

E-HBA:
Using Action Policies for Expert Advice and Agent Typification
—
Appendix

Stefano V. Albrecht

The University of Edinburgh
Edinburgh, United Kingdom
s.v.albrecht@sms.ed.ac.uk

Jacob W. Crandall

Masdar Institute of Science and Technology
Abu Dhabi, United Arab Emirates
jcrandall@masdar.ac.ae

Subramanian Ramamoorthy

The University of Edinburgh
Edinburgh, United Kingdom
s.ramamoorthy@ed.ac.uk

This document is an appendix to [1]. It contains a listing of all parameter settings used in the study as well as plots for each experiment.

Contents

1	Parameter Settings	3
1.1	E-HBA	3
1.2	Leader-Follower-Trigger Agents	3
1.3	Co-Evolved Decision Trees / Neural Networks	3
2	Plots of Results	4
2.1	Legend	4
2.2	Results	4

List of Figures

1	Legend for bar plots	4
2	Player 2's type included in Θ_{2^*} (random type)	5
3	Player 2's type included in Θ_{2^*} (fictitious player)	6
4	Player 2's type <i>not</i> included in Θ_{2^*} (random type)	7
5	Player 2's type <i>not</i> included in Θ_{2^*} (fictitious player)	8

1 Parameter Settings

1.1 E-HBA

- 5 unique experts (i.e. $|\Phi_1| = 5$)
- 5 unique types (i.e. $|\Theta_2^*| = 5$)
- Depth of planning horizon: $h = 5$

1.2 Leader-Follower-Trigger Agents

- Maximum number of joint actions in target solutions: 2
- Target solution admissible if average payoff \geq maximin value (for each player)

1.3 Co-Evolved Decision Trees / Neural Networks

- Number of populations: 2 (one for each player)
- Individuals per population: 50 (first population randomly generated)
- Fitness = average payoff after 20 rounds ($\in [1, 4]$) – average similarity ($\in [0, 1]$)
- Each individual evaluated against random 40% of other population
- Resampling method: linear ranking
- Decision Trees:
 - Tree depth: 3 (up to three previous actions of other player)
 - Similarity: percentage of nodes with same action choice
 - Evolutions: 300 (evolution with highest average fitness used)
 - Random mutation of single node (flipping action): 5% of population
 - Random crossing of sub-trees (preserving tree depth): 30% of population
- Neural Networks:
 - Input layer: 4 nodes (up to two previous joint actions)
 - Hidden layer: 5 nodes
 - Output layer: 1 node (probability of action 1)
 - Each node fully-connected with nodes of next layer
 - Standard sigmoidal threshold function
 - Similarity: 1 – average difference of output for each input
 - Evolutions: 1000 (evolution with highest average fitness used)
 - Random mutation of single input weight (standard normal shift): 20% of population
 - Random crossing of nodes (preserving network structure): 10% of population

2 Plots of Results

2.1 Legend

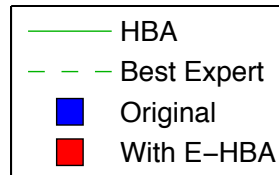
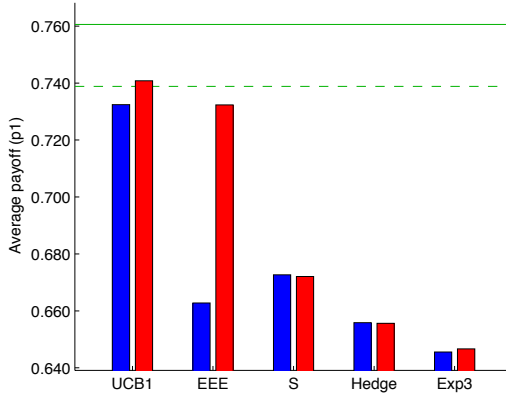


Figure 1: **Legend for bar plots.** The bar plots show the average payoffs of player 1 (E-HBA) over 5,000 rounds. The minimum and maximum achievable payoffs per round were 0 and 1, respectively.

