# IBM Research Report 

# Utility Functions in Repeated Games 

Nimrod Megiddo<br>IBM Research Division<br>Almaden Research Center<br>650 Harry Road<br>San Jose, CA 95120-6099<br>USA

# Utility Functions in Repeated Games 

Nimrod Megiddo*

October 29, 2011


#### Abstract

The repeated games that are analyzed in game theory do not seem to exist in reality. The issue is that stages of a repeated game are played in the context of a histories of different lengths and hence the utilities of their outcomes may not be the same. The differences in utility have to do with several factors, including reduced risks, wealth accumulation, and changing attitudes toward other players. A new analysis is proposed for explaining cooperative behavior in the finitely-repeated prisoners' dilemma. A gift-giving optimization problem is introduced.


## 1 Introduction

Many real-life games are played repeatedly in the sense that the same situation of decision, conflict, or cooperation arises repeatedly with the same players, same strategy spaces, and same material outcomes. It seems that, in reality, repeated games are much more common than non-repeated ones, but many authors of applications papers treat such situations as one-shot games. In standard game theory, it is further assumed that in a repeated game the preferences are also the same each time the stage game is played, regardless of the past history of the play. Furthermore, the preferences are represented by utility functions, and the utility payoffs received in the stage games are aggregated in various ways into utility payoffs in the "supergame." In a finitely-repeated game, the most common aggregation rule is to add up the utilities of the stage games, with or without discount over time. In many laboratory experiments with repeated games, the players are given the same material payoff at each stage, and these payoffs are assumed to reflect the true utilities of the players. Interestingly, von Neumann and Morgenstern [4, p. 19] caution that "it would be an unnecessary complication, as far as our present objectives are concerned, to get entangled with problems of the preferences between events in different periods in the future" adding in a footnote: "It is well known that this presents very interesting, but yet extremely obscure, connections with the theory of saving and interest, etc." We question here the appropriateness of above-described approach. We discuss three challenges to the standard game-theoretic approach and explain them with simple examples. The challenges stem from the facts that when a game is played repeatedly, then (i) risk is reduced due to laws of large numbers, (ii) wealth can grow or shrink, depending on various choices made by players and Nature, and (iii) the attitude of one player toward another can change as a consequence of the choices the other has made, thereby affecting the player's preferences over both future and past ${ }^{1}$ material outcomes.

[^0]
## 2 Reduced risk

Utility functions that obey the axioms of von Neumann and Morgenstern [4] are aimed at reflecting the preferences of players with respect to risk. Arguably, these utility functions are exactly the ones that have the property that the utility from a probability distribution over material outcomes equals the expected value of the utilities of the outcomes. In other words, these utility functions are linear in terms of the probability values. The "expected-utility hypothesis" has been challenged by many people on various grounds, but this is not the issue on which we focus here.

Example 1. Consider a situation where a player has to decide whether or not to bet $\$ 100$ in a lottery, where he gets back either $\$ 250$ or nothing with equal probabilities. Denoting the von Neumann - Morgenstern utility function of the player by $u(\cdot)$, the player prefers to bet if and only if $\frac{1}{2} u(\$ 250)+\frac{1}{2} u(\$ 0)>u(\$ 100)$. Thus, some risk-averse players would prefer to keep the $\$ 100$ rather than bet on getting $\$ 250$ with probability $\frac{1}{2}$. If the offer to bet is presented repeatedly many times, and if we assume the player's utility function does not change during the play, then under the standard repeated-game approach, regardless of the number of opportunities to bet, the player either always bets or never bets, and that choice depends only on the comparison between a sure $\$ 100$ and $\$ 250$ with probability $50 \%$. This conclusion misses the important fact that risk is reduced when the bet may be made multiple times. If the player can repeat the same abovementioned bet multiple times, then the probability that he will not lose any money increases dramatically with the number of stages. The optimal choice would depend not only on the utility values of the outcomes $\$ 0, \$ 100$ and $\$ 250$, but also on many other combinations of these numbers. Suppose, for example, that player's utility function from money is $u(x) \equiv \ln (1+x)$ (assuming $x \geq 0)$. In this case, the player would not bet in the one-shot game because $\frac{1}{2} u(\$ 250)+\frac{1}{2} u(\$ 0)=2.76$, whereas $u(\$ 100)=4.62$. However, if the player has $n$ opportunities to bet, ${ }^{2}$ then for a large $n$ (e.g., $n>1000$ the probability distribution of the total return from the bets would be approximately Gaussian with expectation $125 n$ and standard deviation $125 \sqrt{n}$. Thus, with very high probability, the return would be greater than $125 n-500 \sqrt{n}$, which is much larger than the cost of the bets, namely, 100, $n$.

