Online Appendix for Algorithmic Cooperation

Bernhard Kasberger[*] Simon Martin[†] Hans-Theo Normann[‡]
Tobias Werner[§]

July 2023

Contents

[*]kasberger@dice.hhu.de. Düsseldorf Institute for Competition Economics (DICE) at Heinrich-Heine-Universität Düsseldorf

[†]simon.martin@dice.hhu.de. Düsseldorf Institute for Competition Economics (DICE) at Heinrich-Heine-Universität Düsseldorf, CESifo Research Affiliate

[‡]normann@hhu.de. Düsseldorf Institute for Competition Economics (DICE) at Heinrich-Heine-Universität Düsseldorf, Max Planck Institute for Research on Collective Goods, Bonn

[§]werner@mpib-berlin.mpg.de. Center for Humans and Machines at the Max Planck Institute for Human Development

S.1 Additional tables

Table S.1: List of $k = 1$ strategies for manual classification

ID	Label	DD (00)	DC (01)	CD (10)	CC (11)	Remark
1	AllC	C	C	C	C	
2		C	C	C	D	
3		C	C	D	C	
4	CDAlt	C	C	D	D	
5		C	D	C	C	
6		C	D	C	D	
7	WSLS	C	D	D	C	
8	WShLSh	C	D	D	D	
9		D	C	C	C	Conv. State = (DD)
10		D	C	C	D	
11	TFT	D	C	D	C	
12		D	C	D	D	
13		D	D	C	C	
14		D	D	C	D	
15	Grim	D	D	D	C	Conv. State = (CC)
16	AllD	D	D	D	D	
17	AllD'	D	D	D	C	Conv. State $\neq (CC)$
18	AllC'	D	C	C	C	Conv. State $\neq (DD)$

Table S.2: Data sources for experiments with human subjects

δ	$R = 32$	$R = 40$	$R = 48$
0.5	DBF2011/19, GS2022 (9)	DBF2011/19, GS2022 (6)	DBF2011/19, GS2022 (11)
0.75	DBF2011/19, GS2022 (11)	DBF2011/19, GS2022 (6)	DBF2011/19, GS2022 (11)
0.90	DBF2019 (7)	KMNW2023 (6)	KMNW2023 (6)
0.95	DBF2019, RR2018 (8)	KMNW2023 (6)	KMNW2023 (6)

Note: This table lists the studies with human subjects whose data we used for our comparisons (DBF2011: Dal Bó and Fréchette (2011), DBF2019: Dal Bó and Fréchette (2019), GS2022: Ghidoni and Suetens (2022), RR2018: Romero and Rosokha (2018), KMNW2023: this paper). In parentheses, we indicate the number of independent observations for each treatment combination. A single independent observation corresponds to a matching group.

Table S.3: Manual vs. SFEM

δ	R	Method	AllC	CDAlt	WSLS	WShLSh	TFT	Grim	AllD	Other	σ	LL	MSE
0.50	32	Manual	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0			
		SFEM	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	1.00	0.0	0.0
	40	Manual	0.0	0.0	0.0	0.0	0.0	0.0	99.9	0.1			
		SFEM	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	1.00	0.0	0.0
	48	Manual	0.1	0.0	0.0	5.3	0.0	0.0	90.6	4.0			
		SFEM	0.0	0.0	0.0	5.3	0.0	0.0	93.3	1.4	1.00	−0.6	1.7
0.75	32	Manual	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0			
		SFEM	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	1.00	0.0	0.0
	40	Manual	0.0	0.0	0.3	0.1	0.0	0.0	98.8	0.8			
		SFEM	0.0	0.0	0.3	0.1	0.0	0.0	99.6	0.0	1.00	−0.1	0.2
	48	Manual	2.2	0.1	4.4	37.8	1.1	0.0	47.2	7.2			
		SFEM	4.4	1.8	4.4	37.8	2.2	0.0	48.6	0.8	0.98	−5.4	6.5
0.90	32	Manual	1.7	1.0	1.2	0.1	5.3	0.0	67.5	23.2			
		SFEM	3.2	1.0	1.2	0.4	3.7	0.0	79.7	10.7	0.99	−3.5	38.8
	40	Manual	6.1	1.0	18.4	7.2	6.8	0.0	43.1	17.4			
		SFEM	8.9	1.5	18.7	7.7	7.7	0.0	47.8	7.7	0.99	−5.2	15.7
	48	Manual	10.6	0.0	15.3	56.2	8.1	0.0	3.3	6.5			
		SFEM	11.9	1.2	15.6	56.2	10.2	0.0	3.8	1.0	0.99	−5.4	4.7
0.95	32	Manual	9.0	2.3	10.3	2.2	17.1	0.0	33.6	25.5			
		SFEM	12.0	1.4	10.4	2.1	14.9	0.0	39.6	19.5	0.99	−4.8	10.9
	40	Manual	13.3	0.4	18.4	8.9	18.5	0.0	23.4	17.1			
		SFEM	18.0	0.9	18.8	9.2	13.8	0.0	27.0	12.4	0.99	−4.9	9.9
	48	Manual	7.1	0.2	9.9	64.2	6.0	0.0	3.1	9.5			
		SFEM	10.3	1.3	10.1	64.2	7.3	0.0	3.8	2.9	0.99	−5.3	7.2