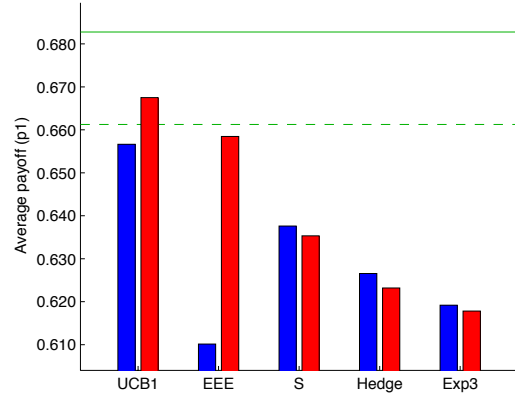
2.2 Results

We use the following abbreviations:

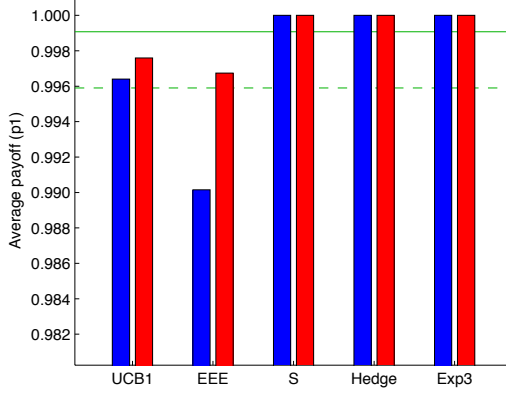
- LFT – Leader-Follower-Trigger Agents
- CDT – Co-Evolved Decision Trees
- CNN – Co-Evolved Neural Networks