Thus, even a risk averse player would prefer to bet in the above-described example if he had many opportunities to bet the same way. Similarly, a risk-seeking player may prefer to bet in the example below but not too many times. The player would not be willing to bet many times if a loss becomes more certain.

Example 2. Suppose a player is offered to bet $\$ 100$ in a lottery that returns $\$ 300$ with probability $1 / 4$, and his utility satisfies $(1 / 4) u(\$ 300)+(3 / 4) u(\$ 0)>u(\$ 100)$. Thus, the player would like to bet in this lottery. However, if the game is repeated $n$ times, e.g., $n=1,000$, then with high probability the total return from betting $n$ times would be less than $75 n+200 \sqrt{n}$, whereas the total cost of the bets would be $100 n$, so the player would not like to bet $n$ times.

Thus, in general, when a risk-seeking player is offered to bet $n$ times, he may calculate an optimal number of bets $x$ as follows. Suppose the initial wealth of the player is $w$, the bet costs $b$ dollars and with probability $p$ it returns $R$ and zero otherwise. Denote the utility function from money by $u(\cdot)$. Denote by $U(x)$ the utility from $x$ independent bets with parameters $b, p$, and $R$. Thus,

$$
U(x)=\sum_{i=0}^{x}\binom{x}{i} p^{i}(1-p)^{x-i} u(w-x \cdot b+i \cdot R) .
$$

Hence, if the player were offered $n$ independent bets, he would pick $x \in\{0,1, \ldots, n\}$, so as to maximize $U(x)$.

Remark 1. In a strict sense, a single utility function has to be applied to the outcomes of one complete "game" rather than at each stage of the game, and it should also reflect the player's beliefs about the future

[^1]as we explain below. Suppose the game consists of a single bet. If the player believes that, with very high probability, he will later become richer independently of the current bet, then he may be willing to take a risk that otherwise he would currently not take. Similarly, if a utility function has to be applied to stages within a repeated game, then the number of stages, among other factors, has to be taken into account.

## 3 Wealth variation

In Section 2 we assumed, for simplicity, that the player decides in advance how many bets to take. Of course, in standard repeated games the player is informed of the payoff after each stage. The element of reduced risk exists even when the player is informed of past payoffs because when the player decides to bet or not to bet in the next stage, he knows that he will have more chances to bet in the future. However, in the betting examples of Section 2, if the player's decision to bet or not to bet in the next stage may depend the results of past bets, then there exists an important factor, which may affect the decision, namely, the current wealth of the player. A player may be more inclined to take risks bet after he has accumulated a large wealth as opposed to after he has lost substantial amounts. It is also possible sometimes that a player would take larger risks after having lost, hoping for better luck to cover the losses.

### 3.1 A simple 2-stage betting example

Example 3. Consider a 1-player game as follows. The player is given $\$ 200$ in cash and is asked to choose between two bets as follows. L: pay $\$ 100$ and win $\$ 300$ with probability $50 \%$ and zero otherwise, and H: pay $\$ 200$ and win $\$ 700$ with probability $50 \%$ and zero otherwise. Some risk-averse players may choose L even though the expected payoff from $H$ is higher, because by choosing $L$ they get at least $\$ 100$ for sure, whereas by choosing H they may get zero. Suppose the player plays the above described game twice. We argue that the choice of the player in the second round may depend on the result of the first round. The player may choose L in the first round, and if he wins the bet, having accumulated $\$ 400$, he may prefer H in the second round. On the other hand, if he loses in the first round he may prefer $L$ in the second round. The table below displays the possible strategies in the two-stage game, which are

LL Play L in both stages.
HH Play H in both stages.
LH Play L in one of the stages and H in the other one.
LHL Play L in the first stage and if a win, then play $H$ in the second stage; otherwise, play L in the second stage.

LLH Play L in the first stage and if a win, then play L in the second stage; otherwise, play H in the second stage.

HHL Play H in the first stage and if a win, then play H in the second stage; otherwise, play L in the second stage.