Table S.4: Average cooperation by ChatGPT

δ	$R = 32$	$R = 40$	$R = 48$
0.50	76.34***	81.34***	75.48***
0.75	75.98***	79.84*	78.30*
0.90	79.14***	81.38*	82.18*
0.95	74.96	81.34*	80.62

Note: This table shows the average cooperation rate by ChatGPT, as well as the significance of a two-sided Mann-Whitney-U-test relative to human cooperation. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

S.2 Algorithmic strategies with memory one

We now reconsider the strategies used by algorithms as described in Section 5.2 of the main paper, focusing on memory-one strategies for $k = 1$ algorithms. Note that an algorithm with memory length one ($k = 1$) cannot learn a strategy that requires higher memory. In our main SFEM specification, this possibility is not ruled out upfront and hence might arise due to noise introduced in the SFEM procedure.

The memory one strategies we investigate now are the following: AllC, AllD, CtoD, TFT, DFFT, Grim, WSLS, DCAlt, and WShLSh. The results of the adjusted SFEM are shown in Table S.5. Comparing with the baseline results from Table 9, we find that the SFEM is very robust to our adjustments. Even when strategies with memory two were enabled, SFEM very rarely classified observed strategies as such, and hence omitting them is inconsequential. Our interpretation of these results is that our main specification and the following results are robust to the adjustments made in this robustness check.

Table S.5: Distribution of strategies by experiment for $k = 1$, memory one strategies only.

δ	R	AllC	AllD	TFT	DTFT	WSLS	WShLSh	σ
0.50	32	0.0	100.0	0.0	0.0	0.0	0.0	1.00
	40	0.0	100.0	0.0	0.0	0.0	0.0	1.00
	48	0.1	93.3***	0.0	1.3***	0.0	5.3***	1.00
0.75	32	0.0	100.0	0.0	0.0	0.0	0.0	1.00
	40	0.0	99.6***	0.0	0.0	0.3*	0.1***	1.00
	48	4.4***	48.6***	2.2***	0.8**	4.4***	37.8***	0.98
0.90	32	3.4***	79.7***	3.7***	10.3***	1.2***	0.6***	0.99
	40	9.1***	47.8***	7.7***	7.4***	18.7***	7.8***	0.99
	48	11.9***	3.8***	10.2***	1.0**	15.6***	56.2***	0.99
0.95	32	12.1***	39.6***	14.9***	19.4***	10.4***	2.1***	0.99
	40	18.1***	27.0***	13.8***	12.2***	18.8***	9.1***	0.99
	48	10.3***	3.8***	7.3***	2.9***	10.1***	64.2***	0.99

S.3 Human cooperation: Levels of cooperation across supergames

As a robustness check, we investigate the levels of human cooperation in our own experiments (where $\delta \in \{0.9, 0.95\}$ and $R \in \{40, 48\}$) for varying subsets of supergames. In Table S.6, we consider all supergames (1); the first supergame only (2); the last three quarters of supergames (3), and the last supergame (4).