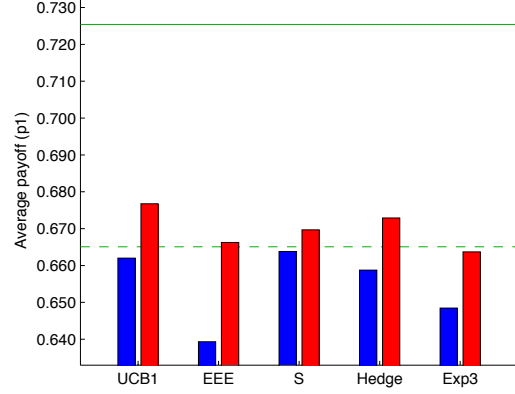
(a) LFT – No-Conflict



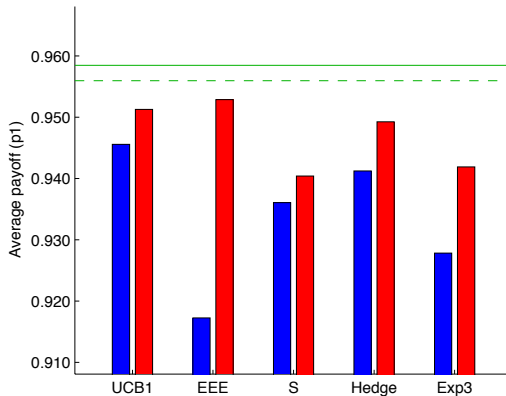
(b) LFT – Conflict



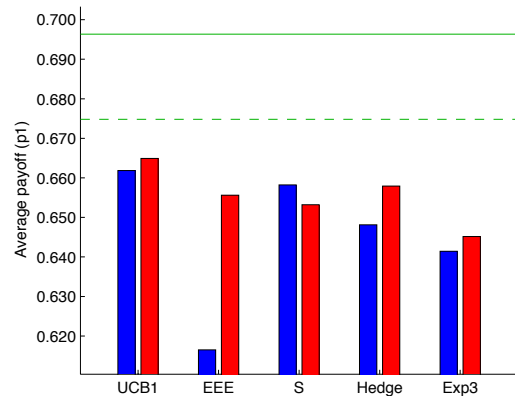
(c) CDT – No-Conflict



(d) CDT – Conflict

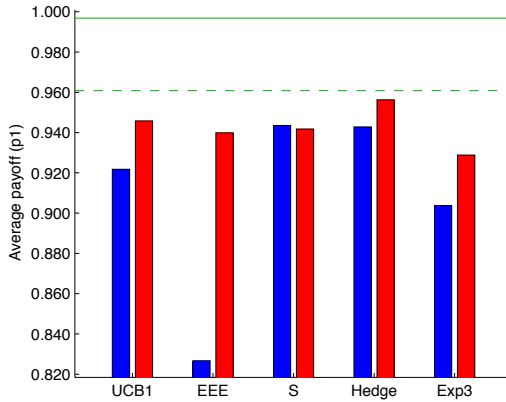


(e) CNN – No-Conflict

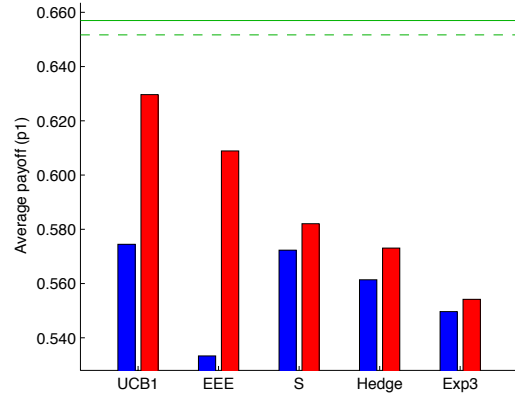


(f) CNN – Conflict

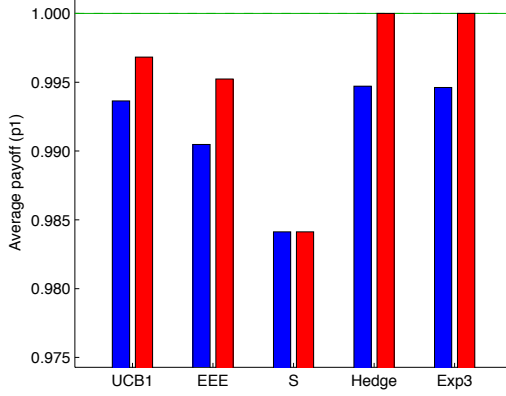
Figure 2: **Player 2's type included in Θ_2^* . Player 2 controlled by random type.**