HLH Play $H$ in the first stage and if a win, then play $L$ in the second stage; otherwise, play $H$ in the second stage.

| Strategy | Expected Payoff | Possible payoffs, each with probability $1 / 4$ |
| :--- | :---: | :---: |
| LL | 500 | $(800,500,500,200)$ |
| HH | 700 | $(1400,700,700,0)$ |
| LH | 600 | $(1100,800,400,100)$ |
| LHL | 550 | $(1100,500,400,200)$ |
| LLH | 550 | $(800,800,500,100)$ |
| HHL | 650 | $(1400,700,400,100)$ |
| HLH | 650 | $(1100,800,700,0)$ |

The preference order over the probability distributions in the previous example is not necessarily the same as the order reflected by the expected dollar rewards, and the utility may depend also on the actual sequence of payments rather than just the total payment.

### 3.2 Multistage betting

Example 4. Suppose a player with utility function $u(x)=\ln (1+x)$ is offered repeatedly to choose one of the following:
(i) do nothing,
(ii) bet $\$ 100$ for a return of $\$ 150$ or $\$ 0$ with equal probabilities, or
(iii) bet $\$ 200$ for a return of $\$ 280$ or $\$ 0$ with equal probabilities.

If the wealth of the player is $\$ 1000$ and the bet is offered once, then the utility values of the three choices are, respectively, $6.91,6.93$, and 6.92 , hence the player chooses (ii). If the wealth is $\$ 3000$, and the bet is offered once, then the utility values of the three choices are, respectively, $8.007,8.014$, and 8.017 , hence the player chooses (iii). If the wealth is only $\$ 200$, and the bet is offered once, then the utility values of the three choices are, respectively, $5.30,5.24$, and 3.09 , hence the player choose (i). The last stage of a finitely repeated game is essentially a one-shot game. When the final offer to bet is made, the choice of the player depends on the wealth he has accumulated up to that point. In particular, if he has accumulated $\$ 3000$, then he will bet (iii). If he has accumulated $\$ 1000$, then he will bet (ii), and if his wealth shrank to $\$ 200$, then he will not bet at all.

Consider a finitely-repeated betting problem. Because the utility from money is generally a nonlinear function, the computation of a betting strategy that maximizes the utility from the final wealth requires more than just considering the next bet. The problem can of course be solved with dynamic programming (see also [2]). Suppose the player is offered to bet $n$ times and denote the stages by $1,2, \ldots, n$. Consider a simple problem where the player has an initial wealth of $w_{0}$, and for a fixed strategy of betting, denote the random variables of the wealth of various times by $W_{1}, \ldots, W_{n}$. The utility from the final wealth $W_{n}$ depends on the player's expectations from what happens after the betting is over. Denote by $u(w)$ the utility function for any possible final wealth $w$. Suppose at each stage the player can choose between doing nothing and betting $b$ dollars in a lottery that pays $R$ dollars with probability $p$ and zero otherwise. Thus, if at some stage $i$ the wealth satisfies $W_{i}<b$, then the player cannot bet any more and the final utility is $u\left(W_{i}\right)$.

For every stage $i$ and every possible $w \geq b$, denote by $u(i, w)$ the utility that the player has from having a wealth $w$ at the end of stage $i$. The option of playing $n-i$ more stages has some value, and the player can calculate the utility from this option by solving a dynamic programming problem as follows. optimal possible present-value utility from the final First, by definition, $u(n, w)=u(w)$. Obviously,

$$
u(n-1, w)= \begin{cases}u(n-1, w) & \text { if } w<b \\ \max \{u(n, w), p u(n, w-b+R)+(1-p) u(n, w-b)\} & \text { if } w \geq b\end{cases}
$$

Example 5. Consider a simple problem where the player has an integer initial wealth of $W$, and his utility from the final wealth $w$ is $u(w)=w^{2}$. Suppose at each stage the player can choose between doing nothing or betting one dollar, where the return is 2 with probability 0.45 and zero with probability 0.55 . If at some point the wealth $w$ is zero, then the player cannot bet any more and the final utility is zero. Suppose the player is repeatedly offered the same bet $n$ times. For every possible integer $w \geq 0$ and stage $i$, denote by $u(i, w)$ the utility from of a wealth $w$ after stage $i$. Denote $u(n, w)=u(w)=w^{2}$. Obviously,

