The Game Theory of You and Your Future Self

Shiran Yuan*

Duke Kunshan University, Kunshan, Jiangsu 215316, China
sy250@duke.edu
https://shiran-yuan.github.io

Abstract. It is a well known consensus in behavioral science and psychology that humans are very irrational regarding making predictions about the future. While some prior works in game theory have took this factor into account by modifying an agent's utility based on temporal factors, none have considered separating the same agent's utilities and choices in different times, thereby effectively modeling them as separate agents. This paper attempts to shed light into this methodology, exploring its effects on game theoretical models. In particular, we propose splitting a single player into multiple agents, and subsequently modeling interactions between them, as well as their interactions with other players. By tuning "foresight functions," we can control the model's degree and direction of deviation from standard, rational assumptions, explaining actions such as shortsighted decisions and saving for the future.
This work is a course report for ECON 206 (Session 2 Spring 2024) at Duke Kunshan University.

Keywords: behavioral economics · irrational behavior · quasi-rationality · delayed gratification · impulsivity · foresight

1 Introduction

A rational person would always prefer a reward with larger utility to one with smaller utility, regardless of the temporal relation between the two. However, numerous research in behavioral science show that this is clearly not the case in reality, backed by evidence such as the famous "Stanford marshmallow experiment" on children [1]. Furthermore, experimental research also shows that this irrationality (and its relation with perceptual factors) influences decisions in game theory, such as in the trust game [2] or the ultimatum game [3]. While some current research exists on this sort of quasi-rational approaches, none have explored the implications this has by treating the present and the future separately.

In this work, we contribute to existing literature in game theory by proposing a method for investigating temporal factors quasi-rationality. We further introduce five new games that constructively illustrate our method's applications and implications.

* **Acknowledgments**: I am thankful to the instructor and classmates for pertinent discussions.

2 Formulation

2.1 A Tale of Two Agents

We now formally present the method for splitting a player into different agents. Let player i have the set of actions $A_i = \{a_n\}$, and utility functions $u_i(a_n)$. However, if we consider the player as two different agents i' and i'', both with the same set of actions, we have their own utility functions $u_i'(a_n)$ and $u_i''(a_n)$. They respectively represent the now-agent, or the current effective utility, and the then-agent, or the future utility.

Here, agent i' represents the present utility of player i. The different between player i and agent i' (the now-agent) is that the decisions of the now-agent would have to take the utility of the then-agent into consideration. Specifically, we define the following term in the utility:

Definition 1. *We define a "foresight function" $\phi(i', i'')$ to be a coefficient of a "foresight term," which represents the degree to which the then-agent impacts decisions of the now-agent. Hence:*

$$u_i'(a_n) = u_i(a_n) + \phi(i', i'')u_i''(a_n) \tag{1}$$

Hence, $\phi = 0$ represents a player who does not care about the future, $\phi \to \infty$ represents a player who only plans for the future, and $\phi = 1$ represents a rational player.

A Delicious Illustration In practice many decisions involve the disabling of a future option through current choices. For instance, in the Stanford marshmallow experiment[1], if the now-agent chooses "eat marshmallow," the then-agent would have no "eat marshmallow" choice. This is represented through a negative infinity utility, indicating unreachability. Hence we have a normal-form game (with full coordination and known choices, and hence only fixed equilibria are available), which we term the marshmallow game, as we show in Table 1.

Table 1. The marshmallow game, where the row player is the now-agent and the column player is the then-agent. $b > a$.

	Eat	Wait
Eat	$-\infty$, $-\infty$	a, 0
Wait	ϕb, b	0, 0

We see that when the then-agent chooses Wait there is a trivial Nash Equilibrium (Eat, Wait), which we name *indulgence*. If $\phi < a/b$, that would be the sole

[1] For the reader's information: the experiment leaves children in a room with a marshmallow, and if the children resist the temptation of that marshmallow, they will get an extra one as reward. It is a classical demonstration of self-control and delayed gratification.

Nash Equilibrium since the now-agent has no incentive to pick Wait over Eat. However, for $\phi > a/b$, we have another Nash Equilibrium, (Wait, Eat), which we name *delayed gratification.*[2]

2.2 Retirement as a Game

The marshmallow game serves as a toy example due to its simplicity: (Eat, Eat) and (Wait, Wait) are both Pareto inefficient, and hence will not occur in a game with full coordination and known choices. As a more nontrivial example, consider the following retirement game, which takes the unpredictability of future actions into account: we simplify life as three stages: young, midage, and old, and simplify economic decisions as three choices: save, no-op, and spend. Since the actions of the "old" stage is pretty much predefined (all choices except one will be Pareto inefficient), we ignore it and make "young" the now-agent and "midage" the then-agent.

