Identifying the Opponent's Strategy by Behavioural Analysis in Repeated 2x2 Games

Joris Teunisse
Supervisors: Mehdi Dastani and Bas Testerink

Utrecht University

August 19, 2020

Outline

- Introduction
- Research questions
- Experiment framework
- Agent implementation
- Results
- Discussion

Introduction - Fundamental concepts

- This thesis studies how *agents* interact with each other.
- Specifically, we study interactions for which the agents have a specific *goal* in mind.
- Each agent applies a *strategy* that selects which actions should be taken.
- This type of interaction can be studied systematically by using the matrix game framework.

Introduction - Matrix games (1/2)

- How does a matrix game model an interaction?
  - Agents can take a finite number of actions.
  - These actions are taken simultaneously: this is called a *joint action*.
  - Rewards quantify the progress towards an agent's goal.

	α	β
a	$3, 3$	$0, 5$
b	$5, 0$	$1, 1$

Table 1: A matrix game.

Introduction - Matrix games (2/2)

- Two related concepts are of importance to this thesis.
  - In Nash equilibria, no agent could have improved its reward independently.
  - In Pareto optima, no joint action would have lead to an improved or equal reward for all agents.

	α	β
a	$3, 3$	$0, 5$
b	$5, 0$	$1, 1$

Table 1: A matrix game.

Introduction - Repeated games

- Repeated games extend the functionality of matrix games.
  - Playing multiple iterations of a *stage game* allows the agents to adapt to their opponents' behaviour.
- We specifically chose repeated 2x2 games, as they possess the minimal properties required for this research.

Introduction - Motivation

- In the context of repeated 2x2 games, agents generally aim to achieve the highest possible reward over time.
- To this end, some agents approximate the actions of their opponent.
- We were interested in an agent whose initial aim is to find out *exactly* which strategy its opponent applies, in order to exploit this information afterwards.

Research questions - Definitions

- A *candidate set* is a finite collection of strategies that an agent could be applying.
- An agent's strategy has been *identified* if the strategy was successfully derived from the agent's behaviour.
- A strategy's *distinguishability* represents how well it can be differentiated from its peers.

Research questions - Declarations

- How can an agent's strategy be identified efficiently from a set of candidate strategies, based upon the agent's behaviour in repeated 2x2 games?
  - What are properties of strategies and stage games that contribute to distinguishability?
  - What is the influence of the stage game's properties on the efficiency of identification?
  - What is the influence of the candidate set on the efficiency of identification?
  - Which methods for behavioural analysis are appropriate for which combinations of candidate sets and stage games?

Research questions - Hypotheses

- Distinguishability will be higher in stage games that provide many strategic options.
- Complex strategies will be harder to identify than simple ones.
- Both the stage game and the candidate set will influence the distinguishability of a strategy.
- Applying a strategy that specifically targets efficient identification will lead to improved results.

Experiment framework - Strategies (1/3)

- We selected strategies of varied complexity.
  - Static strategies: Pure 0, Pure 1, Random, Nash.
  - Dynamic strategies (non-learning): TFT, ϵ-FP, Relative FP.
  - Dynamic strategies (learning): ϵ-greedy Q, Boltz-Q.
- We will briefly discuss one strategy per category.

Experiment framework - Strategies (2/3)

- **Pure 0:** Take the first action.
- Pure 1: Take the second action.
- Random: Take either action with equal probabilities.
- Nash: Take the action which represents your part of a random NE.
- **TFT:** Initially take the action which represents your part of a random Pareto optimum. Afterwards, mimic the action of your opponent: take the first action if your opponent has taken the first action in the previous iteration, and the second action otherwise.