$$
\begin{aligned}
u(n-1, w) & = \begin{cases}0 & \text { if } w=0 \\
\max \left\{w^{2}, 0.45(w+1)^{2}+0.55(w-1)^{2}\right\}=w^{2}+\max \{0,1-0.2 w\} & \text { if } w \geq 1\end{cases} \\
u(n-2, w) & = \begin{cases}0 & \text { if } w=0 \\
0.45\left(2^{2}-0.2 \cdot 2+1\right)+0.55 \cdot 0=2.07 & \text { if } w=1 \\
\max \left\{w^{2}, 0.45\left((w+1)^{2}-0.2(w+1)+1\right) 0.55\left((w-1)^{2}-0.2(w-1)+1\right)\right\} \\
=w^{2}+2 & \text { if } w \geq 2\end{cases} \\
u(i, 0) & =0 \\
u(i-1,1) & =\max \{u(i, 1), 0.45 u(i, 2)\} \\
u(i-1,2) & =\max \{u(i, 2), 0.45 u(i, 3)+0.55 u(i, 1)\} .
\end{aligned}
$$

### 3.3 Discounting

Discounting in repeated games is usually incorporated in order to represent either a loss of interest payment on delayed payoffs, or to account for the probability that the game will terminate. Thus, if a payment $R$ is actually paid $t$ time units later, and the discount factor is $\lambda(0<\lambda<1)$, then the "present value" of the payment is $\lambda^{t} R$. Next, suppose the interest rate is zero but the game may stop at some point, and the conditional probability that the game will continue to be played beyond stage $t$, given that it has been played until stage $t$, is equal to $\gamma$. In this case, the expected dollar present value of a payment $R$ at time $t$ is also equal to $\gamma^{t} R$. However, in terms of utility, these two models are quite different from one another. Suppose at each stage the player may bet $b$ dollars and then receive $R$ dollars with probability $p$, and zero otherwise. If $w \geq b$, then his wealth in stage 2 will be $w-b+R$ with probability $p$, and $w-b$ with probability $1-p$.

### 3.3.1 Discounting for interest rate

Consider first the interest-rate model and assume the game never stops. The dollar present value of a stream of payments $R_{1}, R_{2}, \ldots$ is equal to $\sum_{t} \lambda^{t} R_{t}$, and its present-value utility is equal to

$$
\rho\left(R_{1}, R_{2}, \ldots\right) \equiv u\left(\sum_{t=1}^{\infty} \lambda^{t} R_{t}\right) .
$$

Denote by $u(w)$ the present utility the player has from a wealth $w$ without the option of playing the infinite game. Denote by $U(w)$ the present utility from the same wealth together with the option of playing the infinite game. If the player does not bet in the first round, then the wealth is carried over to the next stage. If from that point and on an optimal policy is followed, then the utility $U(w)$ is equivalent to getting a dollar payment of $u^{-1}(U(w))$ at the beginning of stage 2 without the option of playing the infinite game. Because of the interest rate, the present value of that payment is $\lambda u^{-1}(U(w))$, hence the present utility of this option is equal to $u\left[\lambda u^{-1}(U(w))\right]$. On the other hand, if the player chooses to bet in the first round, and to subsequently follow an optimal policy, then his present utility is

$$
p U(w-b+R)+(1-p) U(w-b),
$$

which is equivalent to receiving a payment of

$$
u^{-1}[p U(w-b+R)+(1-p) U(w-b)]
$$

in stage 2, with a present value of

$$
\lambda u^{-1}[p U(w-b+R)+(1-p) U(w-b)]
$$

It follows that the optimal policy is characterized the following ("Bellman") equation:

$$
U(w)=\max \left\{u\left(\lambda u^{-1}[U(w)]\right), u\left(\lambda u^{-1}[p U(w-b+R)+(1-p) U(w-b)]\right)\right\}
$$

The effect of the discount factor depends on the kind of the player's utility function, namely, the relation of $u(\lambda w)$ to $u(w)$. Consider, for example, the function $u(w)=\ln w$. In this case,

$$
u(\lambda w)=\ln (\lambda w)=\ln \lambda+\ln w=\ln \lambda+u(w)
$$

and

$$
U(w)=\ln \lambda+\max \{U(w), p U(w-b+R)+(1-p) U(w-b)\}
$$

On the other hand, if $u(w)=w^{c}$, then

$$
u(\lambda w)=\lambda^{c} u(w)
$$

and

$$
U(w)=\lambda^{c} \max \{U(w), p U(w-b+R)+(1-p) U(w-b)\}
$$

Of course, for other kinds of utility functions $u(w)$ there may not be a clean relation like the above-stated ones but, in any case, a numerical solution should be possible.

