An Empirical Study on the Practical Impact of Prior Beliefs over Policy Types

—

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 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 Colour Codes . . . 4
2.2 Results . . . 4

List of Figures

1 Colour codes for heat matrices . . . 4
2 Colour codes for payoff plots . . . 4
3 Leader-Follower-Trigger Agents – Random Type – No-Conflict Games . . . 5
4 Leader-Follower-Trigger Agents – Random Type – Conflict Games . . . 6
5 Leader-Follower-Trigger Agents – Fictitious Player – No-Conflict Games . . . 7
6 Leader-Follower-Trigger Agents – Fictitious Player – Conflict Games . . . 8
7 Leader-Follower-Trigger Agents – Conditioned Fictitious Player – No-Conflict Games 9
8 Leader-Follower-Trigger Agents – Conditioned Fictitious Player – Conflict Games . . 10
9 Co-Evolved Decision Trees – Random Type – No-Conflict Games . . . 11
10 Co-Evolved Decision Trees – Random Type – Conflict Games . . . 12
11 Co-Evolved Decision Trees – Fictitious Player – No-Conflict Games . . . 13
12 Co-Evolved Decision Trees – Fictitious Player – Conflict Games . . . 14
13 Co-Evolved Decision Trees – Conditioned Fictitious Player – No-Conflict Games . . 15
14 Co-Evolved Decision Trees – Conditioned Fictitious Player – Conflict Games 16
15 Co-Evolved Neural Networks – Random Type – No-Conflict Games . . . 17
16 Co-Evolved Neural Networks – Random Type – Conflict Games . . . 18
17 Co-Evolved Neural Networks – Fictitious Player – No-Conflict Games . . . 19
18 Co-Evolved Neural Networks – Fictitious Player – Conflict Games . . . 20
19 Co-Evolved Neural Networks – Conditioned Fictitious Player – No-Conflict Games . 21
20 Co-Evolved Neural Networks – Conditioned Fictitious Player – Conflict Games . . . 22

1 Parameter Settings

1.1 HBA

- 10 hypothesised types for player 2 (i.e. $|\Theta_2^*| = 10$)
- True type of player 2 always included in Θ_2^*
- Value/LP priors: $t = 5$ and $b = 10$

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 Colour Codes

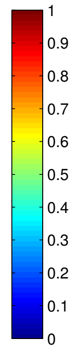


Figure 1: **Colour codes for heat matrices.** Performance criteria are measured over *time slices*, which are a temporal partitioning of a play. Each element (r, c) in the heat matrix corresponds to the percentage of time slices in which the prior belief r produced significantly higher values for the criterion c than the Uniform prior, averaged over all plays in all tested games. (Note: binary criteria such as convergence and Nash equilibrium assume 0/1-values in each time slice.) All significance statements are based on paired right-sided t-tests with a 5% significance level.

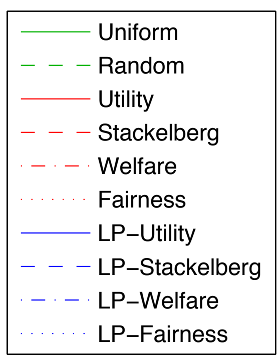
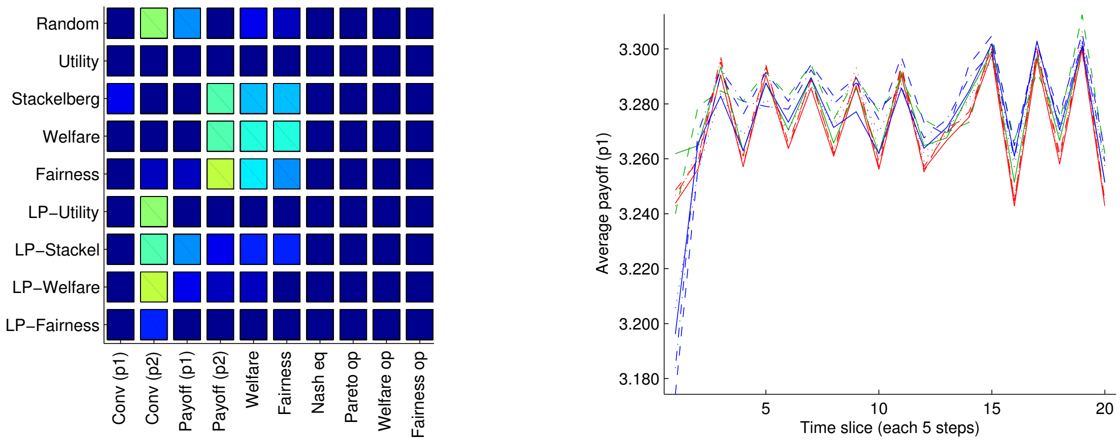


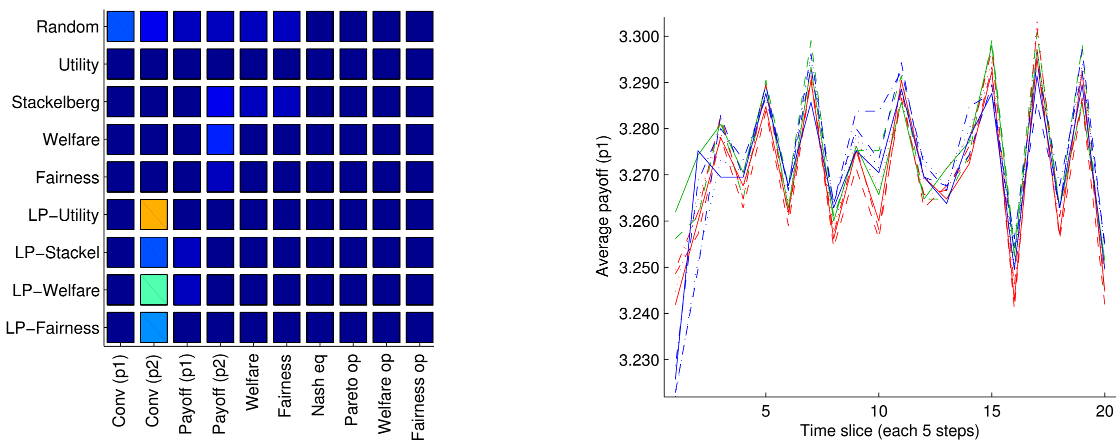
Figure 2: **Colour codes for payoff plots.** The payoff plots show the average payoffs of player 1 (i.e. HBA) in each time slice. Every play against a random type lasted 100 rounds and was partitioned into 20 time slices, consisting of 5 consecutive time steps each. Every play against a (conditioned) fictitious player lasted 10,000 rounds and was partitioned into 100 time slices, consisting of 100 consecutive time steps each. The minimum and maximum achievable payoffs per round were 1 and 4, respectively.

2.2 Results

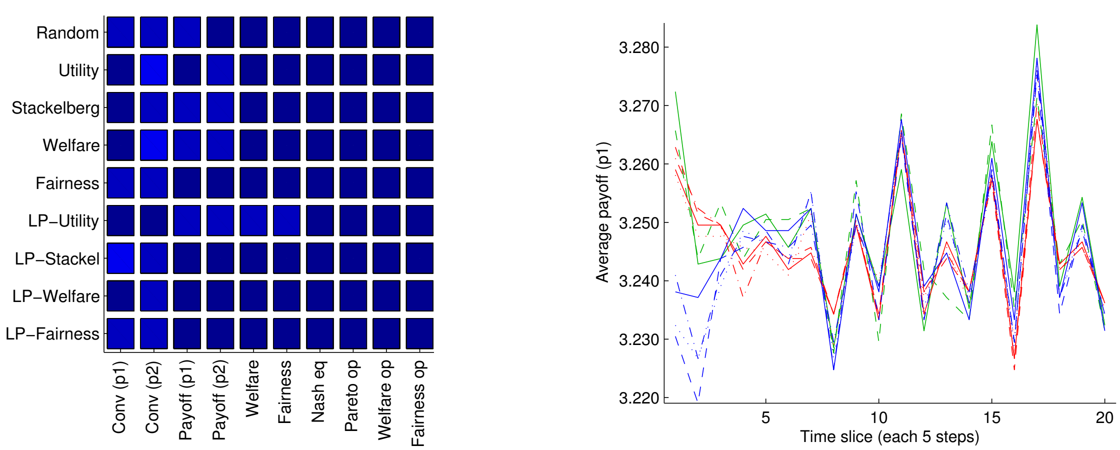
Note: A planning horizon of h means that HBA made predictions for the next h actions of player 2.



(a) $h = 5$

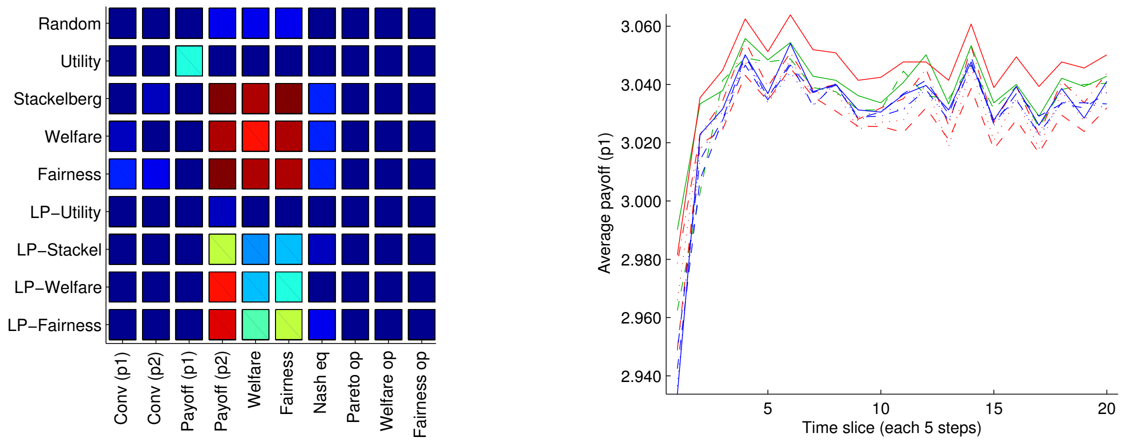


(b) $h = 3$

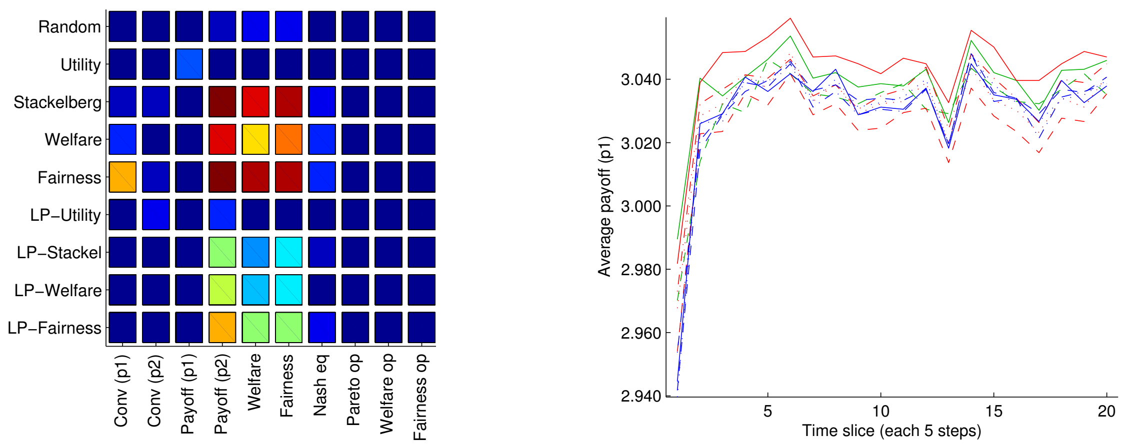


(c) $h = 1$

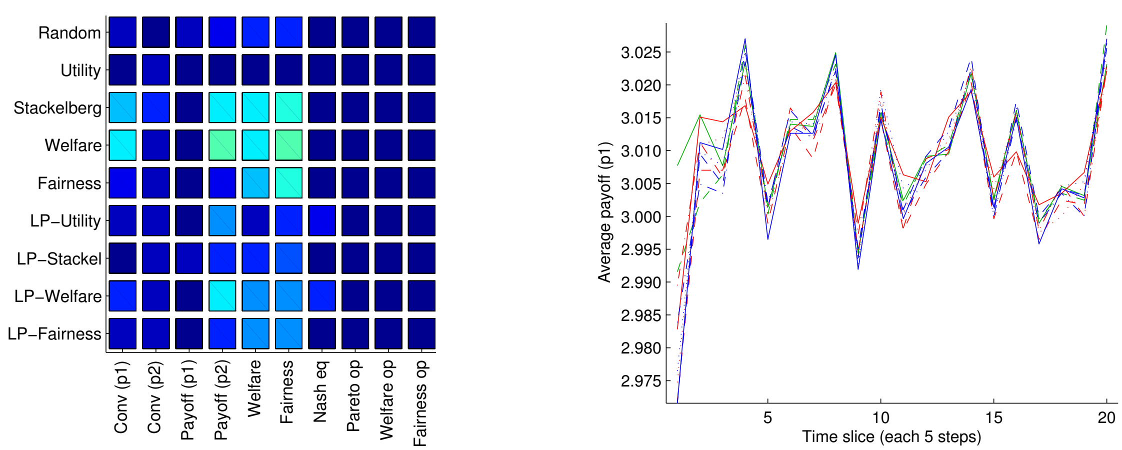
Figure 3: Player 1 controlled by HBA using **leader-follower-trigger agents** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **random type**. Results from **no-conflict games**.