Experiment framework - Strategies (3/3)

- ϵ-FP: Apply the Fictitious Play algorithm. If the difference between the expected rewards is smaller than ϵ, take either action with equal probabilities. If not, take the action with the highest expected reward.
- Relative FP: Apply the Fictitious Play algorithm. Normalise the expected rewards between 0 and 1: take either action with the resulting probabilities.
- ϵ-**greedy Q:** Apply the Q-learning algorithm: take either action according to the ϵ-greedy exploration method.
- Boltz-Q: Apply the Q-learning algorithm: take either action according to the Boltzmann exploration method, rather than ϵ-greedy.

Experiment framework - Stage games

0	α	β
a	$3, 2$	$0, 0$
b	$0, 0$	$2, 3$

1	α	β
a	$0, 0$	$-1, 1$
b	$1, -1$	$-9, -9$

2	α	β
a	$1, 1$	$0, 0$
b	$0, 0$	$0, 0$

3	α	β
a	$2, 2$	$0, 0$
b	$0, 0$	$1, 1$

4	α	β
a	$1, 1$	$0, 3$
b	$3, 0$	$2, 2$

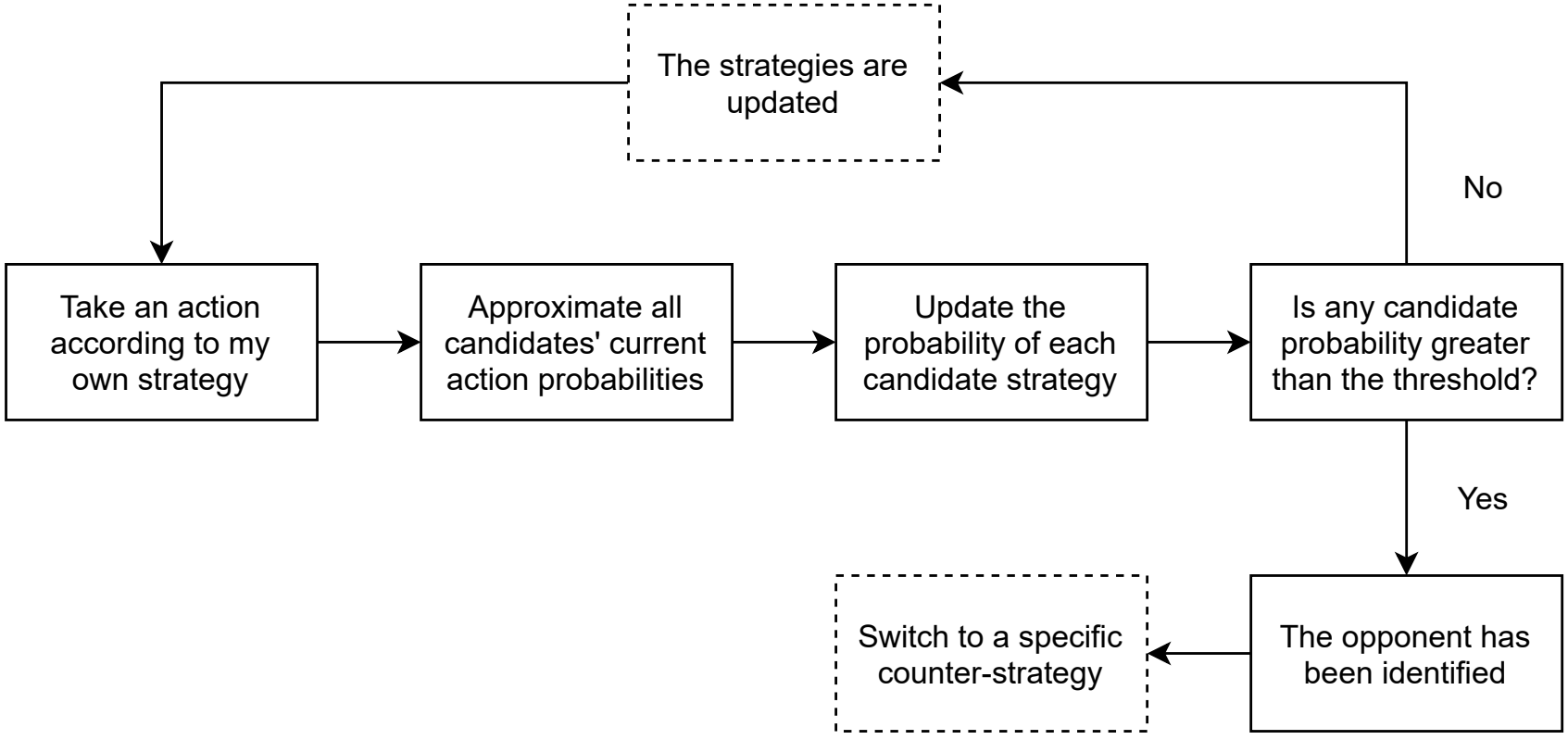
5	α	β
a	$1, -1$	$-1, 1$
b	$-1, 1$	$1, -1$