In modeling, we let the savings status start from 0. The "save" action adds one to the status, "no-op" makes no changes, and "spend" subtracts one from the status. The actions respectively give utilities u_0, u_1, and u_2, where $u_0 < u_1 < u_2$, and during the old stage the utilities corresponding to the savings statuses of 0, 1, and 2 are respectively v_0, v_1, and v_2, where $v_0 < v_1 < v_2$. Foresight functions are ϕ_{ym}, ϕ_{mo}, and ϕ_{yo}. Given all the above, we have:

Table 2. The retirement game, where the row player is the now-agent (young) and the column player is the then-agent (midage).

	Save	No-op	Spend
Save	$(1+\phi_{ym})u_0 + \phi_{yo}v_2$, $u_0 + \phi_{mo}v_2$	$u_0 + \phi_{ym}u_1 + \phi_{yo}v_1$, $u_1 + \phi_{mo}v_1$	$u_0 + \phi_{ym}u_2 + \phi_{yo}v_0$, $u_2 + \phi_{mo}v_0$
No-op	$u_1 + \phi_{ym}u_0 + \phi_{yo}v_1$, $u_0 + \phi_{mo}v_1$	$(1+\phi_{ym})u_1 + \phi_{yo}v_0$, $u_1 + \phi_{mo}v_0$	$-\infty$, $-\infty$
Spend	$u_2 + \phi_{ym}u_0 + \phi_{yo}v_0$, $u_0 + \phi_{mo}v_0$	$-\infty$, $-\infty$	$-\infty$, $-\infty$

In this game the information is asymmetric: the column player knows the row player's action, but not the other way around (because "spend" or "no-op" decisions in the present does not force a future action). This can be implemented by first finding Nash Equilibria normally and then tossing out subgame imperfect outcomes. Obviously the agents' preferences now depend on the relations between the parameters. Mathematically, all three permutations of preferences are possible for both players.

[2] (Wait, Eat) is a Weak Nash Equilibrium if $\phi = a/b$, but in this work for the sake of simplicity we avoid discussions of borderline cases for ϕ.

The Samuelson Variant We can conduct simplification based on related assumptions. For instance, if we assume that utility decays exponentially with time (known famously as *exponential discounting*[3], as introduced by Samuelson [5]), we have that $\phi_{yo} = \phi_{ym}\phi_{mo}$. Then we can set the following:

$$x_{rs} = u_r + \phi_{mo} v_s \tag{2}$$

More generally we have the following useful conclusion:

Theorem 1. *Given exponential discounting and three temporal agents i', i'', and i''' (in that temporal order), we have $\phi(i', i''') = \phi(i', i'')\phi(i'', i''')$, and thus:*

$$u'_i(a_n) = u_i(a_n) + \phi(i', i'')(u''_i(a_n) + \phi(i'', i''')u'''_i(a_n)) \tag{3}$$

Proof. Given a discounted utility function $u(t) = u^* \exp(\pi t)$ (where u^* is the expected future utility, and π is the discount factor), we can directly deduce the foresight functions. We slightly abuse notation here by assuming i', i'', and i''' are linear timestamps. Thus, from (1) we have that $\phi(i', i'') = \frac{u'_i(a_n) - u_i(a_n)}{u''_i(a_n)} = \exp(\pi(i'' - i'))$. Completion of the proof is left as an exercise to the reader.

This gives us the following simplified game:

Table 3. The retirement game's Samuelson variant (assuming exponential discounting of utility and thus Theorem 1).

	Save	No-op	Spend
Save	$u_0 + \phi_{ym}x_{02}$, x_{02}	$u_0 + \phi_{ym}x_{11}$, x_{11}	$u_0 + \phi_{ym}x_{20}$, x_{20}
No-op	$u_1 + \phi_{ym}x_{01}$, x_{01}	$u_1 + \phi_{ym}x_{10}$, x_{10}	$-\infty$, $-\infty$
Spend	$u_2 + \phi_{ym}x_{00}$, x_{00}	$-\infty$, $-\infty$	$-\infty$, $-\infty$

This game can be further explored for specific parameters via computational approaches. We attach an implementation of a preliminary demo based on Nashpy, hosted on Google Colab[4].