(a) $h = 5$

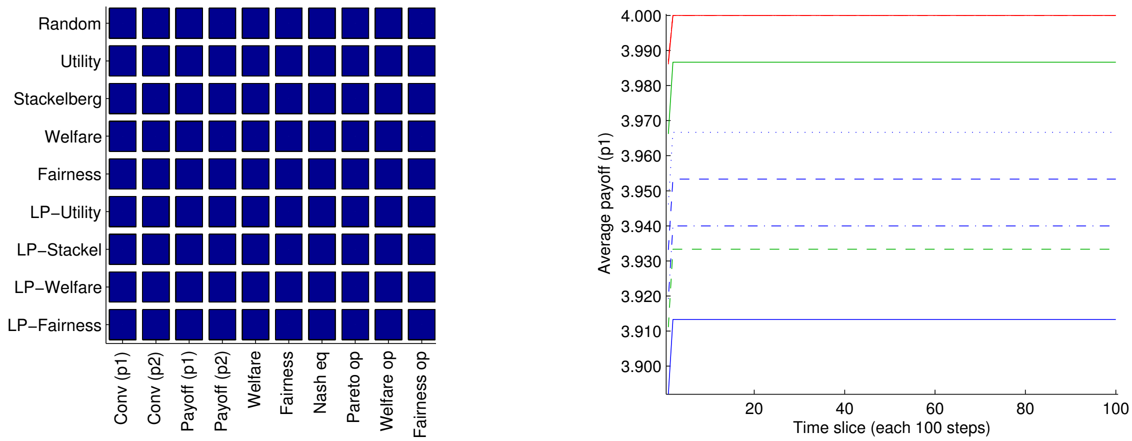


(b) $h = 3$

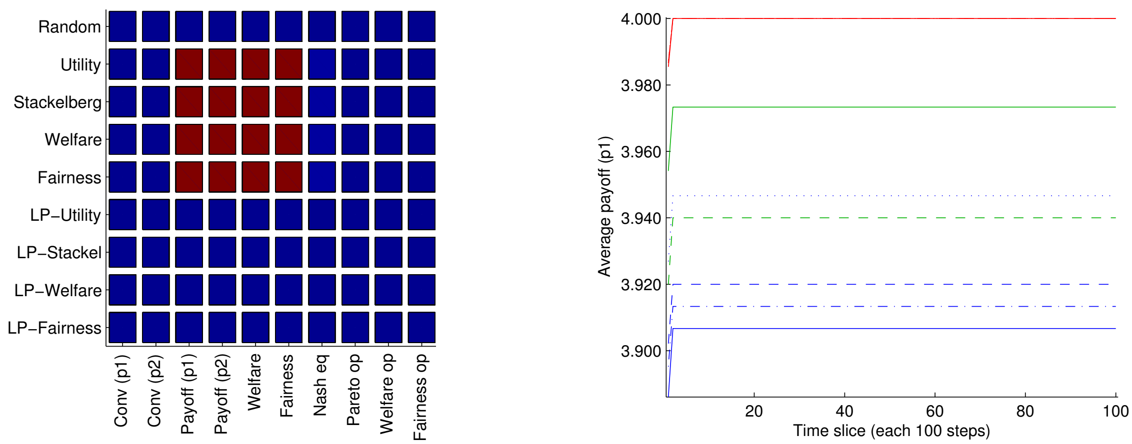


(c) $h = 1$

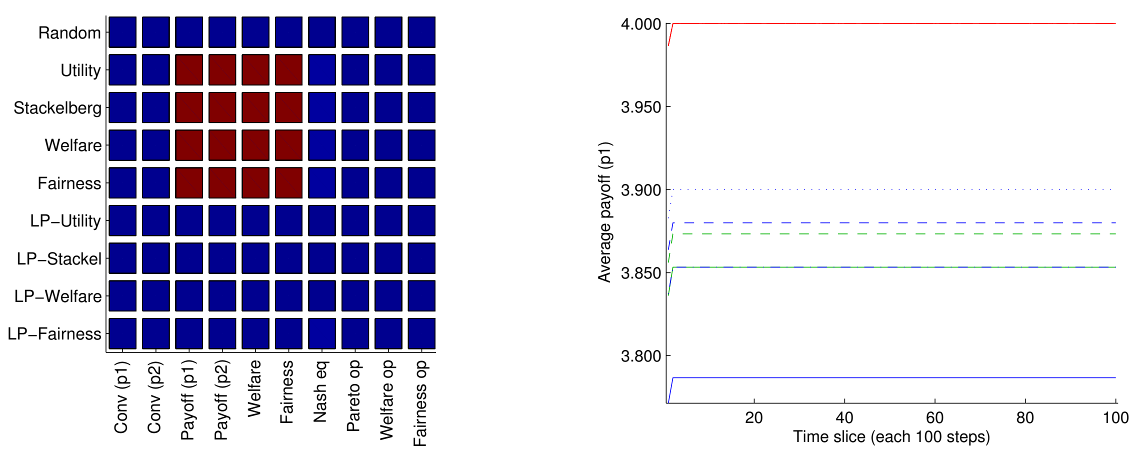
Figure 4: Player 1 controlled by HBA using **leader-follower-trigger agents** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **random type**. Results from **conflict games**.



(a) $h = 5$

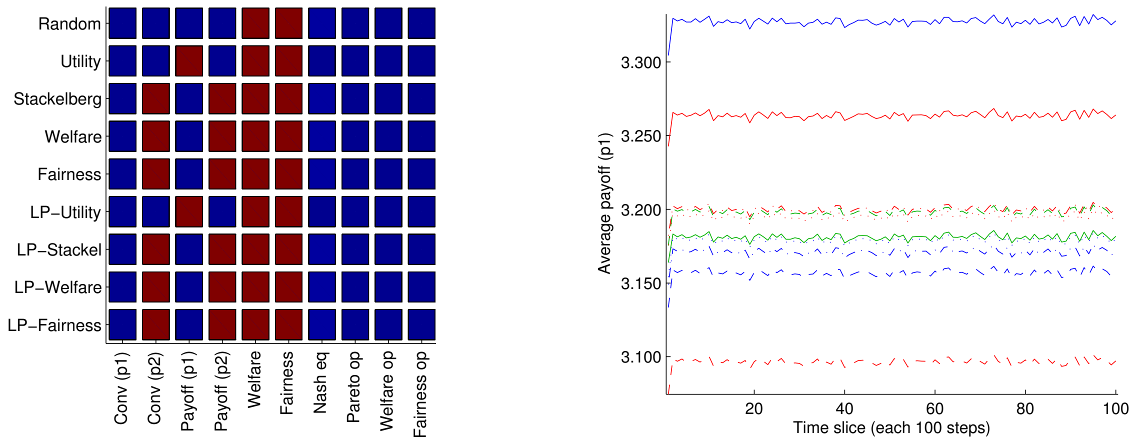


(b) $h = 3$

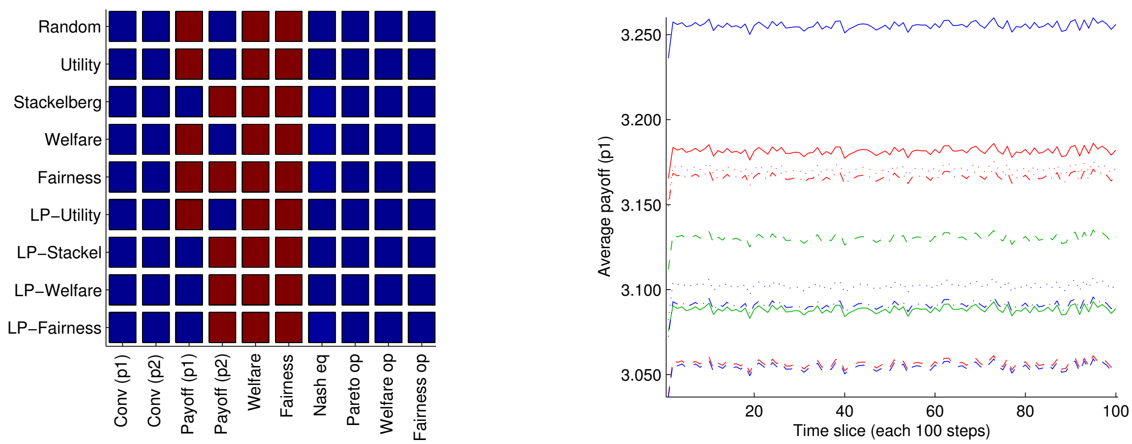


(c) $h = 1$

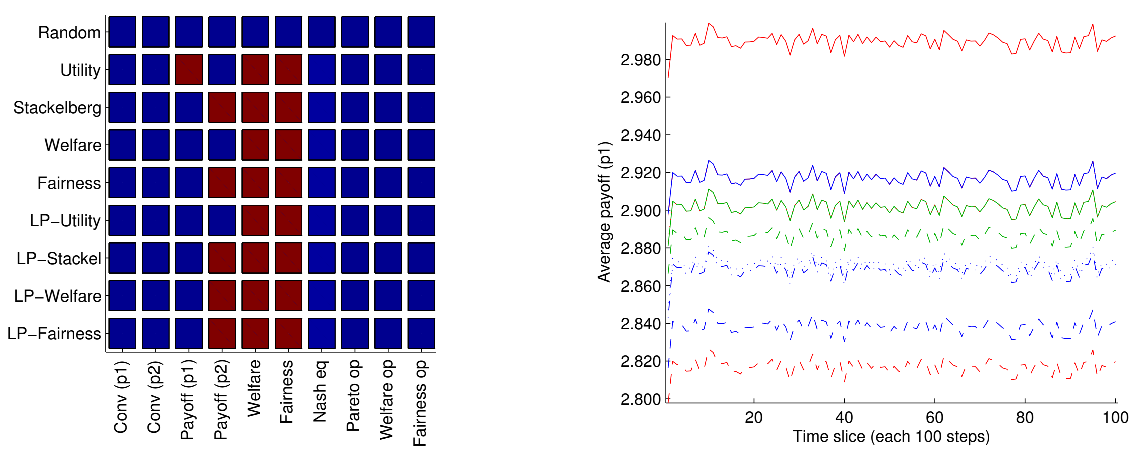
Figure 5: Player 1 controlled by HBA using **leader-follower-trigger agents** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **fictitious player**. Results from **no-conflict games**.