6	α	β
a	$3, 3$	$0, 5$
b	$5, 0$	$1, 1$

7	α	β
a	$4, 4$	$1, 3$
b	$3, 1$	$2, 2$

- Joint actions can be Nash equilibria, Pareto optima or both.
- Note that mixed Nash equilibria have not been visualised.

Agent implementation - Identification



- The goal of our Identification Agent (IA) is to identify the opponent's strategy as efficiently as possible.
- To this end, it calculates the probability of each candidate strategy given the history of joint actions.

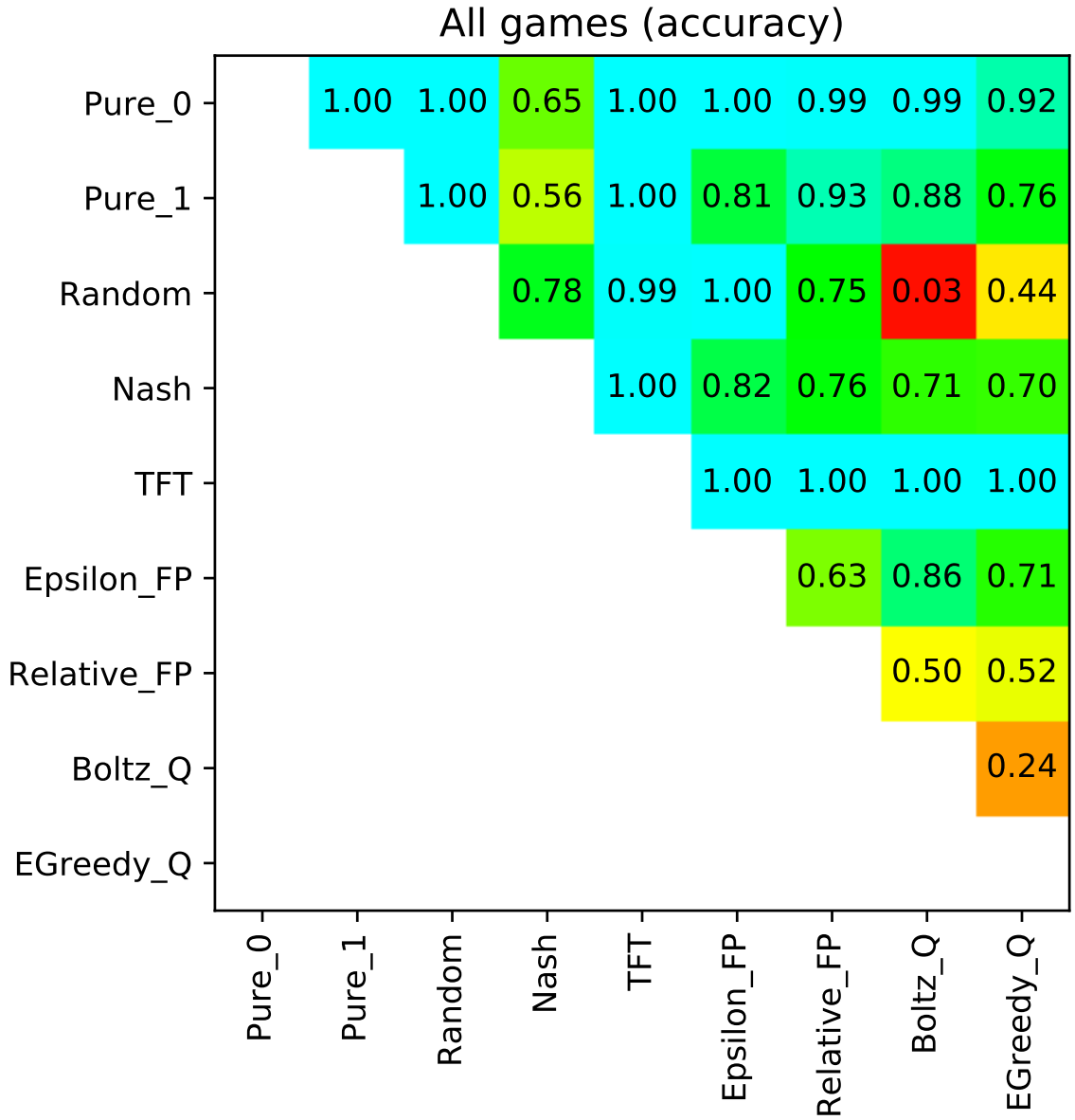
Agent implementation - Look-ahead strategy