### 3.3.2 Discounting for failure of the game to continue

If the interest rate is zero, and the game may stop as described above, then with probability $(1-\gamma) \gamma^{t}$ the game is played exactly $t$ stages, and the total payment is $\sum_{i=1}^{t} R_{i}$. The utility from the resulting probability distribution over payments is proportional to the expected value of the utility from the payments, e.g.,

$$
\sigma\left(R_{1}, R_{2}, \ldots\right) \equiv \sum_{t=1}^{\infty} \gamma^{t} u\left(\sum_{i=1}^{t} R_{i}\right)
$$

As before, denote by $U(w)$ the utility from a wealth of $w$ together with the option of playing the game until it terminates. If the player does not bet in the first stage, then the wealth is carried over to the next stage, and if from that point and on an optimal policy is followed, then the utility is $U(w)$. However, the conditional probability that the game will continue, given that it has not stopped is $\gamma$. If the game stops, then utility is $u(w)$. Thus, if the player does not bet in the first stage, the expected utility is equal to

$$
\gamma U(w)+(1-\gamma) u(w)
$$

If the player does bet in the first round, then the expected utility is equal to

$$
p[\gamma U(w-b+R)+(1-\gamma) u(w-b+R)]+(1-p)[\gamma U(w-b)+(1-\gamma) u(w-b)] .
$$

It follows that
$U(w)=\max \{\gamma U(w)+(1-\gamma) u(w), p[\gamma U(w-b+R)+(1-\gamma) u(w-b+R)]+(1-p)[\gamma U(w-b)+(1-\gamma) u(w-b)]\}$.
To simplify, denote for every $x$,

$$
H(x)=\gamma U(x)+(1-\gamma) u(x)
$$

Thus,

$$
\begin{aligned}
H(w) & =\gamma U(w)+(1-\gamma) u(w) \\
& =\max \{H(w), p H(w-b+R)+(1-p) H(w-b)\}
\end{aligned}
$$

Note that if the player does not bet, then

$$
H(w)=\gamma U(w)+(1-\gamma) u(w)=\gamma u(w)+(1-\gamma) u(w)=u(w)
$$

so, in particular, for $w<b, H(w)=u(w)$. If the player does bet, then

$$
H(w)=p H(w-b+R)+(1-p) H(w-b)
$$

Thus,

$$
H(b)=\max \{u(b), p H(R)+(1-p) H(0)\}=\max \{u(b), p H(R)\}
$$

With these boundary conditions, the entire function $H$ can be evaluated, and then $U$ can be evaluated.

### 3.4 Non-accumulable rewards

The notion of wealth variation does not seem to arise when the rewards cannot be accumulated. In the analysis above we assumed that the player receives dollar payoffs at each stage, the payoffs accumulate, and the player has a utility value of current wealth together with the future of the game. That model may not be appropriate for situations where the rewards, or portions thereof, cannot be accumulated. However, as we suggest later, emotions and feelings play an important role in the formation of preferences and hence utility functions. Therefore, even if the rewards have to be enjoyed or consumed at the time they are given, they may sometimes have a cumulative effect, which is not necessarily linear. For example, the player may have a utility function $u(m)$ where $m$ is the number of times the player receives the reward. Such a function may be too naïve though. On the one hand, the marginal value of a reward may be diminishing. On the other hand, if a person has got used to receiving (and consuming or enjoying) a certain reward every day, for example, a free cup of coffee every morning, then one denial of the reward may have a disutility that increases with the number of times the reward has been given.

## 4 Repeated games

The difficulties in formulating utilities in multistage decision problem are compounded once there are other players in the situation. As mentioned above, a repeated game in standard game theory is a multistage game where at each stage the players play the same basic game with the same utility payoffs, and the utilities from the individual stages are aggregated in a certain way into a utility from the play in the"supergame." We have argued that this definition may not apply if some player's utility function is not linear in terms of the amount of money. Therefore, we now consider repeated games, where the material payoffs are the same at each stage, but the utility function is not necessarily linear on terms of the amount of money. Suppose the material payoffs are made in dollars. We first consider games where the utility of each players depends only on his own total dollar payoff. Assume, for simplicity, that the stage game is a $2 \times 2$ bimtarix game, with dollar payoffs $a_{i j}$ to Player 1 and $b_{i j}$ to Player 2.

### 4.1 A finitely-repeated game

Suppose the game is played $n$ times, and the interest rate is zero. Denote by $W_{k}^{t}$ the wealth of player $k$ $(k=1,2)$ after stage $t(t=1, \ldots, n)$. Denote by $u_{k}(\cdot)(k=1,2)$ the utility function of Player $k$ from money.