(a) $h = 5$

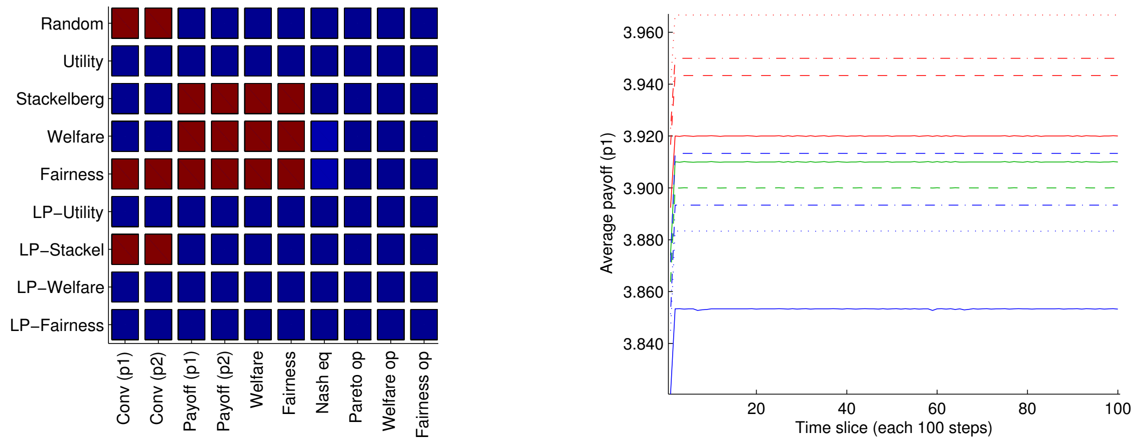


(b) $h = 3$

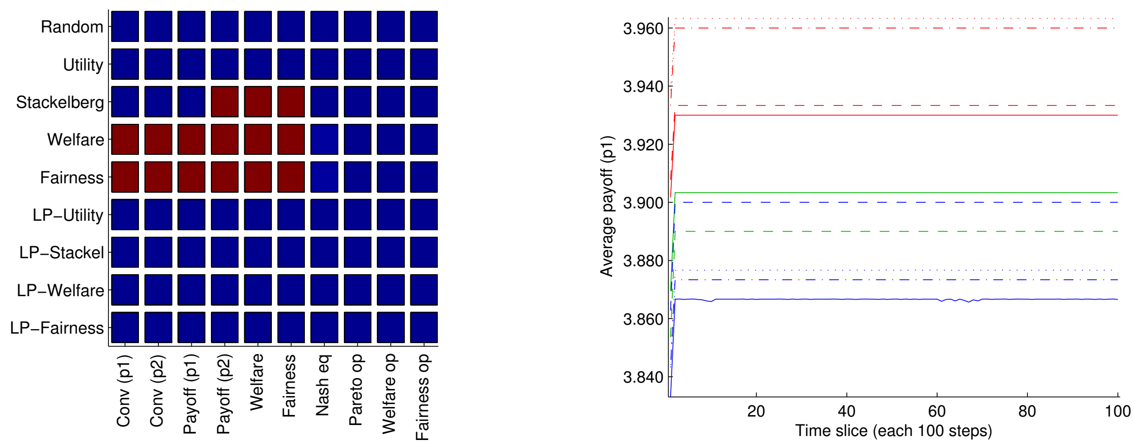


(c) $h = 1$

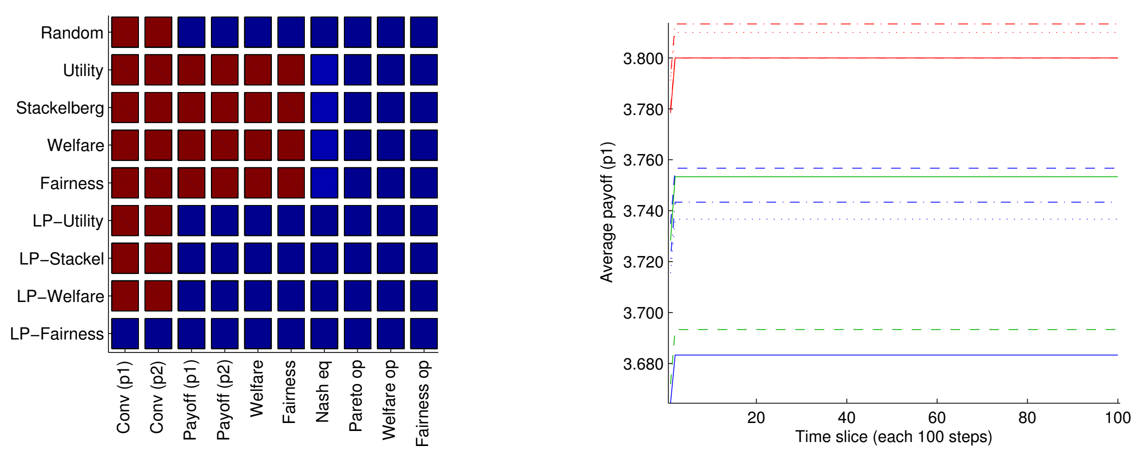
Figure 6: Player 1 controlled by HBA using **leader-follower-trigger agents** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **fictitious player**. Results from **conflict games**.



(a) $h = 5$

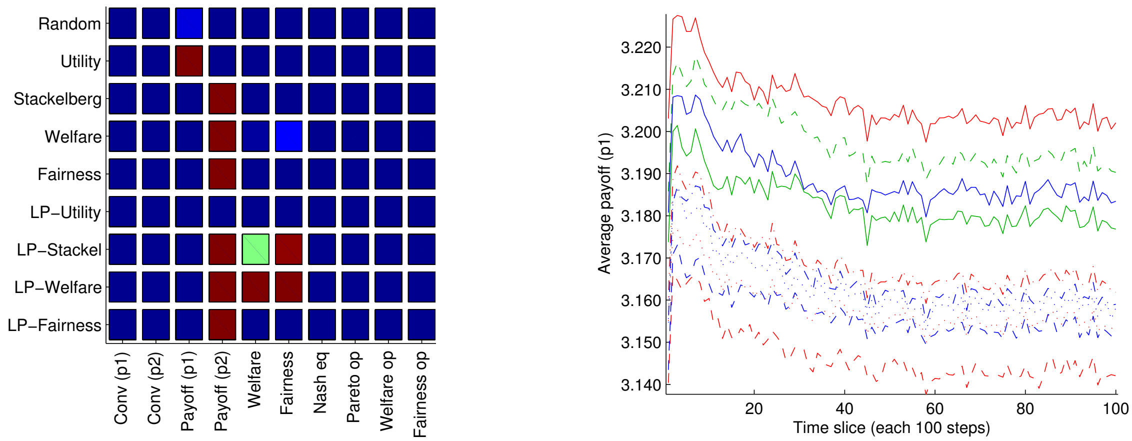


(b) $h = 3$

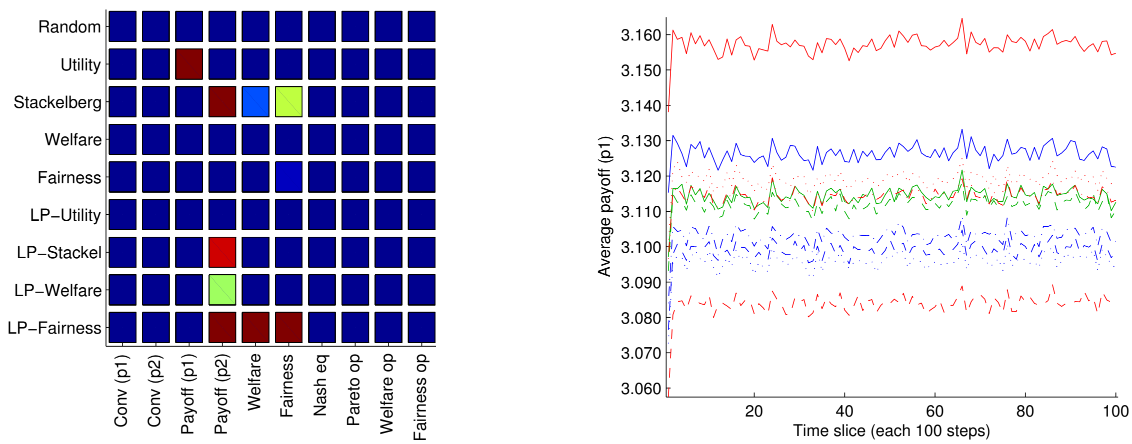


(c) $h = 1$

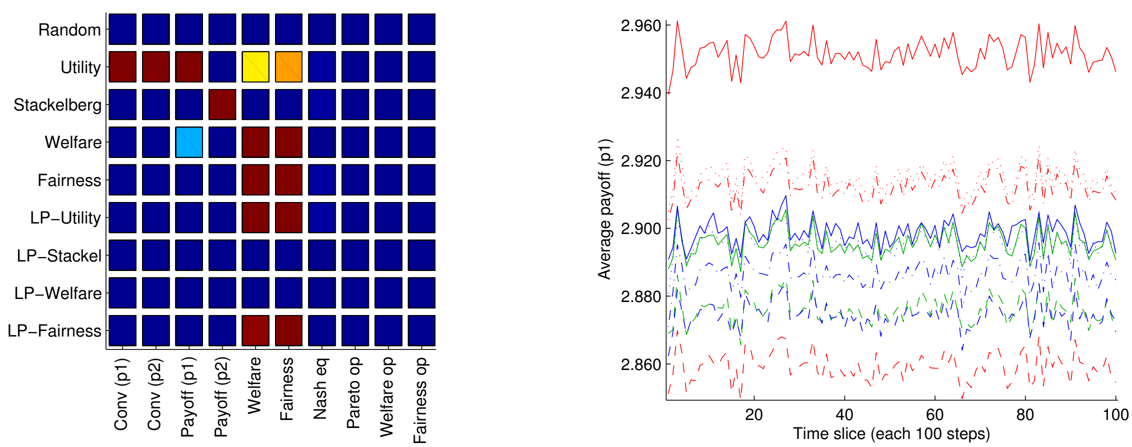
Figure 7: Player 1 controlled by HBA using **leader-follower-trigger agents** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **conditioned fictitious player**. Results from **no-conflict games**.