3 Applications

3.1 The Ultimatum Game with Expected Requital (TIGER)

Now we demonstrate how this model can affect games with multiple players. Consider the Ultimatum Game with two players, A and B. It is well-known

[3] The recent consensus seems to be that utility discounts follow a hyperbolic pattern instead [4], which would be less mathematically convenient in our case, and we still adopt the exponential approach. The difference between the two models is out of the scope of this paper.

[4] *https://colab.research.google.com/drive/1-tCV1nPlWuQdbgknTMBTsqp58M-48VNT?usp=sharing*

that experimental results show the responder often acts irrationally by rejecting offers (which is not subgame optimal). Some have hypothesized that this is due to social factors [6], which relates to evolutionary processes based on repeated instantiations of the game. We create a minimum model of this by incorporating four steps: A as proposer, B as responder, B as proposer, and A as responder. We assume that A is a *Homo economicus*, and as a responder, would always accept the offer. However, we model B more realistically by assuming a mechanism of requital: if A made an unfair proposal, B might retailiate with an unfair proposal; and if A made a fair proposal, B might reciprocate with a fair proposal. Otherwise B is also rational and would accept all offers as the responder. We analyze the game from A's perspective, modeling B as a probabilistic agent but A as two temporal agents A' and A", as previously mentioned.

In this case, if we focus only on the subgame for the first iteration, A would always propose an unfair proposal. However, given information regarding B's action, A could sacrifice his current benefits by giving a fair proposal and expecting future payoffs. We thus have the TIGER as follows:

Table 4. The TIGER. The row player is A' and the column player is B. a is half of the utility split, and b is the amount the proposer keeps in an unfair proposal, so $a < b \leq 2a$. We do not consider temporal effects for B's utility.

	Fair	Unfair
Fair	$(1+\phi)a,\ 2a$	$(1+2\phi)a-\phi b,\ a+b$
Unfair	$b+\phi a,\ 3a-b$	$2\phi a+(1-\phi)b,\ 2a$

The sole Nash Equilibrium is (Unfair, Unfair). However, if we model requital by adding variability in B's response (analogous to trembling-hand perfect equilibria [7]), by giving a probabilty of $p > 0.5$ that B chooses the same action as A, we have different expectations for A. Now Fair gives expectation $p(1+\phi)a+(1-p)((1+2\phi)a-\phi b)$, while Unfair gives expectation $(1-p)(b+\phi a)+p(2\phi a+(1-\phi)b)$. Solving, we have the following equilibrium point for player A:

$$p = \frac{1+\phi}{2\phi} \tag{4}$$

If (4) holds then A is indifferent between the two choices. Hence we see that agents with $\phi > 1$ (high foresight) will choose Fair splits if they believe p is high enough, and this threshold for p decreases as foresight increases. Of course, real humans rarely value the future more than the present (have a ϕ higher than 1). However, if we tweak the payoffs by letting the future game contain more utility, even nearsighted agents can rationally choose the Fair split. This can be done by increasing the payoffs of the second iteration of the game, and is mathematically equivalent to increasing the foresight index ϕ. As in real life actions like ones in the ultimatum game often determine life-long reputations, and hence carry more weight than a transient benefit, this explains why our effective ϕ is often larger than 1, prompting us to make fair offers.

The conclusion of our analysis is that under our temporal model, a person who conducts requital (B) can influence a rational person (A) to choose fairness, even if the latter does not have preferences for requital. A rational agent with more foresight would be more prone to choose a Fair split, while a shortsighted choose Unfair splits, thus explaining experimental results such as Achtziger, Alós-Ferrer, and Wagner [8], which suggests people with more self-control (and hence presumably foresight) would give fairer proposals, and vice versa. This simultaneously explains results from [2] which suggestions trustworthiness (p) influences our temporal decisions.

3.2 The Joint Account Game

The TIGER only analyzes the actions of one agent, and does not change the Nash Equilibrium (although finding a way of circumventing it). We now propose a new two-player game with temporal factors to demonstrate the impact of our model on the Nash Equilibrium, and to split both players into agents. Our game is loosely inspired by the trust game.

Consider two players A and B. At the start of the game, both deposit one dollar into an account. They will choose to steal the two dollars or not. If neither steals the two dollars, they will be rewarded with $x+1$ dollars each. Else, if one chooses to steal he will have all, and if both steal they get their money back after some time (hence the temporal factor). We have the following vanilla payoff matrix:

Table 5. The vanilla joint account game. Row player is A and column player is B.

	Steal	Co-op
Steal	0, 0	1, -1
Co-op	-1, 1	x, x

This is a very flexible symmetric game that spans the 2×2 game taxonomy, depending on x. When $x < -1$ this is Prisoner's Delight, when $-1 < x < 0$ this is a Deadlock, when $0 < x < 1$ this is a Prisoner's Dilemma, and when $x > 1$ this is a Stag Hunt.