- In contrast to the strategies discussed thus far, the look-ahead strategy specifically targets efficient identification.
- By simulating the consequences of its actions, this strategy estimates which action leads to faster identification on average.
- This generally results in actions that would provoke different responses from each candidate strategy.

Results - Setup

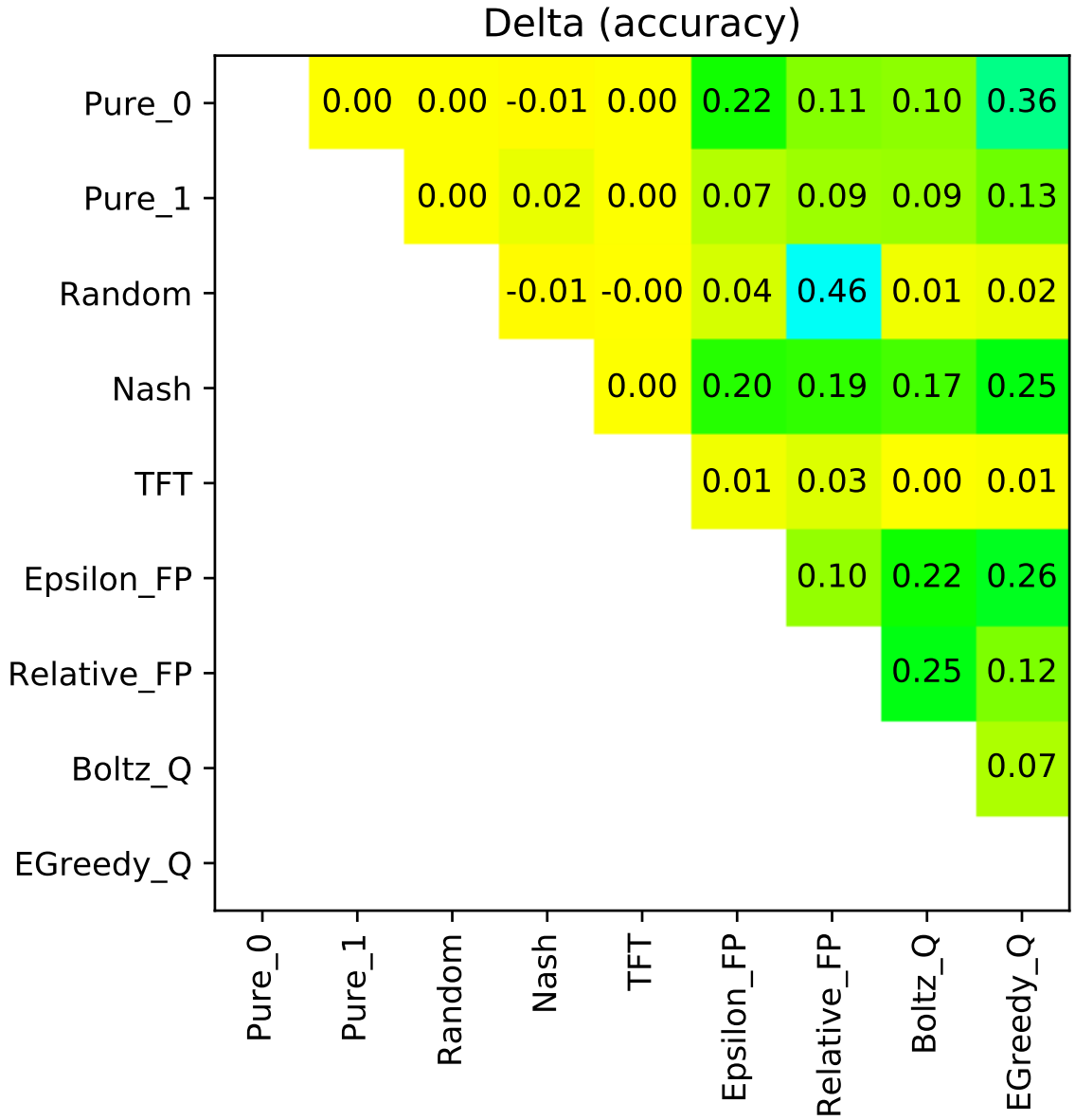
- In each experiment, the IA attempted to identify the strategy of its opponent within 10 iterations.
- We researched *candidate pairs*: candidate sets of size 2.
- The IA's performance was judged by using two main statistics: its *accuracy* and the *mean iterations* it required.
- The statistics average over 100 individual experiments.