(a) $h = 5$

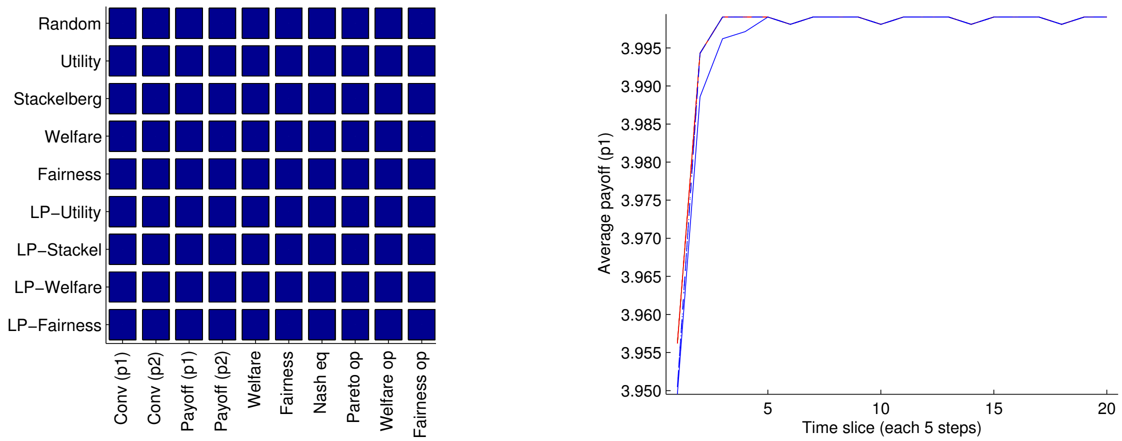


(b) $h = 3$

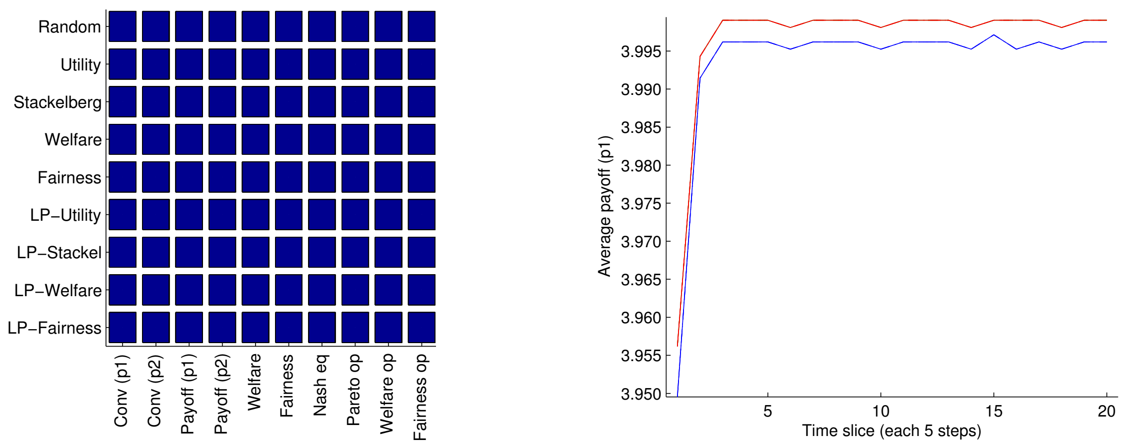


(c) $h = 1$

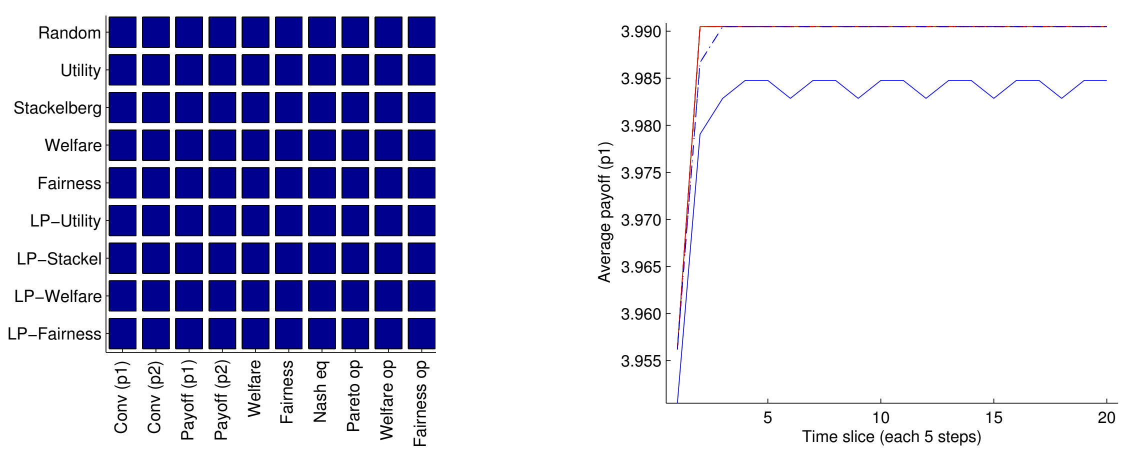
Figure 8: Player 1 controlled by HBA using **leader-follower-trigger agents** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **conditioned fictitious player**. Results from **conflict games**.



(a) $h = 5$

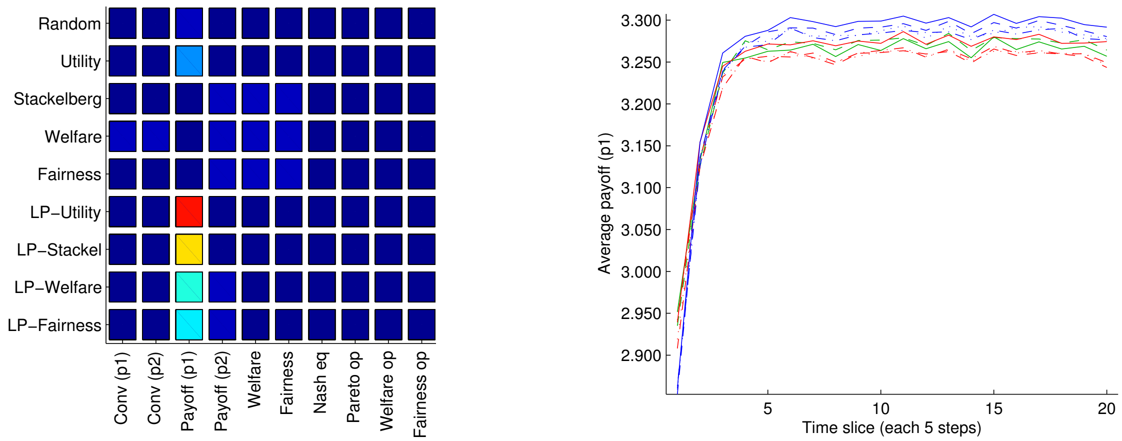


(b) $h = 3$

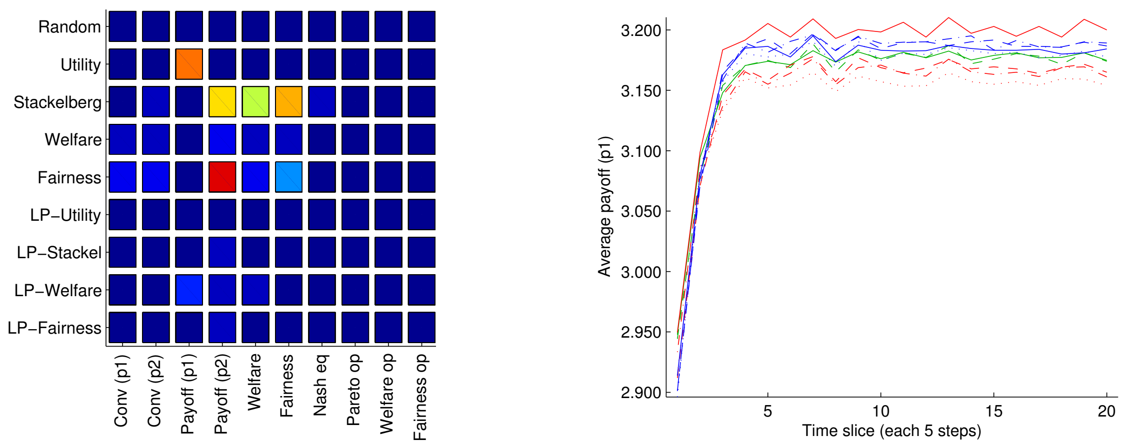


(c) $h = 1$

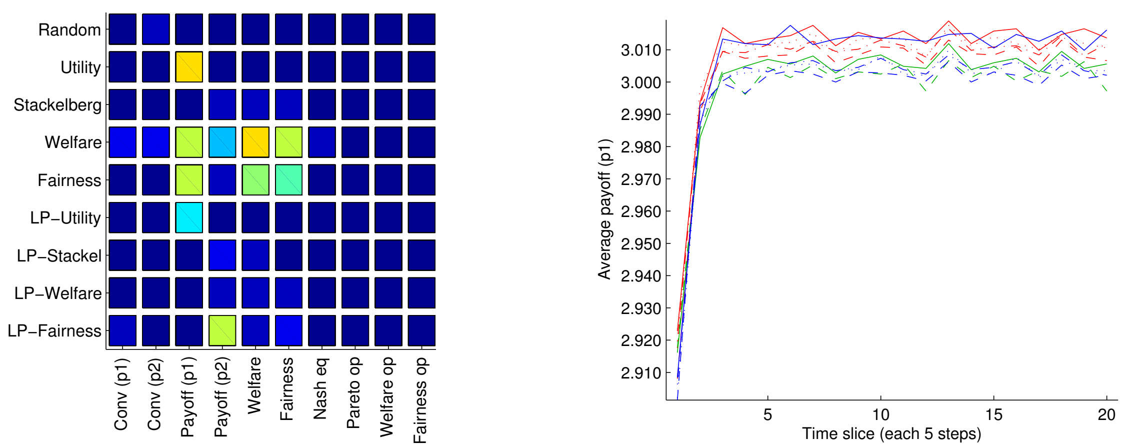
Figure 9: Player 1 controlled by HBA using **co-evolved decision trees** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **random type**. Results from **no-conflict games**.



(a) $h = 5$

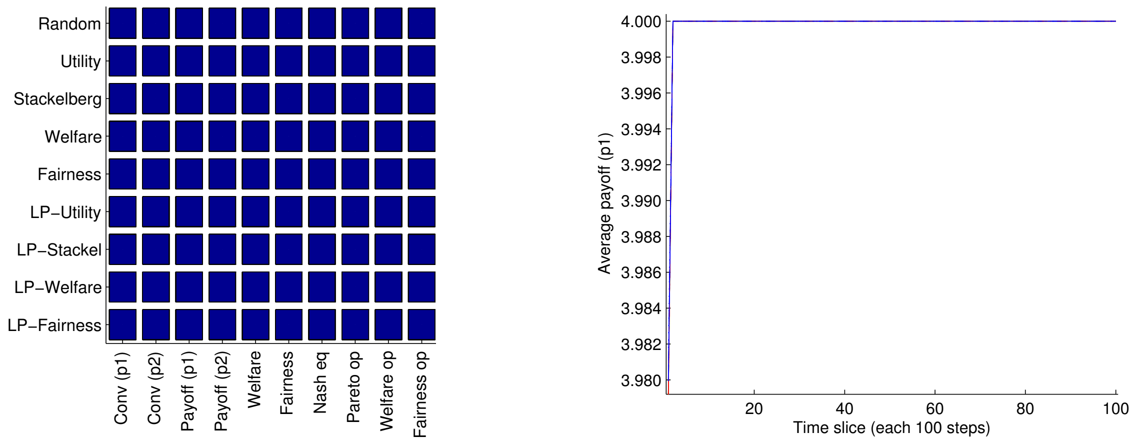


(b) $h = 3$

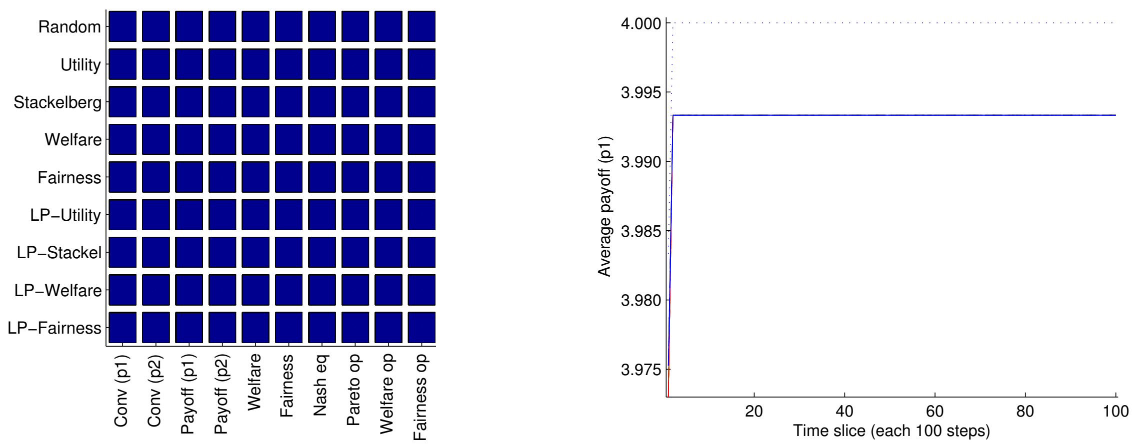


(c) $h = 1$

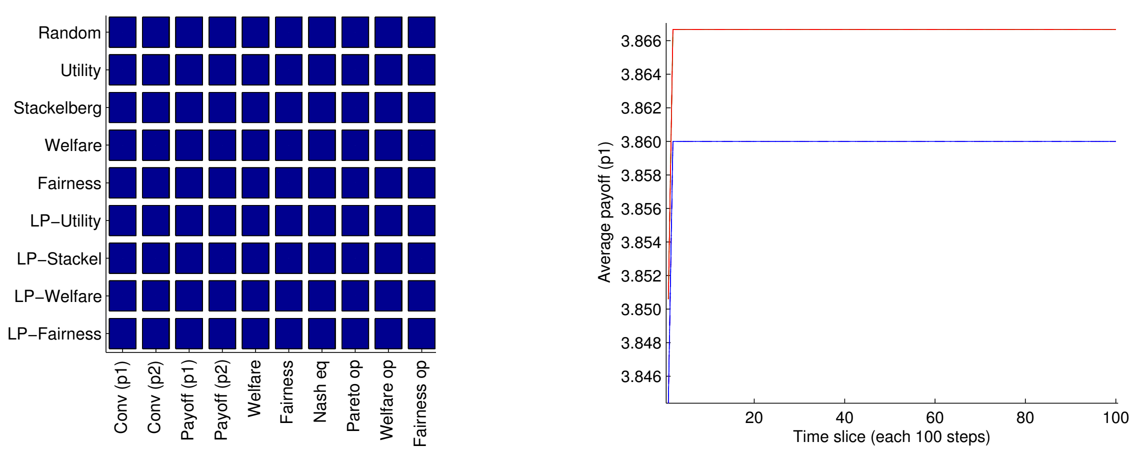
Figure 10: Player 1 controlled by HBA using **co-evolved decision trees** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **random type**. Results from **conflict games**.