At the last stage, $n$, the players essentially play a single $2 \times 2$ bimatrix game with matrices $A$ and $B$, where $A_{i j}=u_{1}\left(W_{1}^{n-1}+a_{i j}\right)$ and $B_{i j}=u_{2}\left(W_{2}^{n-1}+b_{i j}\right)$. The wealths $W_{k}^{n-1}$ depend on the past history of the play including the initial wealths $W_{k}^{0}$.

At stage $n-1$ the players essentially play a 2 -stage game as follows. At the first stage (i.e., stage $n-1$ ) if the strategy choices are $(i, j)$, then the wealths are updated in the following way:

$$
\begin{aligned}
& W_{1}^{n-1}=W_{1}^{n-2}+a_{i j} \\
& W_{2}^{n-1}=W_{2}^{n-2}+b_{i j}
\end{aligned}
$$

Suppose in stage $n$ Player 1 plays a mixed strategy $(x, 1-x)$, where $x$ depends on $j$, and Player 2 plays a mixed strategy $(y, 1-y)$ where $y$ depends on $i$. Then, the final wealths are the following:

$$
\begin{aligned}
& W_{1}^{n}=W_{1}^{n-1}+a_{I J} \\
& W_{2}^{n}=W_{2}^{n-1}+b_{I J}
\end{aligned}
$$

where $(I, J)$ are the random choices according to the mixed strategies, and the utilities are
$x y \cdot u_{1}\left(W_{1}^{n-1}+a_{11}\right)+(1-x) y \cdot u_{1}\left(W_{1}^{n-1}+a_{21}\right)+x(1-y) \cdot u_{1}\left(W_{1}^{n-1}+a_{12}\right)+(1-x)(1-y) \cdot u_{1}\left(W_{1}^{n-1}+a_{22}\right)$
and
$x y \cdot u_{2}\left(W_{2}^{n-1}+b_{11}\right)+(1-x) y \cdot u_{2}\left(W_{2}^{n-1}+b_{21}\right)+x(1-y) \cdot u_{2}\left(W_{2}^{n-1}+b_{12}\right)+(1-x)(1-y) \cdot u_{2}\left(W_{2}^{n-1}+b_{22}\right)$.
If there was an acceptable theory (for example, some notion of an equilibrium) of how the players would play at stage $n$, as a function of the wealths $W_{k}^{n-1}$, then it could be extended to suggest how they would play at stage $n-1$, and this analysis could be continued by backwards induction to the start of the game.

### 4.2 An infinitely repeated game

Suppose the $2 \times 2$ game described above is repeated infinitely many times with a discount factor $\lambda$. Denote by $U_{k}(w)$ the utility of player $k$ from starting to play the game with a wealth $w$. If the players choose $(i, j)$ in the first stage then their updated wealths are $W_{1}^{1}=W_{1}^{0}+a_{i j}$ and $W_{2}^{1}=W_{2}^{0}+b_{i j}$. The utilities of having such wealths at the beginning of stage 2 , assuming the players will continue to play forever, are $U_{k}\left(W_{k}^{1}\right)$, $k=1,2$. They are equivalent to having at the beginning of stage 2 the wealths $u_{k}^{-1}\left[U_{k}\left(W_{k}^{1}\right)\right]$ and terminating the to play of the game. These equivalent wealth values at the beginning of stage 2 are in turn equivalent to wealths of $\lambda u_{k}^{-1}\left[U_{k}\left(W_{k}^{1}\right)\right]$ at the beginning of stage 1 , with utility values $u_{k}\left(\lambda u_{k}^{-1}\left[U_{k}\left(W_{k}^{1}\right)\right]\right)$.

## 5 Changing emotions and attitudes in repeated games

We argue that in a repeated game, the past history of the play affects the states of mind of the player when the decide what to do in the next stage. The issue arises even when there is only one decision maker and some choices are made by Nature, i.e., some random device. When there are additional players, the attitudes of players' towards each other are affected by past play. This issue has been pointed out a previous papers [3].

### 5.1 A loss after a win versus a win after a loss

Sequences of bets with the same total gain may have different utility values. Consider, for example, the sequence depicted in Figure 1. The gamble depicted in Figure 2 gives the same probability distribution


Figure 1: A random device decides which problem to present


Figure 2: A random device decides which problem to present
over the total monetary payoffs. However, the simplification offered by the latter eliminates the emotional differences. It seems that some human players feel better when they win back their previous losses than when they lose their previous gains.