The key here is that since A and B might have different foresight coefficients, they can have different preferences towards the future. We hence have the following adjusted game:

Table 6. The joint account game, adjusted for temporal factors.

	Steal	Co-op
Steal	0, 0	1, $-\phi_B$
Co-op	$-\phi_A$, 1	$\phi_A x$, $\phi_B x$

From player A's perspective (player B is analogous), if player B steals A will also steal. If player B cooperates, A would steal if $\phi_A < 1/x$ and . As other cases are nonsensical or trivial, we further limit $x > 0$ and $\phi \in [0, 1]$. Hence we see the following cases, assuming complete information regarding utilities:

1. $\phi_A, \phi_B < 1/x$, in which case this is a Prisoner's dilemma, and the sole Nash Equilibrium is (Steal, Steal).
2. $\phi_A < 1/x < \phi_B$, in which case A will always steal, and B, believing A will steal, will also steal. This is the asymmetric Prisoner's Dilemma with the same sole Fixed-Strategy Nash Equilibrium.
3. $1/x < \phi_A, \phi_B$, in which case the game becomes a Stag Hunt again. Besides the same equilibrium of mutual defection, there now is a new equilibrium of mutual cooperation.

The difference of our case with any of the above is that the utilities of the other player are unknown. Hence we can scale this up to arrive at metagames for the now-agents:

Table 7. The joint account metagame. The row player is agent A' and the column player is agent B'. PD, aPD, and SH mean Prisoner's Dilemma, Asymettric PD, and Stag Hunt. Shown are the utilities of the Nash Equilibria.

	Small ϕ_B	Large ϕ_B
Small ϕ_A	PD 0, 0	aPD 0, 0
Large ϕ_A	aPD 0, 0	SH $\phi_A x$, $\phi_B x$ / 0, 0

Hence the conclusion is that this game benefits players with more foresight (larger ϕ) because that is the sole way of achieving a Nash Equilibrium that is not mutual-defection.

4 Conclusion

In this paper, we introduced a method for reconciling temporal effects to utility with traditional *Homo economicus* models. We created five completely games (the Marshmallow Game, the Retirement Game, the Retirement Game's Samuelson Variant, the TIGER, and the Joint Account Game) to demonstrate how this model can explain human behavior with temporal factors. Results show our method's explanatory power for human behavior involving expectations.

Bibliography

[1] W. Mischel, Y. Shoda, and M. L. Rodriguez, "Delay of gratification in children," *Science*, vol. 244, no. 4907, pp. 933–938, 1989. [Online]. Available: https://www.jstor.org/stable/pdf/1704494.pdf

[2] L. Michaelson, A. De la Vega, C. H. Chatham, and Y. Munakata, "Delaying gratification depends on social trust," *Frontiers in psychology*, vol. 4, p. 355, 2013. [Online]. Available: https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2013.00355/full

[3] E. Halali, Y. Bereby-Meyer, and A. Ockenfels, "Is it all about the self? the effect of self-control depletion on ultimatum game proposers," *Frontiers in human neuroscience*, vol. 7, p. 240, 2013. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnhum.2013.00240/full

[4] H. Ersner-Hershfield, G. E. Wimmer, and B. Knutson, "Saving for the future self: Neural measures of future self-continuity predict temporal discounting," *Social cognitive and affective neuroscience*, vol. 4, no. 1, pp. 85–92, 2009. [Online]. Available: https://watermark.silverchair.com/nsn042.pdf

[5] P. A. Samuelson, "A note on measurement of utility," *The review of economic studies*, vol. 4, no. 2, pp. 155–161, 1937. [Online]. Available: https://www.jstor.org/stable/pdf/2967612.pdf

[6] M. A. Nowak, K. M. Page, and K. Sigmund, "Fairness versus reason in the ultimatum game," *Science*, vol. 289, no. 5485, pp. 1773–1775, 2000. [Online]. Available: https://www.science.org/doi/full/10.1126/science.289.5485.1773

[7] R. Selten, *Reexamination of the perfectness concept for equilibrium points in extensive games.* Springer, 1988. [Online]. Available: https://link.springer.com/chapter/10.1007/978-94-015-7774-8_1

[8] A. Achtziger, C. Alós-Ferrer, and A. K. Wagner, "The impact of self-control depletion on social preferences in the ultimatum game," *Journal of Economic Psychology*, vol. 53, pp. 1–16, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167487015001476