(a) $h = 5$

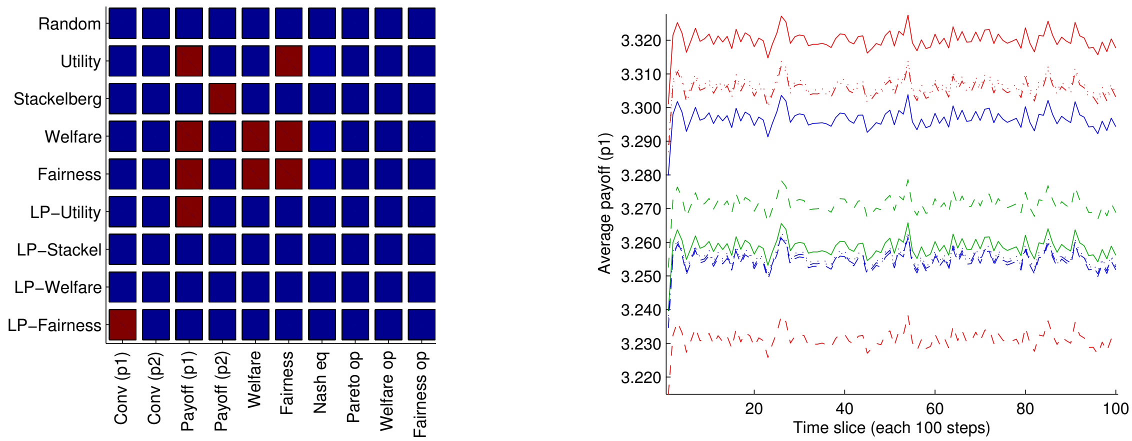


(b) $h = 3$

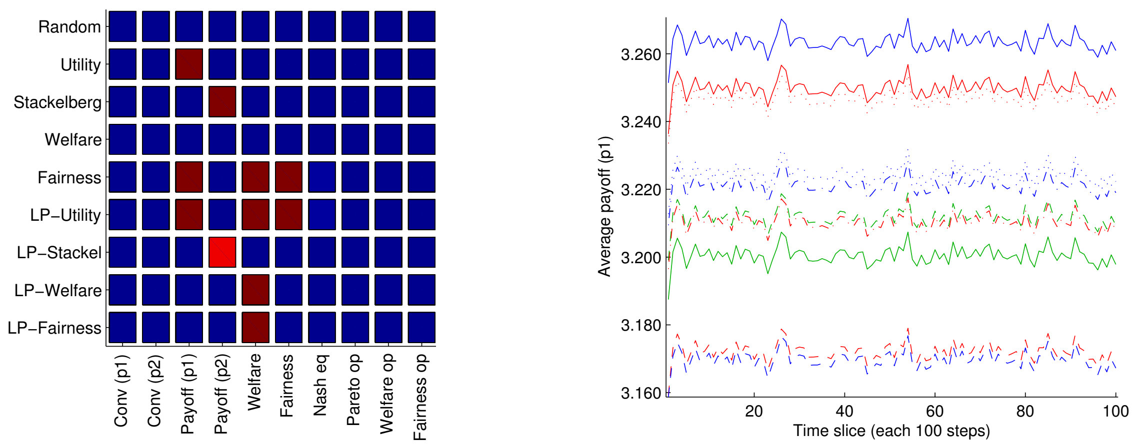


(c) $h = 1$

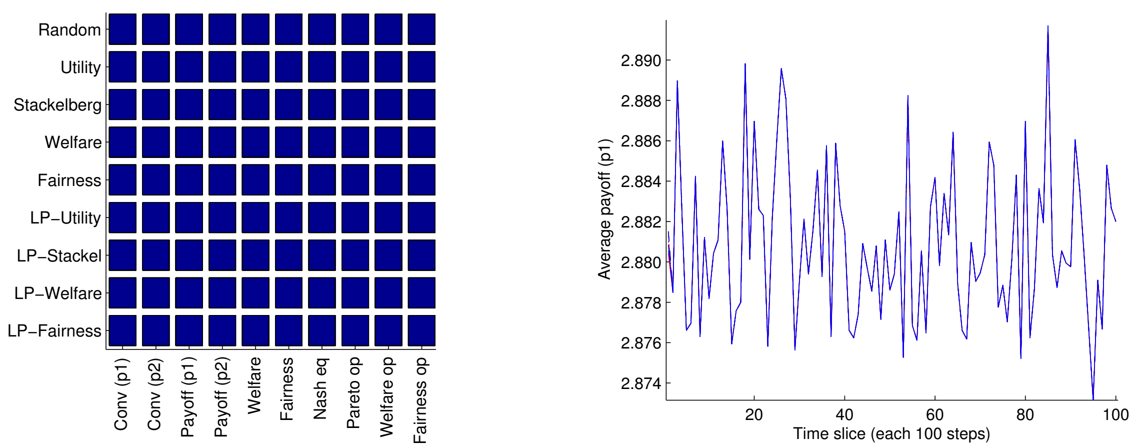
Figure 11: Player 1 controlled by HBA using **co-evolved decision trees** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **fictitious player**. Results from **no-conflict games**.



(a) $h = 5$

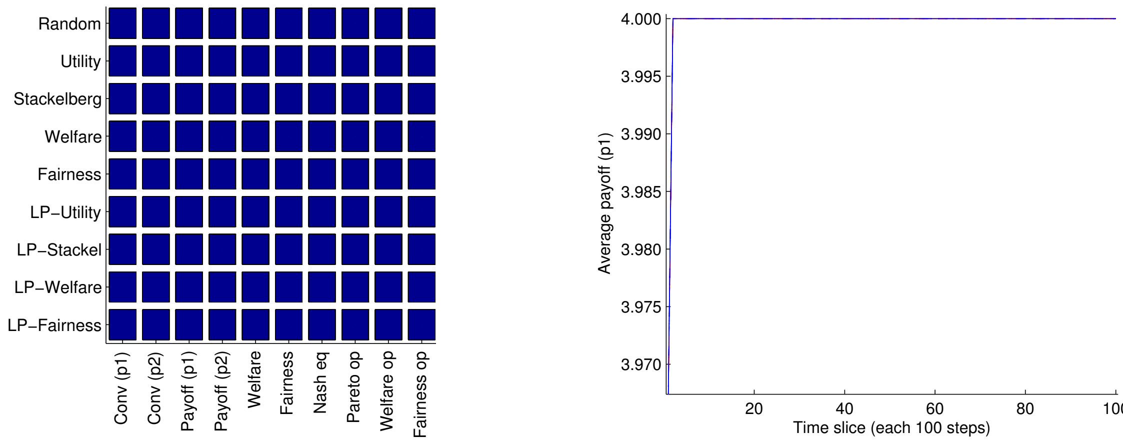


(b) $h = 3$

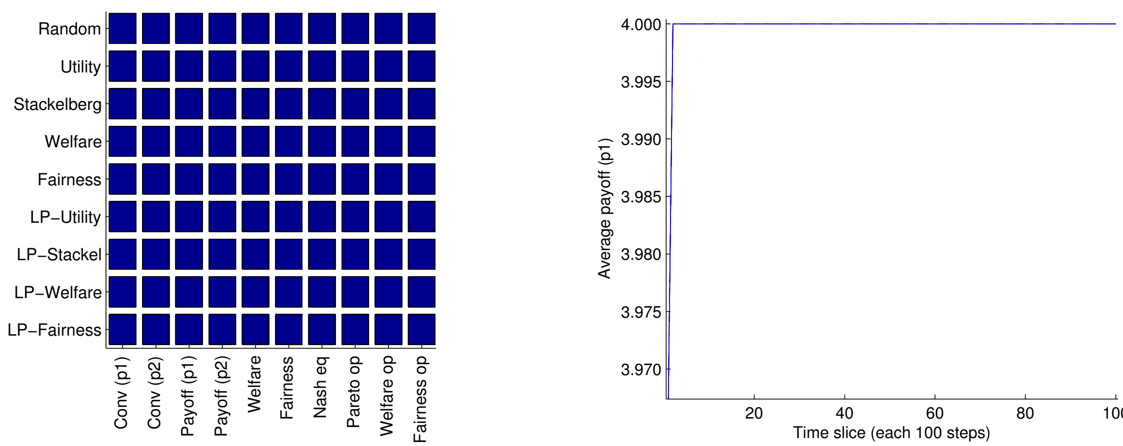


(c) $h = 1$

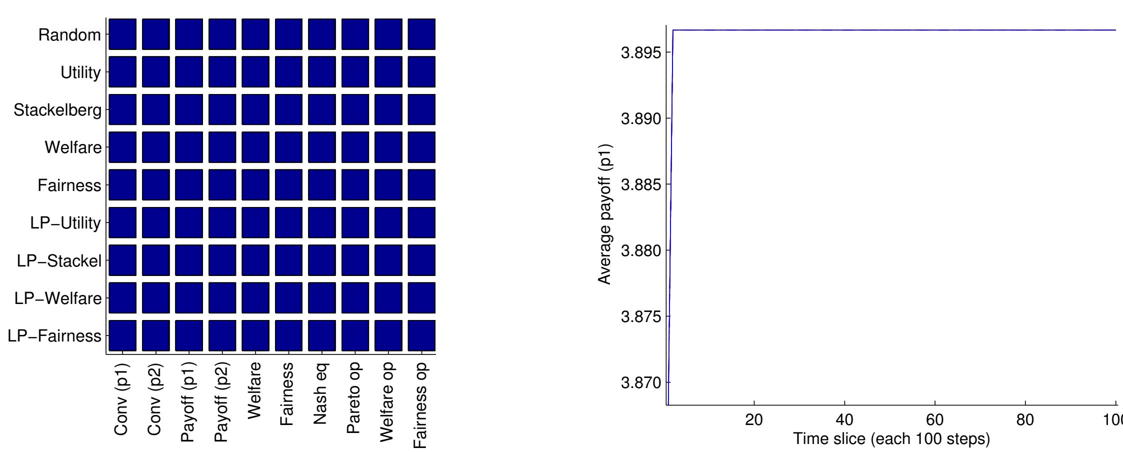
Figure 12: Player 1 controlled by HBA using **co-evolved decision trees** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **fictitious player**. Results from **conflict games**.