First of all, the state a player changes during the play as a result of the play. One obvious example is that the wealth of the player varies in a sequence of gambles, and obviously that may change the amount a player is willing to pay for a bet that returns $\$ 1,000$ with probability 0.5 . Therefore, we argue that a repeated matching pennies game with a reward of $\$ 1,000$ in each stage is not a repeated game according to the standard definition of game theory.

### 5.2 Impact on other players

If one player's utility takes into account how a material outcome affects another player, then the utility of the first player may also depend on the past history of the play, especially the choices of the other player, which impact the attitude of the first player towards the other one. The general issue as it occurs in extensive-form game has been pointed out in [3]. Here, we examine it as it applies to repeated games. In particular, we demonstrate below that it may give rise to cooperative play in the finitely-repeated Prisoner's Dilemma.


Figure 3: One-sided gift giving game

### 5.3 Repeated gift-giving games

### 5.3.1 A one-sided game

Consider the trivial game described in Figure 3. In this game Player I is passive while Player II chooses whether or not to give a $\$ 100$ gift to Player I. We later expand the strategy space of Player I, but first consider Player I's utilities in a repeated version of this game. The situation after the game has been played, for example, ten times, can be nontrivial. Consider the various situations that may arise:
(i) The gift was given at each stage during the first ten stages. In this case, getting a gift at the 11th stage does not add to the overall utility as much as the utility of the gift when it was given the first time. On the other hand, the disappointment of not getting the gift at the 11th stage after having received it 10 times may be more significant than not getting the gift in the first stage.
(ii) No gift was given any stage during the first ten stages. In this case, the surprise of getting the gift in the 11th stage can be much more significant thane getting the gift in the first stage. However, not getting the gift at the 11th stage is much less significant than not getting it in the first stage.
(iii) The gift was given in stages 1 and 2 , but no gift was given during stages $3-10$. The surprise of getting the gift at the 11th stage could be different from the surprise of getting the gift at the 11th stage in case (ii).
(iv) The gift was given at stages $1,2,4,7,9$. In this case, Player I may believe than Player II has decided during the first ten stages whether or not to give a gift by tossing a fair coin. Therefore, Player I will not be too surprised or too disappointed at stage 11 .

Now, if Player II is not at all interested in giving any gift to Player I, than there is nothing interesting here. Otherwise, an interesting optimization problem of timing may arise if Player II has a limited amount of money to spend on these gifts and he is interested in maximizing the utility of Player I from the gifts.

### 5.3.2 A two-sided game

Next, consider the two-sided gift giving game described in Figure 4. Note that the total dollar payoffs are zero when no gifts are exchanged and when both Players give gifts. However, it is likely that players would have different utilities in these two situations. Also, when a player receives a gift and does not give one, his utility may be lower than when he receives a gift and gives one. Thus, there may be quite different versions of this game in terms of utility. One example is shown in Figure 5. When the game is player repeatedly, the formation of utilities can be much more complicated compared to the repeated one-sided game as the interpretation of the other player's past play has to be with respect to the player's own past play, giving rise to mutual expectations, surprises and disappointments.


Figure 4: A two-sided gift-giving game


Figure 5: A two-sided gift-giving game in utilities

### 5.4 Optimization of giving gifts

The utility from a stream of gifts may be viewed as the sum of utilities from surprises and disappointments. Given the history of nonnegative gift payments, $x_{1}, \ldots, x_{i-1}$, the utility of the next gift can be represented as

$$
u_{i}(x)=u_{i}\left(x ; x_{1}, \ldots, x_{i-1}\right)
$$

Thus the total utility in $n$ stages is equal to

$$
U\left(x_{1}, \ldots, x_{n}\right)=\sum_{i=1}^{n} u_{i}\left(x ; x_{1}, \ldots, x_{i-1}\right) .
$$

Given a budget $B$, we wish to maximize

$$
U\left(x_{1}, \ldots, x_{n}\right)=\sum_{i=1}^{n} u_{i}\left(x ; x_{1}, \ldots, x_{i-1}\right)
$$

subject to

$$
\sum_{i=1}^{n} x_{i}=B
$$

More specific utility function are of the kind described below. Suppose at stage $i$ the recipient has a level of anticipation $y_{i-1}$ and the utility from the gift $x_{i}$ depends on this level, for example,

$$
\begin{equation*}
u_{i}\left(x_{i}, y_{i-1}\right)=x_{i}-y_{i-1} \tag{1}
\end{equation*}
$$

or

$$
u_{i}\left(x_{i}, y_{i-1}\right)=\frac{1}{1+e^{-\left(x_{i}-y_{i-1}\right)}} .
$$