Table S.6: Average cooperation by humans by supergame

	(1) All SG		(2) First SG		(3) Last 3/4 SG		(4) Last SG	
δ	$R = 40$	$R = 48$	$R = 40$	$R = 48$	$R = 40$	$R = 48$	$R = 40$	$R = 48$
0.90	0.77	0.89	0.75	0.66	0.80	0.96	0.88	0.99
0.95	0.79	0.88	0.68	0.69	0.84	0.91	0.89	0.90

Note: This table depicts average cooperation rates by humans.

S.4 Sensitivity Analysis of Large Language Models

As a robustness check, we investigate the levels of cooperation for ChatGPT (where $\delta \in \{0.9, 0.95\}$ and $R \in \{40, 48\}$) for varying subsets of rounds. In Table S.7, we consider all rounds (1); the first round only (2), and the last round only (3).

Table S.7: Average cooperation by ChatGPT by round

	(1) All Rounds			(2) First Round			(3) Last Round		
δ	$R = 32$	$R = 40$	$R = 48$	$R = 32$	$R = 40$	$R = 48$	$R = 32$	$R = 40$	$R = 48$
0.50	0.76	0.81	0.75	0.89	0.93	0.88	0.74	0.80	0.72
0.75	0.76	0.80	0.78	0.89	0.91	0.90	0.76	0.78	0.75
0.90	0.79	0.81	0.82	0.90	0.89	0.91	0.79	0.80	0.80
0.95	0.75	0.81	0.81	0.86	0.91	0.89	0.75	0.79	0.77

Note: This table depicts average cooperation rates by ChatGPT.

Similarly, we investigate the determinants of average cooperation for ChatGPT in Table S.8. As the discussion in the main text already suggested, ChatGPT does not appear very responsive to changes in the environment. This is also confirmed by our regression analysis.

Table S.8: Determinants of average cooperation for ChatGPT

	(1)	(2)	(3)	(4)
δ	4.97			
	(3.24)			
R	0.16			
	(0.09)			
GT		1.51		
		(2.20)		
RD		2.02		
		(1.23)		
$\underline{p}$			-3.33	
			(1.77)	
$\delta - \delta^{RD}$				5.49^{*}
				(2.31)
Constant	68.71^{***}	76.34^{***}	79.99^{***}	77.98^{***}
	(4.33)	(1.97)	(0.81)	(0.69)
N	3000	3000	3000	3000

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

S.5 Human experiment

In this appendix, we state our preregistered hypothesis. The experimental literature has shown that cooperation among humans can be expected to increase in the discount factor and the reward parameter (Dal Bó and Fréchette, 2018). Our first hypothesis extends this insight to the parameter region in the δ-R space that has hitherto been unexplored. We address the hypothesis in the main text.

Hypothesis 1. *The cooperation rate among humans increases in R and δ.*

S.6 Q-Learning

Optimization problem. Q-learning agents aim to maximize the stream of future discounted rewards π_t in period t from being in some state $s_t \in S$ by picking an action $a_t \in A$. The state of possible states and action is denoted by S and A, respectively. The Bellman equation can express the optimization problem:

$$V(s_t) = \max_{a_t}\{\mathbb{E}[\pi_t|s_t, a_t] + \delta\mathbb{E}[V(s_{t+1})|s_t, a_t]\} \tag{1}$$

with $\delta \in [0, 1)$ being the discount rate. Importantly, the Q-learning agent does not have prior knowledge about the environment it interacts with. Hence, it does not know which state-action combinations yield the highest reward. In Q-learning, the Bellman equation is rewritten as the Q-function

$$Q(s_t, a_t) = \mathbb{E}[\pi_t|s_t, a_t] + \delta\mathbb{E}[\max_a Q(s_{t+1}, a)|s_t, a_t] \tag{2}$$

We assume that both A and S are finite sets. Then, the Q-function can be expressed as a $|A| \times |S|$ matrix. Each cell of this Q-matrix then represents the approximation of the discounted stream of rewards for a specific state-action combination.