(a) $h = 5$

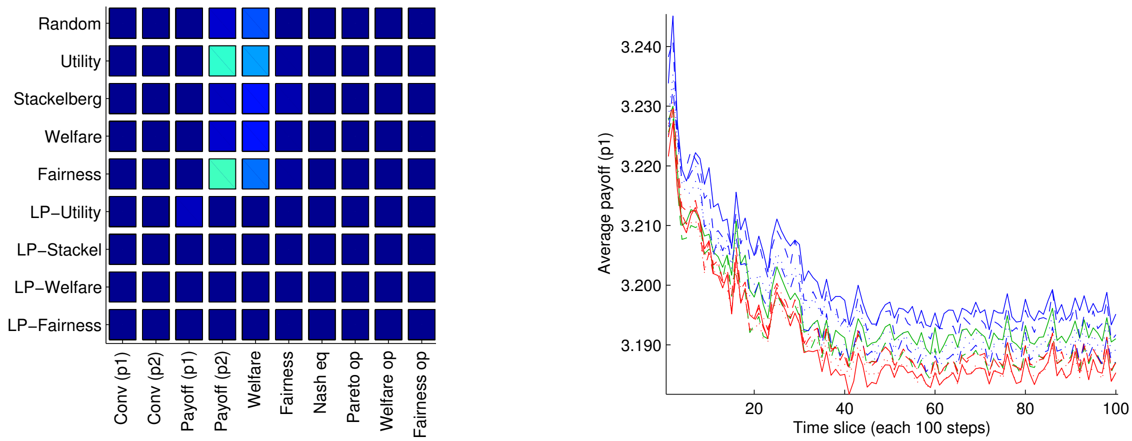


(b) $h = 3$

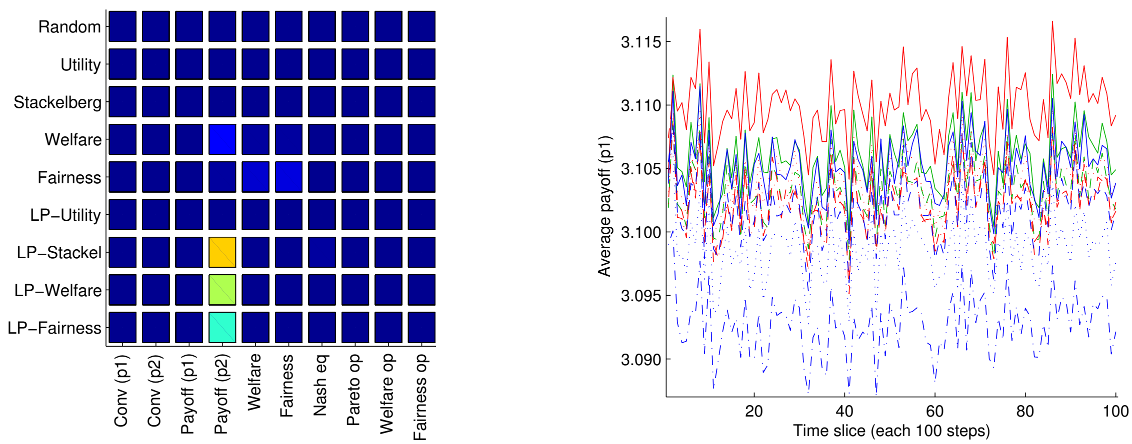


(c) $h = 1$

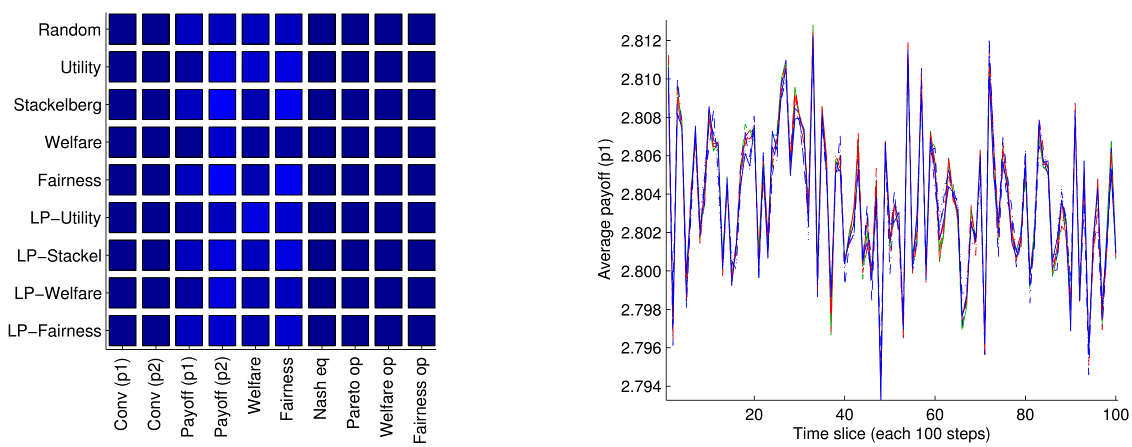
Figure 13: Player 1 controlled by HBA using **co-evolved decision trees** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **conditioned fictitious player**. Results from **no-conflict games**.



(a) $h = 5$

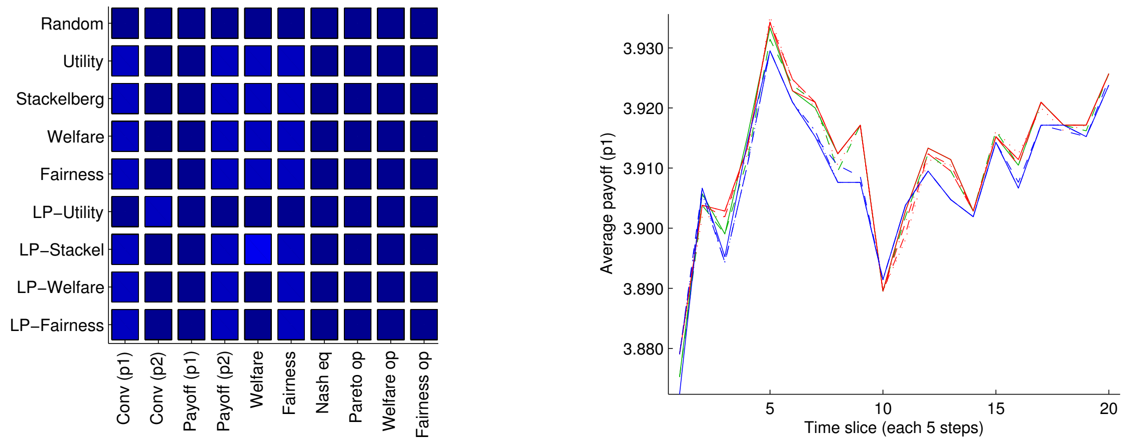


(b) $h = 3$

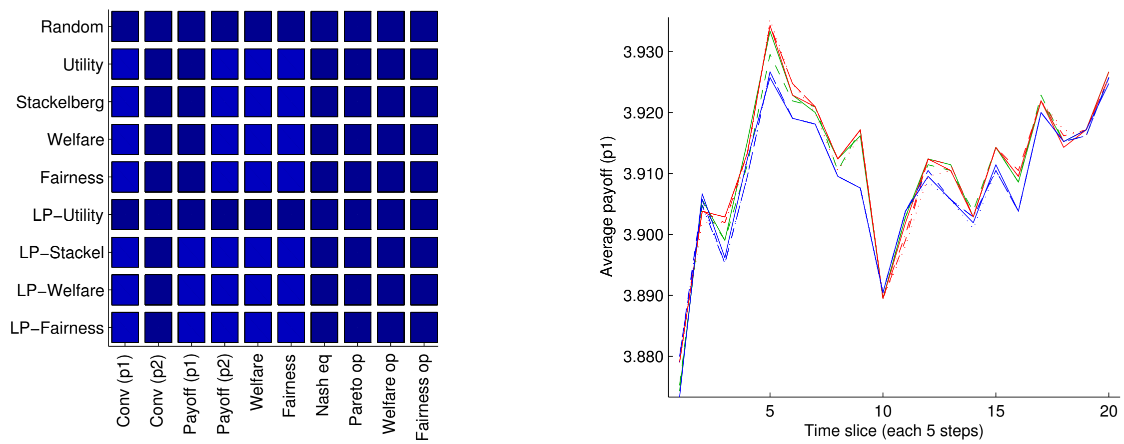


(c) $h = 1$

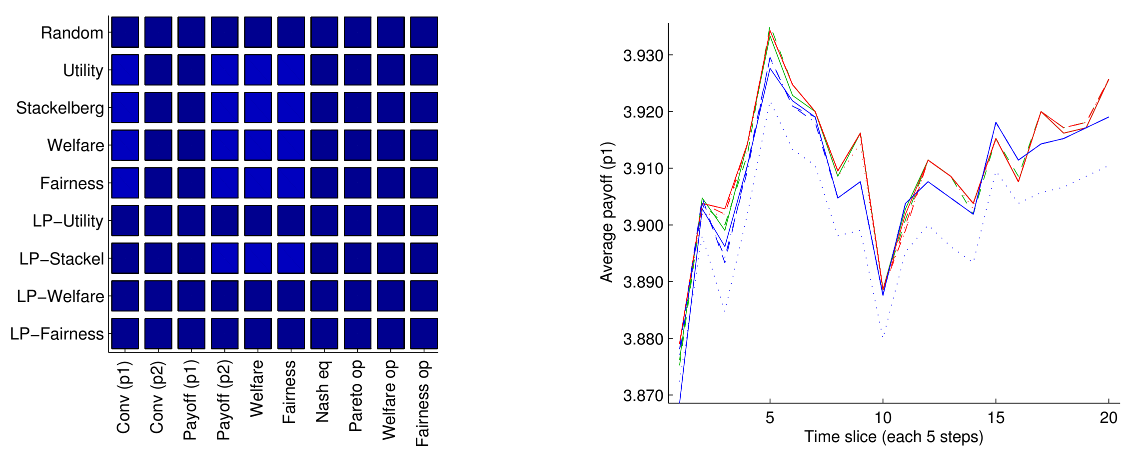
Figure 14: Player 1 controlled by HBA using **co-evolved decision trees** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **conditioned fictitious player**. Results from **conflict games**.