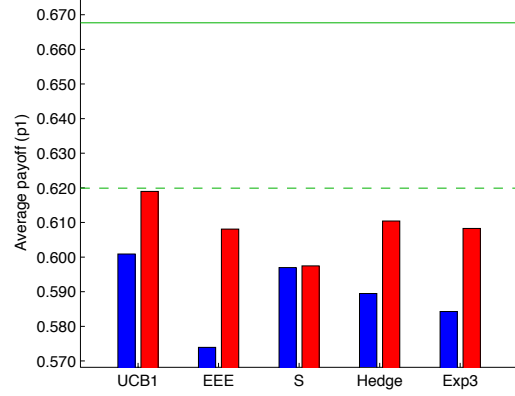
(a) LFT – No-Conflict



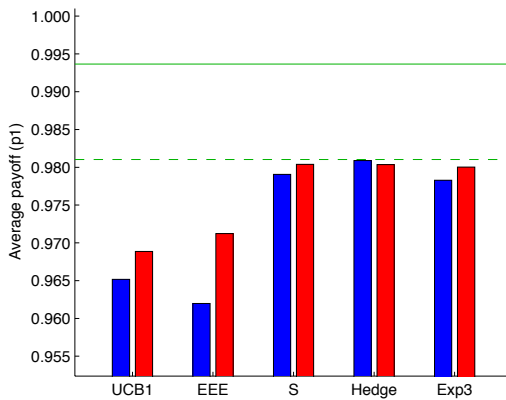
(b) LFT – Conflict



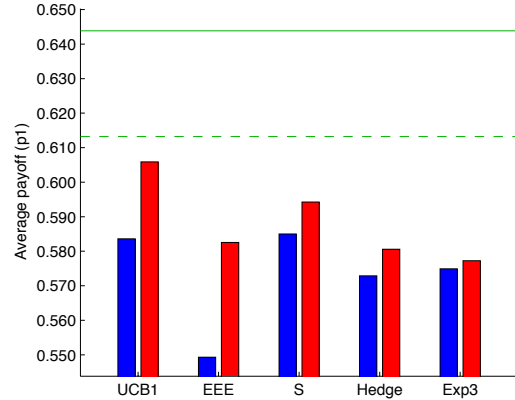
(c) CDT – No-Conflict



(d) CDT – Conflict

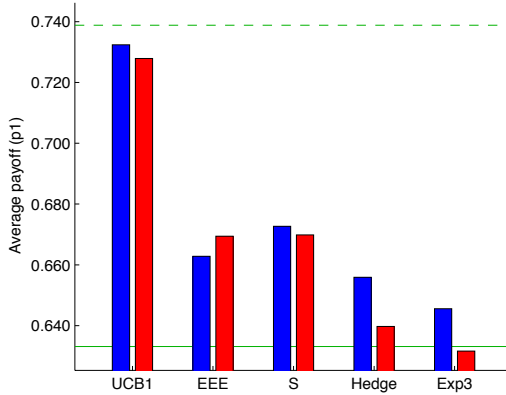


(e) CNN – No-Conflict

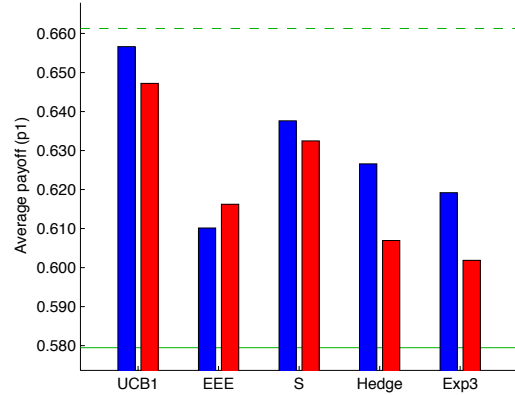


(f) CNN – Conflict

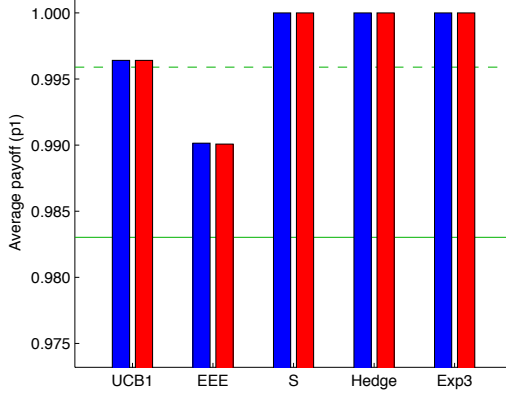
Figure 3: **Player 2's type included in Θ_2^* .** Player 2 controlled by fictitious player.