In each period t the Q-learning agent picks an action a_t and updates the Q-matrix as a weighted average between the past approximation of the state-action value and the newly learned information:

$$Q_{t+1}(s_t, a_t) = (1 - \alpha)Q_t(s_t, a_t) + \alpha(\pi_t + \delta \max_a Q_t(s_{t+1}, a)) \tag{3}$$

The parameter $\alpha \in [0, 1)$ is the learning rate. It governs the weight which is given to newly arriving information. If α is large, more weight is given to the newly learned values. On the other hand, if α is small, the algorithm values the existing approximation of the state-action higher. Note that the agent does not learn α but it is a hyperparameter that we set as a researcher or if we, for instance, think about pricing algorithms, the developer of the algorithm.

Exploration and exploitation. In each period, the agent has to decide which action to choose conditionally on observing the current state of the environment. Given the current approximation of the Q-matrix, the agent should pick the action $a_t^* = \arg\max_a Q(s_t, a)$ as it corresponds to the profit-maximizing choice. At the same time, the approximation of the Q-matrix might not be optimal. The algorithm does not have any prior knowledge about the environment. Hence, it has to explore the environment and try out different actions in different states to gather information. For instance, specific actions might be underexplored in the given period and yield a higher reward than the current approximation might suggest. Thus, the agent faces a trade-off between exploring the action space and exploiting the information it already has about the environment. To circumvent the trade-off, the ϵ-greedy algorithm is used in Q-learning.[1] With probability $1 - \epsilon_t$ the agent picks greedy action a_t^*, e.g., the action that is optimal given the current approximation of the Q-matrix. Conversely, to explore the environment, the agent picks a random action from the action set with probability ϵ_t. We let ϵ_t decay over time and define it as $\epsilon_t = e^{-\beta t}$ for some small $\beta > 0$. At the beginning of the agents' training, the Q-matrix is still uninformative as it has not explored the state-action space sufficiently. Since t is small, ϵ_t is close to 1 and the agent picks random actions frequently to explore the environment. As the training of the agent progresses, the Q-matrix becomes more informative and it chooses the greedy action more often. Note that in the limit, the agent picks a_t^* only as $\lim_{t\to\infty} \epsilon_t = 0$.

Q-learning in the Prisoner's Dilemma. In the Prisoner's Dilemma, the agent can choose between cooperation (C) and defection (D). Thus, the action set is given by $A = \{C, D\}$. We define $m = |A| = 2$ as its cardinality. The reward signal π_t is simply given by the payoff for the agent in the given round.

Following Calvano et al. (2020), we consider state representations where the agents condition their current action in period t on the set of actions from the previous k periods. We denote $\mathbf{a_k}$ as the vector of action by both agents in period

[1] There are other exploration algorithms like Boltzmann exploration. Calvano et al. (2020) show that it yields similar results in market environments; thus, we focus on the ϵ-greedy algorithm.

k. The state space is then $S = \{\mathbf{a_1}, \mathbf{a_2}, ..., \mathbf{a_k}\}$. Note that the cardinality of S grows exponentially in k as $|S| = m^{nk} = 2^{2k}$.

The cardinality of S interacts with the exploration parameter β. Holding β fixed, the probability that a cell is visited by chance through exploration is smaller in larger state spaces. To address this, we choose $\beta(k)$ to keep ν constant for all k, where ν is the expected value of a cell in the Q-matrix being explored purely by randomness (see Calvano et al. (2020), Section III.A).[2]

$$\nu = \frac{(m-1)^n}{m^{kn+n+1}\left(1-e^{-\beta(n+1)}\right)} = \frac{1}{2^{2k+3}\left(1-e^{-3\beta}\right)}.$$

S.7 Instructions

We provide the experimental instructions for humans and large language models for the version with $\delta = 0.95$ and $R = 48$ below. Across treatments, we only varied those two parameters in the instructions. The original instructions for humans are in German and based on the instructions used by Dal Bó and Fréchette (2011). We also provide a translation below. The instructions for LLMs are in English.