(a) $h = 5$

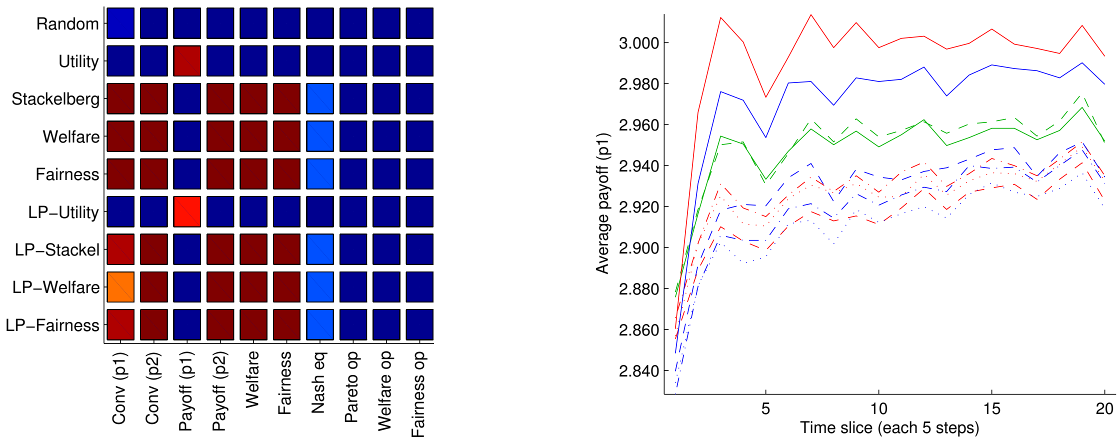


(b) $h = 3$

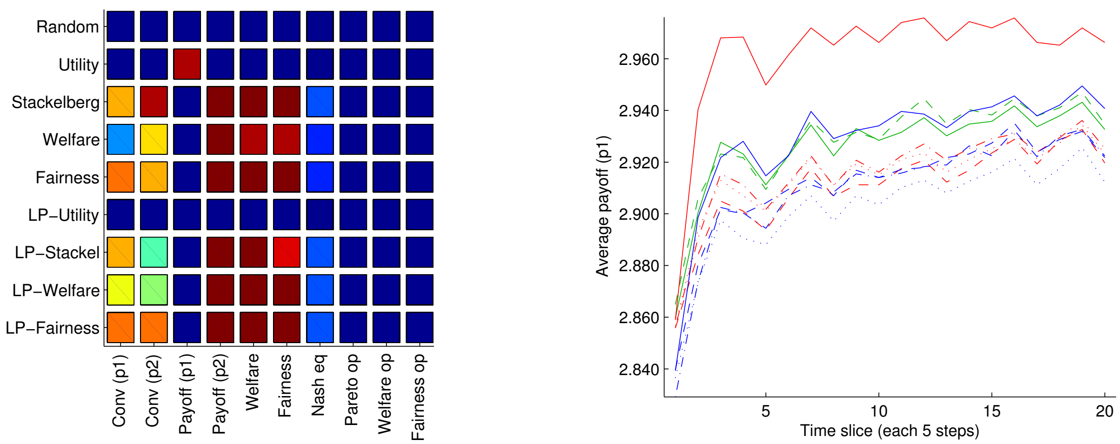


(c) $h = 1$

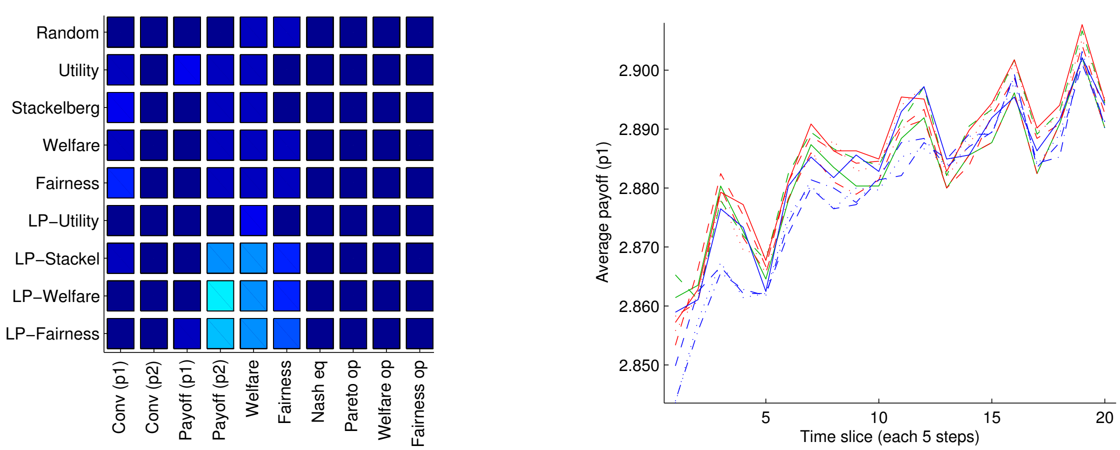
Figure 15: Player 1 controlled by HBA using **co-evolved neural networks** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **random type**. Results from **no-conflict games**.



(a) $h = 5$

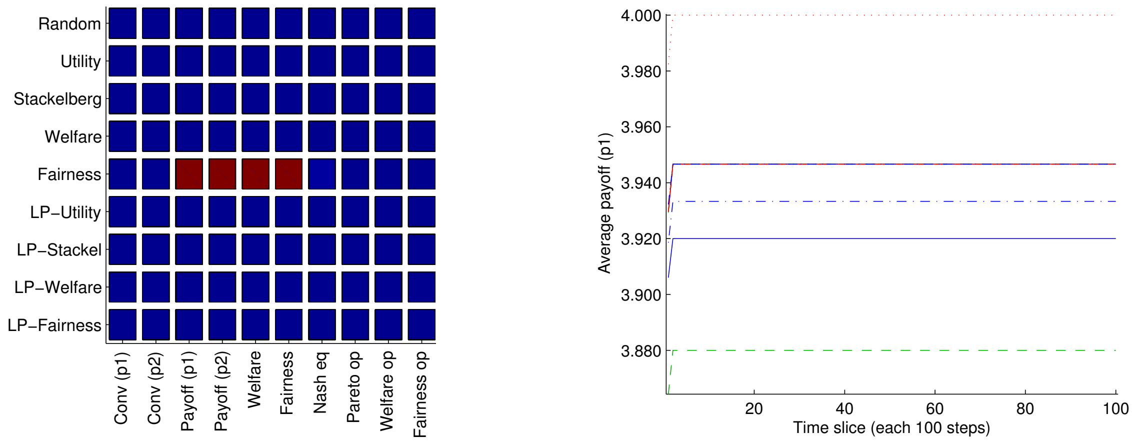


(b) $h = 3$

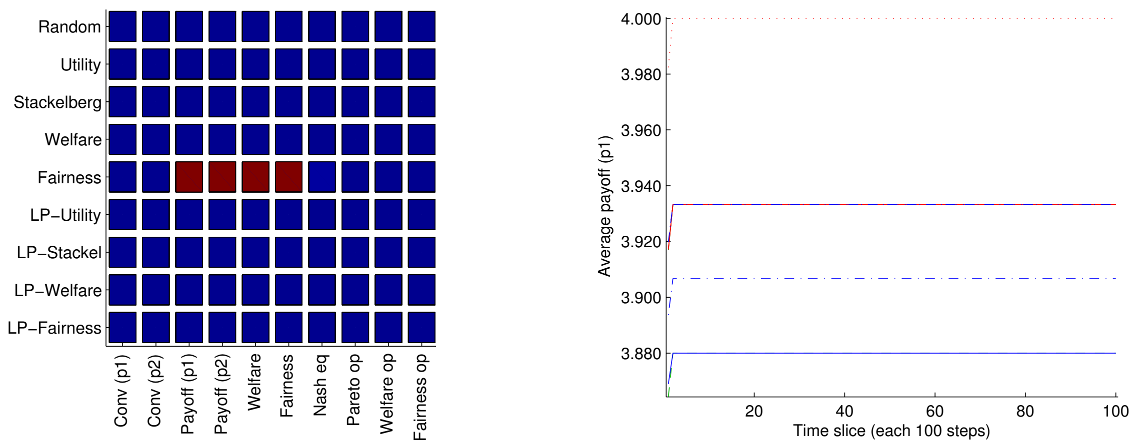


(c) $h = 1$

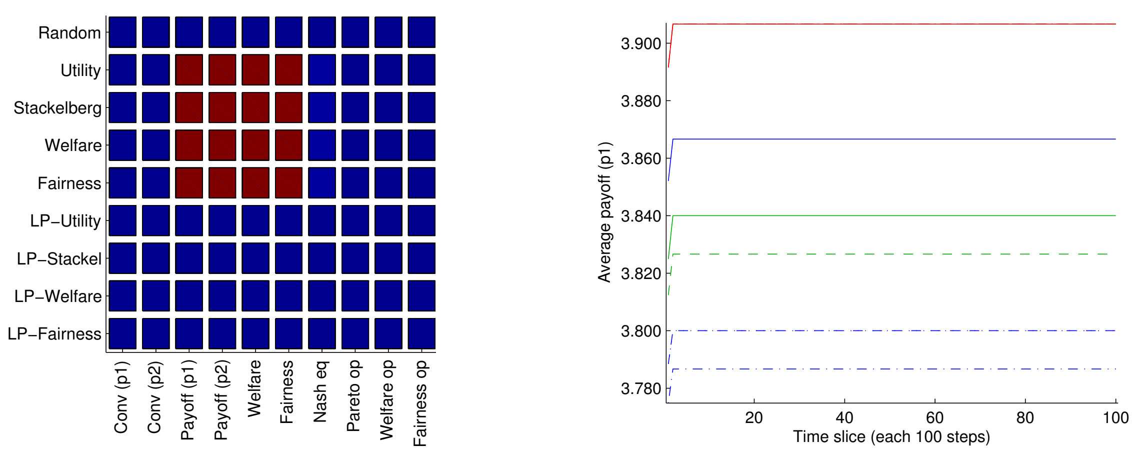
Figure 16: Player 1 controlled by HBA using **co-evolved neural networks** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **random type**. Results from **conflict games**.



(a) $h = 5$

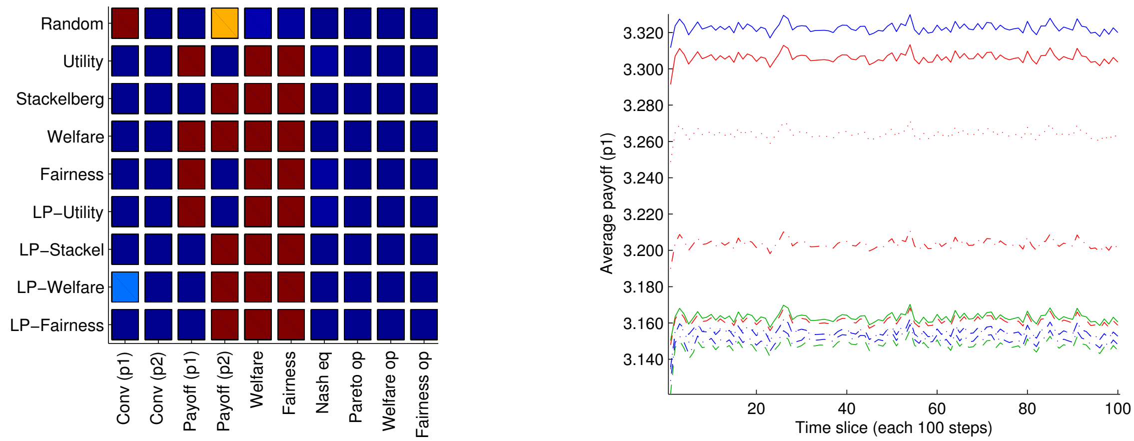


(b) $h = 3$

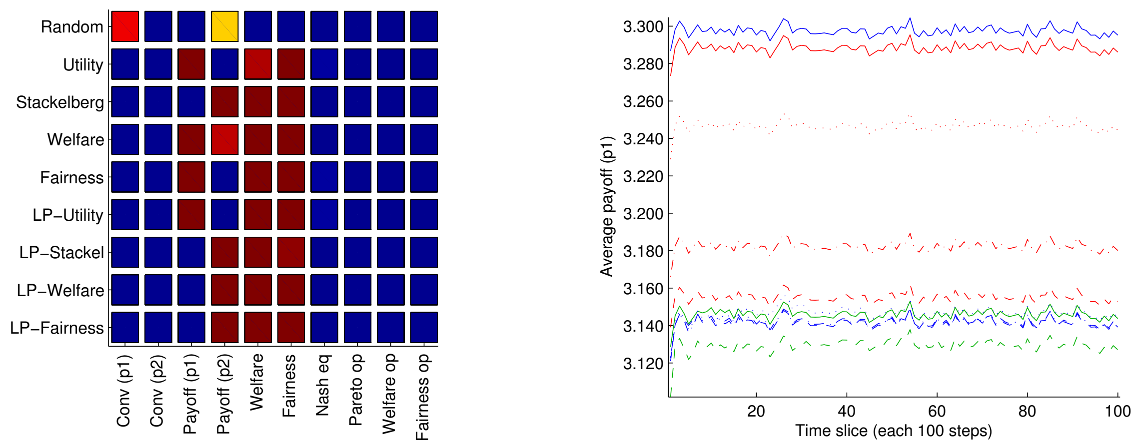


(c) $h = 1$

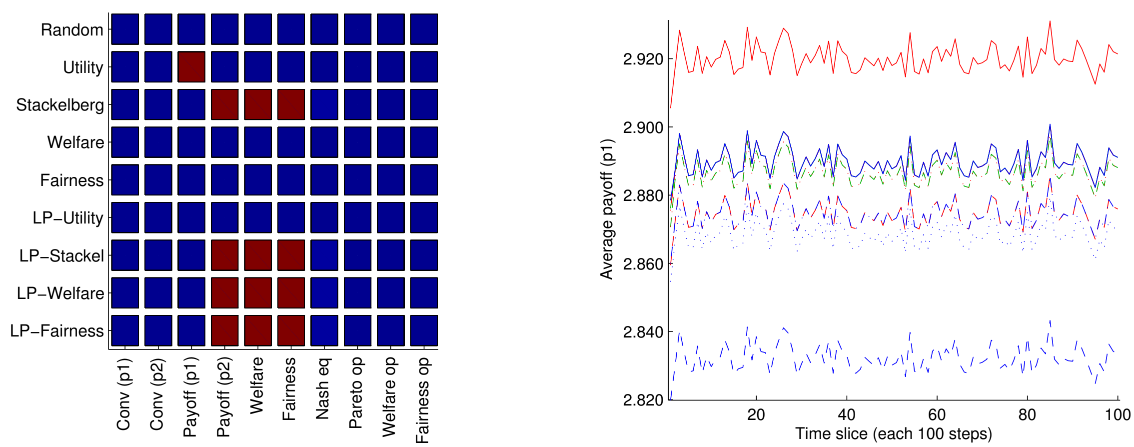
Figure 17: Player 1 controlled by HBA using **co-evolved neural networks** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **fictitious player**. Results from **no-conflict games**.