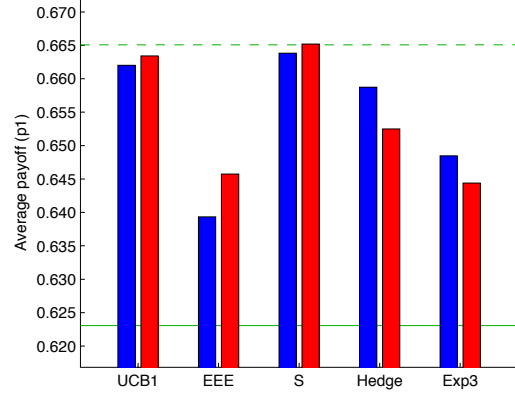
(a) LFT – No-Conflict



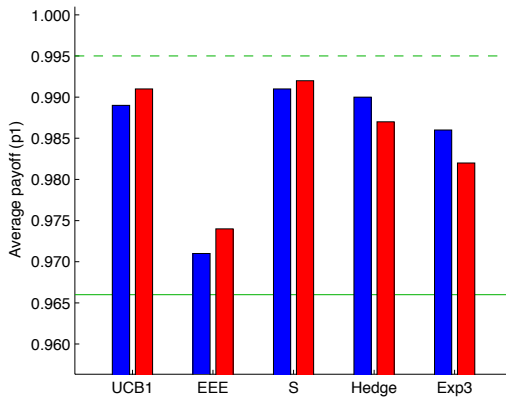
(b) LFT – Conflict



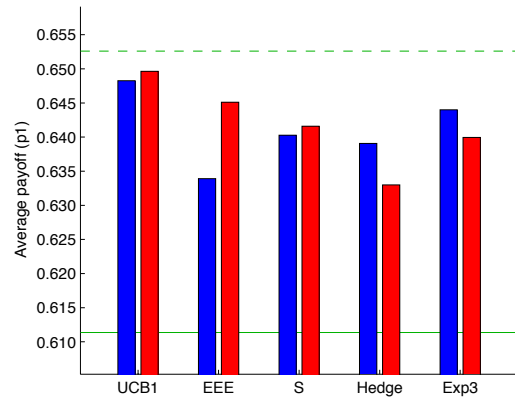
(c) CDT – No-Conflict



(d) CDT – Conflict

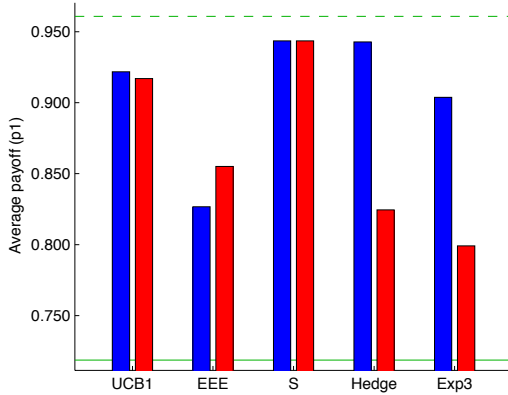


(e) CNN – No-Conflict

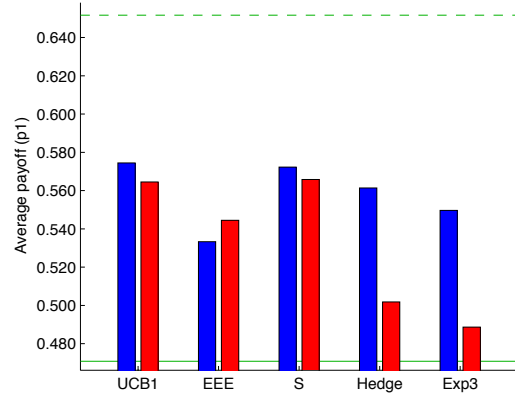


(f) CNN – Conflict

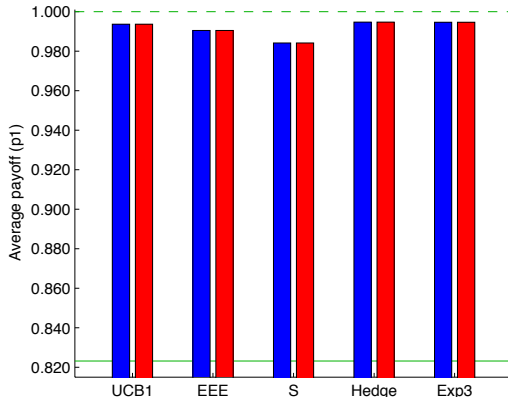
Figure 4: **Player 2's type not included in Θ_2^* .** Player 2 controlled by random type.



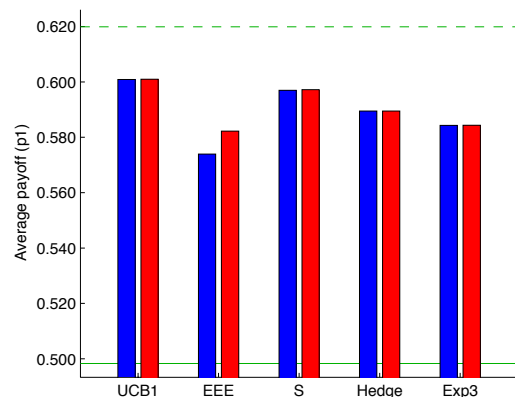
(a) LFT – No-Conflict



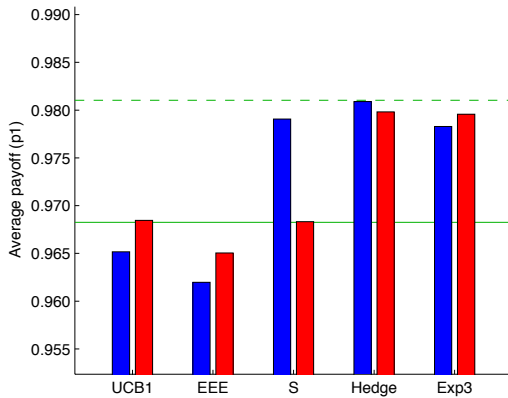
(b) LFT – Conflict



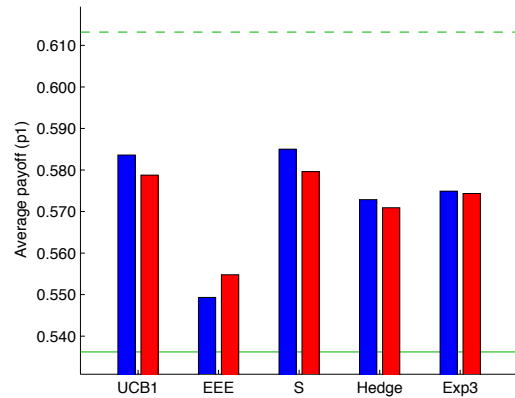
(c) CDT – No-Conflict



(d) CDT – Conflict



(e) CNN – No-Conflict



(f) CNN – Conflict

Figure 5: **Player 2's type *not* included in Θ_2^* . Player 2 controlled by fictitious player.**

References

- [1] S. V. Albrecht, J. W. Crandall, and S. Ramamoorthy. E-HBA: Using action policies for expert advice and agent typification. In *Proceedings of the Second AAI-Workshop on Multiagent Interaction without Prior Coordination*, Austin, Texas, USA, January 2015.