S.7.1 Human instructions

S.7.1.1 Original version

Willkommen:

> Willkommen zu unserem Experiment!
>
> Sie nehmen gleich an einem Experiment zum Thema Entscheidungsfindung teil und werden am Ende des Experiments privat für Ihre Teilnahme bezahlt. Was Sie verdienen, hängt zum Teil von Ihren Entscheidungen, zum Teil von den Entscheidungen anderer und zum Teil vom Zufall ab. Für dieses Experiment gibt es generell 4€ Teilnahmeprämie, plus den Verdienst, der aus den Entscheidungen resultiert.
>
> Das gesamte Experiment wird über Computer abgewickelt, und die gesamte Interaktion zwischen Ihnen wird über die Computer stattfinden.

[2] The formula reported in Calvano et al. (2020), footnote 20, erroneously contains $kn(n+1)$ in the denominator instead.

Bitte schalten Sie jetzt Ihre Mobiltelefone aus. Bitte sprechen Sie nicht und versuchen Sie nicht, während der Sitzung mit anderen Teilnehmern zu kommunizieren.

Wir beginnen mit einer kurzen Einführungsphase. Während der Einweisung erhalten Sie eine Beschreibung der Hauptmerkmale des Experiments und es wird Ihnen gezeigt, wie Sie die Computer benutzen. Wenn Sie während dieser Zeit Fragen haben, heben Sie die Hand und Ihre Frage wird so beantwortet, dass jeder sie hören kann.

Anleitung:

1. In diesem Experiment werden Sie aufgefordert, in mehreren Runden Entscheidungen zu treffen. Sie werden nach dem Zufallsprinzip mit einer anderen Person für eine Reihe von Runden gepaart. Jede Folge von Runden wird als Spiel bezeichnet.
2. Die Länge eines Spiels wird zufällig bestimmt. Nach jeder Runde besteht eine 95-prozentige Wahrscheinlichkeit, dass das Spiel für mindestens eine weitere Runde fortgesetzt wird. Das ist so, als würden wir nach jeder Runde einen zwanzigseiten Würfel werfen und bei “1” bis “19” weitermachen und bei “20” aufhören. Wenn Sie z. B. in Runde 2 sind, beträgt die Wahrscheinlichkeit, dass es eine dritte Runde gibt, 95%, und wenn Sie in Runde 9 sind, beträgt die Wahrscheinlichkeit, dass es eine weitere Runde gibt, ebenfalls 95%.
3. Sobald ein Spiel endet, werden Sie zufällig mit einer anderen Person für ein neues Spiel gepaart.
4. Es werden insgesamt maximal 15 Spiele gespielt. Das erste Spiel, das nach 120 Minuten Spielzeit beendet ist, markiert das Ende des Experiments.
5. Die Auswahlmöglichkeiten und die Auszahlungen in jeder Runde sind wie folgt:

	Entscheidung des anderen Spielers	
Ihre Entscheidung	A	B
A	48, 48	12, 50
B	50, 12	25, 25

Der erste Eintrag in jeder Zelle steht für Ihre Auszahlung, während der zweite Eintrag die Auszahlung der Person darstellt, mit der Sie zusammengeführt werden.

- Sie wählen A und der/die andere wählt A, dann erhalten Sie beide 48.
- Sie wählen A und der/die andere wählt B, dann erhalten Sie 12 und der/die andere 50.
- Sie wählen B und der/die andere wählt A, dann erhalten Sie 50 und der/die andere 12.
- Sie wählen B und der/die andere wählt B, dann erhalten Sie beide 25.

6. Am Ende des Experiments wird ein Spiel zufällig für die Auszahlung ausgelost. Von den in diesem Spiel erzielten Punkten erhalten Sie 1 € für 40 erzielte Punkte. Hinzukommt die oben erwähnte Zahlung von 4€.
7. Bevor wir beginnen, möchten wir Sie daran erinnern: Die Dauer eines Spiels wird nach dem Zufallsprinzip bestimmt. Nach jeder Runde besteht eine 95-prozentige Wahrscheinlichkeit, dass das Spiel für mindestens eine weitere Runde fortgesetzt wird. Sie spielen während des gesamten Spiels (über alle Runden) mit der gleichen Person. Bei einem neuen Spiel werden Sie zufällig mit einer anderen Person für ein neues Spiel gepaart. Für Ihre Auszahlung wird am Ende ein Spiel zufällig ausgelost.