(a) $h = 5$

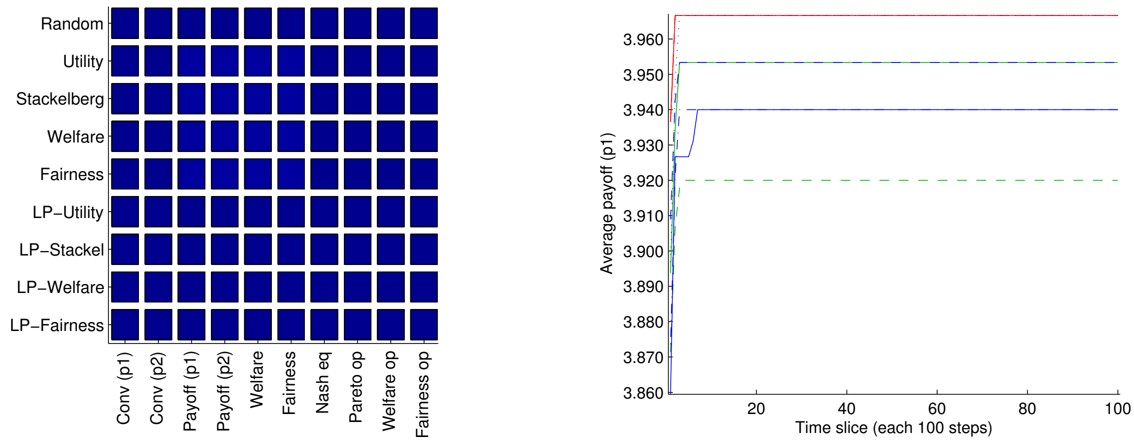


(b) $h = 3$

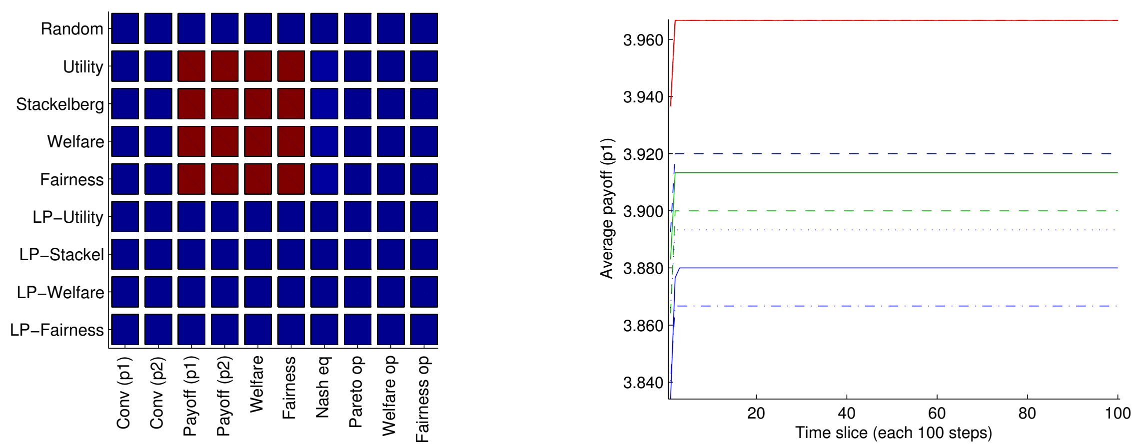


(c) $h = 1$

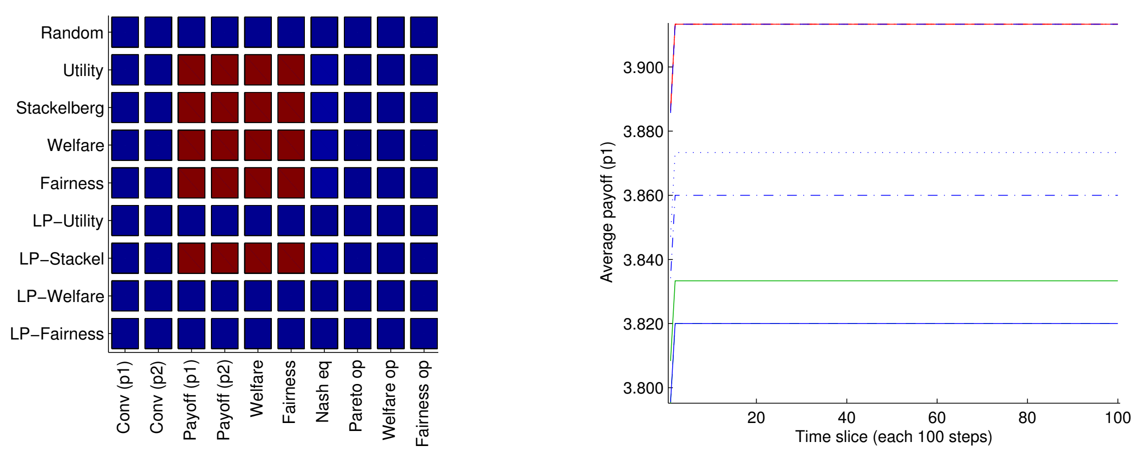
Figure 18: Player 1 controlled by HBA using **co-evolved neural networks** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **fictitious player**. Results from **conflict games**.



(a) $h = 5$

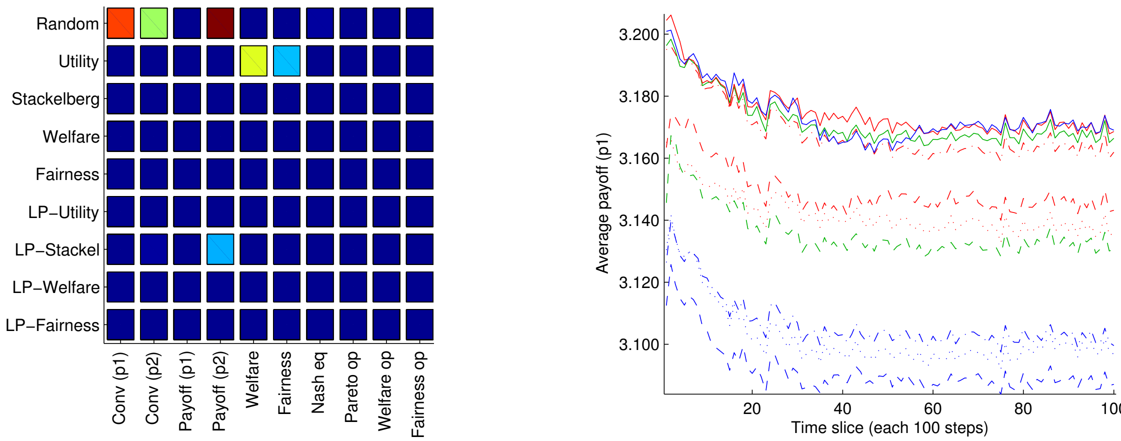


(b) $h = 3$

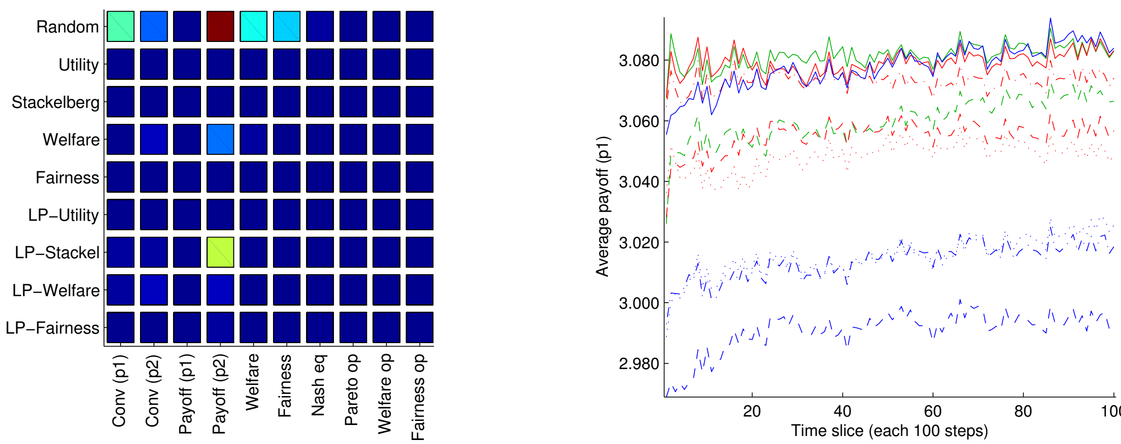


(c) $h = 1$

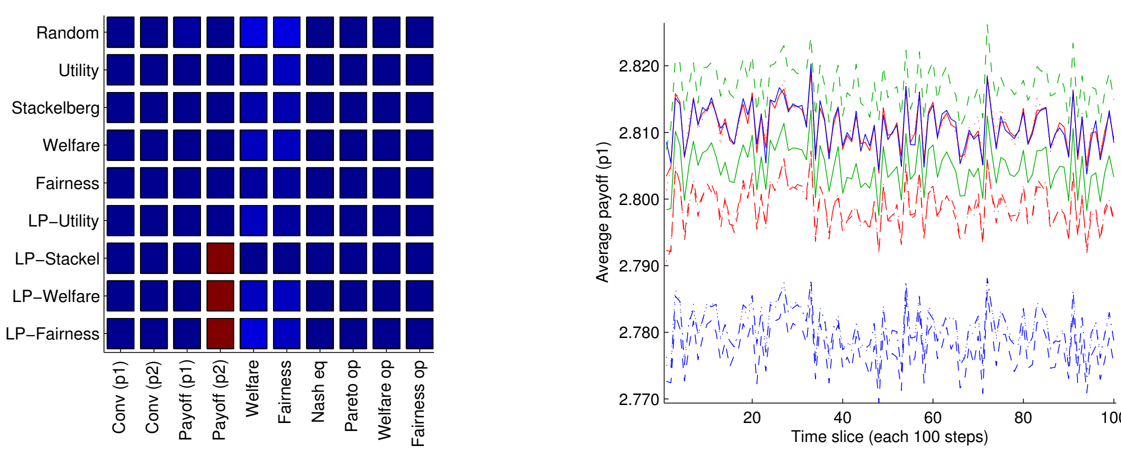
Figure 19: Player 1 controlled by HBA using **co-evolved neural networks** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **conditioned fictitious player**. Results from **no-conflict games**.



(a) $h = 5$



(b) $h = 3$



(c) $h = 1$

Figure 20: Player 1 controlled by HBA using **co-evolved neural networks** and planning horizons $h = 5, 3, 1$. Player 2 controlled by **conditioned fictitious player**. Results from **conflict games**.

References

[1] S. V. Albrecht, J. W. Crandall, and S. Ramamoorthy. An empirical study on the practical impact of prior beliefs over policy types. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, Austin, Texas, USA, January 2015.