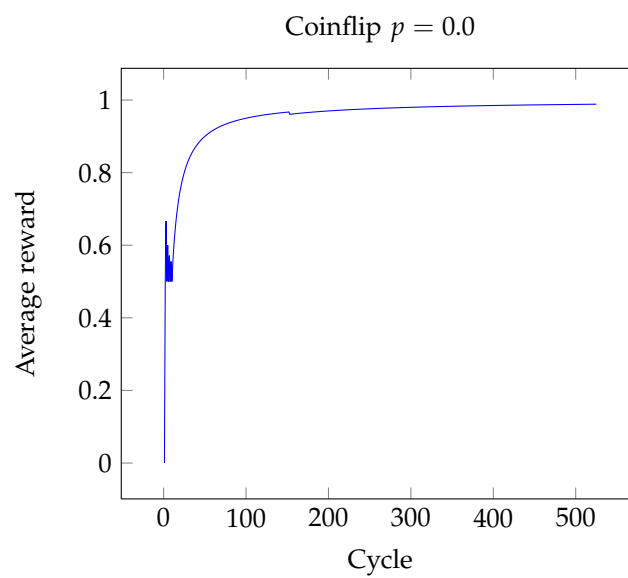
Chapter 4

Upper Confidence Tree

Chapter 5

Revert Function

5.1 Experimentation



5.2 Learning

5.3 Extension

Bibliography

- L. Kocsis and C. Szepesvári. Bandit based monte-carlo planning. *Machine Learning: ECML 2006*, pages 282–293, 2006.
- F.M.J. Willems, Y.M. Shtarkov, and T.J. Tjalkens. The context-tree weighting method: Basic properties. *Information Theory, IEEE Transactions on*, 41(3):653–664, 1995.