Results - All stage games



- In general, distinguishability varies per candidate pair.

Results - Impact of look-ahead



- The look-ahead strategy frequently lead to improved efficiency.

Results - Statistics per strategy

Strategy	Avg. accuracy	Avg. mean	Failures
TFT	1.00	2.84	0
Pure 0	0.94	2.75	1
Pure 1	0.87	2.67	2
ϵ-FP	0.85	3.64	1
Relative FP	0.76	4.41	5
Nash	0.75	3.27	6
Random	0.75	5.79	7
Boltz-Q	0.66	5.02	6
ϵ-greedy Q	0.65	5.50	0

- The TFT strategy was identified with perfect accuracy: presumably due to its unique behaviour.
- Both Nash and Random ranked lower than expected.

Results - Statistics per stage game

Stage game	Avg. accuracy	Avg. mean	Failures
Chicken	0.87	3.61	0
Coordination	0.87	3.76	1
Bach or Stravinsky	0.85	4.04	1
Coordinate First	0.85	4.08	3
Prisoner's Dilemma	0.77	3.90	1
Stag Hunt	0.75	4.20	1
Matching Pennies	0.75	4.30	6
Deadlock	0.70	3.48	1

- Coordinate First lead to high accuracy values despite offering few strategic options.
- The converse was true for Matching Pennies.

Discussion - Conclusions

- Both the stage game and the specific candidate set can have a large influence on distinguishability.
- The look-ahead strategy improved the efficiency of identification for all non-static candidate pairs.
- The stage games' strategic options do not seem to be as influential as we expected.
- Lastly, the distinguishability of a strategy does appear to be related to its complexity: however, unique behaviour also seems to be an important factor.

Discussion - Future work

- Future work could research the benefits of applying different strategies for the purpose of identification.
- Larger candidate sets could also be investigated.
- Finally, this research can be used as a baseline for an agent that applies a specific counter-strategy after identification.

The end

Thank you for listening!