Gibt es noch Fragen?

Quiz

1. Wie viele Punkte verdienen Sie in einer Runde, wenn Sie und der andere Teilnehmer Option A wählen?
2. Wie viele Punkte verdienen Sie in einer Runde, wenn sowohl Sie als auch der andere Teilnehmer Option B wählen?
3. Wie viele Punkte verdienen Sie in einer Runde, wenn Sie Option A wählen und der andere Teilnehmer Option B wählt?
4. Wie viele Punkte verdienen Sie in einer Runde, wenn Sie Option B wählen und der andere Teilnehmer Option A wählt?

5. Nehmen wir an, Sie befinden sich in der vierten Runde eines Spiels. Wie hoch ist die Wahrscheinlichkeit, dass das Spiel nach dieser Runde beendet wird?

S.7.1.2 English translation

Welcome

Welcome to our experiment!

You are about to participate in an experiment on decision-making, and at the end of the experiment you will be paid privately for your participation. What you earn depends partly on your decisions, partly on the decisions of others, and partly on chance. There is generally 4€ participation pay for this experiment, plus the earnings that result from the decisions.

The entire experiment will be conducted through computers, and all interaction between you will be through the computers.

Please turn off your cell phones now. Please do not speak or attempt to communicate with other participants during the session.

We will begin with a brief introductory period. During the briefing, you will be given a description of the main features of the experiment and shown how to use the computers. If you have any questions during this time, raise your hand and your question will be answered so that everyone can hear it.

Instruction

1. In this experiment, you will be asked to make choices in several rounds. You will be randomly paired with another person for a series of rounds. Each sequence of rounds is called a game.
2. The length of a game is determined randomly. After each round, there is a 95 percent chance that the game will continue for at least one more round. This is like rolling a twenty-sided die after each round and continuing at "1" through "19" and stopping at "20". For example, if you are in round 2, the probability that there will be a third round is 95%, and if you are in round 9, the probability that there will be another round is also 95%.

3. Once a game ends, you will be randomly paired with another person for a new game.
4. There will be a maximum of 15 games played in total. The first game to finish after 120 minutes of play marks the end of the experiment.
5. The choices and payouts in each round are as follows:

	Other player's choice	
Your decision	A	B
A	48, 48	12, 50
B	50, 12	25, 25

   The first entry in each cell represents your payoff, while the second entry represents the payoff of the person you are matched with.

   - You choose A and the other person(s) chooses A, then you both get 48.
   - You choose A and the other chooses B, then you get 12 and the other 50.
   - You choose B and the other chooses A, then you get 50 and the other 12.
   - You choose B and the other chooses B, then you both get 25.

6. At the end of the experiment, one game is randomly drawn for payoff. From the points scored in this game, you will receive 1 € for 40 points scored. In addition, there is the above-mentioned payment of 4€.
7. Before we begin, we would like to remind you: the duration of a game is determined randomly. After each round, there is a 95% probability that the game will continue for at least one more round. You will play with the same person during the entire game (over all rounds). When a new game is played, you will be randomly paired with another person for a new game. For your payout, a game is randomly drawn at the end.

Are there any questions?

Quiz

1. How many points do you earn in a round if you and the other participant choose Option A?
2. How many points do you earn in a round if both you and the other participant choose Option B?
3. How many points do you earn in a round if you choose Option A and the other participant chooses Option B?
4. How many points do you earn in a round if you choose Option B and the other participant chooses Option A?
5. Assuming you are in the fourth round of a game, what is the probability that the game will end after this round?

S.7.2 Large Language Model instructions

System Message:

> You are a participant in my experiment. You play a repeated game. The length of the repeated game is determined randomly. Your objective is to maximize your own total payoff across all rounds. The total payoff is the sum of points across all rounds.

Initial Instructions:

> You play the game against another AI. The other player receives the same instructions and tries to maximize its own total payoff.
>
> 1. You have to make choices in several rounds. You play with the same player over all rounds.
> 2. The length of a game is determined randomly. After each round, there is a 95% chance that the game will continue for at least one more round. With a probability of 5% the game ends.
> 3. You can choose two actions: A and B.
> 4. The payoffs in each round are the following:
>   - You choose A and the other player chooses A: You get a payoff of 48 points. The other gets a payoff of 48 points.
>   - You choose A and the other player chooses B: You get a payoff of 12 points. The other gets a payoff of 50 points.

- You choose B and the other player chooses A: You get a payoff of 50 points. The other gets a payoff of 12 points.
- You choose B and the other player chooses B: You get a payoff of 25 points. The other gets a payoff of 25 points.

5. Your total payoff is the sum of points across all rounds. Your objective is to maximize your own total payoff over the course of the entire experiment.
6. You play the game against another AI. The other player receives the same instructions and tries to maximize its own total payoff. After each round, I will let you know your own payoff from the last round and your total payoff across all rounds. Furthermore, I will remind you of the choice of your own choice and tell you about the choice of the other player.

Find a strategy to play this game yourself and plan ahead. Think strategically and step-by-step about the strategy that you want to use in the game and develop a plan for the game against the other player. Reflect on the payoff structure. Also, think about the repeated nature of the game and that the game may end after each round with probability 5% when developing the plan. The plan should help you to achieve your goal which is to maximize the sum of points across all rounds.

Please pretend that you are a participant and can make choices in this experiment.

Quiz:

Before the game, let us do some quizzes to test your understanding of rules. Think step by step before answering. Please answer the following questions:

- If you choose A and the other player chooses A, your payoff is []
- If you choose B and the other player chooses A, your payoff is []
- If you choose A and the other player chooses B, your payoff is []
- If you choose B and the other player chooses B, your payoff is []
- After each round, there is a []% probability that another round is played.

Please only answer what your payoffs are not the other’s.

First Choice Message:

What action do you choose in the first round? Afterward, I will tell you what the other player has chosen.

I will also tell you whether the game ends or another round is played. Remember that the game may end at random after each round with probability 5%.

Please provide your choice in this round in JSON format, with your choice in the “choice” key like { “choice”: “[]” }. Replace [] with either A or B. Please answer in the exact format (no need to tell your reasoning).

Each Round Message:

You picked _own_choice_ in the last round. The other player picked _partner_choice_ in the last round. Your payoff in the last round was _payoff_last_round_. Your current total payoff across all rounds is _payoff_total_. Your objective is to maximize your own total payoff over the course of the entire experiment.

The game continues. Keep in mind that the game may end at random with probability 5%.

Please provide your choice in this round in JSON format, with your choice in the “choice” key like { “choice”: “[]” }. Replace [] with either A or B. Please answer in the exact format (no need to tell your reasoning).

Last Message:

The game has ended. Your total payoff is _payoff_total_. Please reflect on the outcome of the game. Explain the strategy that you used during the game. Before you answer think carefully and consider the entire history of the game.

References

Calvano, Emilio, Giacomo Calzolari, Vincenzo Denicolò, and Sergio Pastorello, "Artificial intelligence, algorithmic pricing, and collusion," *American Economic Review*, 2020, *110* (10), 3267–3297.

Dal Bó, Pedro and Guillaume R. Fréchette, "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence," *American Economic Review*, February 2011, *101* (1), 411–29.

_ **and** _ , "On the Determinants of Cooperation in Infinitely Repeated Games: A Survey," *Journal of Economic Literature*, March 2018, *56* (1), 60–114.

_ **and** _ , "Strategy choice in the infinitely repeated Prisoner's Dilemma," *American Economic Review*, 2019, *109* (11), 3929–52.

Ghidoni, Riccardo and Sigrid Suetens, "The effect of sequentiality on cooperation in repeated games," *American Economic Journal: Microeconomics*, 2022, *14* (4), 58–77.

Romero, Julian and Yaroslav Rosokha, "Constructing strategies in the indefinitely repeated prisoner's dilemma game," *European Economic Review*, 2018, *104*, 185–219.