Suppose further that the next level of anticipation is a function of the current level and the actual gift, for example,

$$
y_{i}=\alpha y_{i-1}+(1-\alpha) x_{i}
$$

If we start with $y_{0}=0$, we have

$$
\begin{equation*}
y_{i}=(1-\alpha) \sum_{j=1}^{i} \alpha^{i-j} x_{j} \tag{2}
\end{equation*}
$$

If we use the linear utility (1), we get the following optimization problem:

$$
\begin{aligned}
\text { Maximize } & \sum_{i=1}^{n} x_{i}-\sum_{i=1}^{n-1} y_{i} \\
\text { subject to } & \sum_{i=1}^{n} x_{i}=B \\
& (1-\alpha) x_{i}+\alpha y_{i-1}-y_{i}=0 \quad(i=1, \ldots, n) \\
& x_{1}, \ldots, x_{n} \geq 0 \\
& y_{0}=0
\end{aligned}
$$

However, by (2),

$$
\sum_{i=1}^{n} x_{i}-\sum_{i=0}^{n-1} y_{i}=\sum_{i=1}^{n} \alpha^{n-i} x_{i}
$$

so it follows that the optimal solution is $x_{n}=B$ and $x_{1}=\cdots=x_{n-1}=0$. A more interesting solution arises when we use for the utility the function

$$
\ell(x, y)=\frac{1}{1+e^{-(x-y)}}
$$

so that $u\left(x_{i}, y_{i-1}\right)=\ell\left(x_{i}, y_{i-1}\right)$, and we get the following optimization problem:

$$
\begin{aligned}
\text { Maximize } & \sum_{i=1}^{n} \ell\left(x_{i}, y_{i-1}\right) \\
\text { subject to } & \sum_{i=1}^{n} x_{i}=B \\
& (1-\alpha) x_{i}+\alpha y_{i-1}-y_{i}=0 \quad(i=1, \ldots, n) \\
& x_{1}, \ldots, x_{n} \geq 0 \\
& y_{0}=0
\end{aligned}
$$

If $\left(x_{1}, y_{1}, \ldots, x_{n}, y_{n}\right)$ is an optimal solution, then there exist multipliers $\lambda$ and $\eta_{i}, i=1, \ldots, n$, such that for every $i$,

$$
\begin{aligned}
& d_{i} \equiv \ell_{x}^{\prime}\left(x_{i}, y_{i-1}\right)=\frac{\exp \left\{-x_{i}+y_{i-1}\right\}}{\left(1+\exp \left\{-x_{i}+y_{i-1}\right\}\right)^{2}} \leq(1-\alpha) \eta_{i}+\lambda \\
& d_{i} \equiv \ell_{x}^{\prime}\left(x_{i}, y_{i-1}\right)=\frac{\exp \left\{-x_{i}+y_{i-1}\right\}}{\left(1+\exp \left\{-x_{i}+y_{i-1}\right\}\right)^{2}}=(1-\alpha) \eta_{i}+\lambda \quad \text { if } x_{i}>0 \\
& -d_{i} \equiv \ell_{y}^{\prime}\left(x_{i}, y_{i-1}\right)=-\frac{\exp \left\{-x_{i}+y_{i-1}\right\}}{\left(1+\exp \left\{-x_{i}+y_{i-1}\right\}\right)^{2}}=-\eta_{i-1}+\alpha \eta_{i} \quad i=2, \ldots, n
\end{aligned}
$$

Thus, for $i=2, \ldots, n$,

$$
\eta_{i}+\lambda \geq \eta_{i-1}
$$

and for $i \geq 2$ such that $x_{i}>0$,

$$
\eta_{i}+\lambda=\eta_{i-1}
$$

If $\alpha=1$ then the anticipation levels at both stage are equal and the optimal quantities are $x_{1}=x_{2}=B / 2$. For $\alpha=0.85$ and $B=1$, the optimal solution is $x_{1}=0.195, x_{2}=0.805$.

### 5.5 Repeated Prisoners' Dilemma

Thousands of papers have been written about the game of Prisoners' Dilemma (PD). In the PD game each of two player can either "cooperate" (C) or "defect" (D). The defect strategy is dominant in terms of material payoffs for both player if both players cooperate, then the material payoff to each of them is greater than the material payoff in case both defect. The analysis and experiments with the repeated PD assume at each stage the players receive utility payments according to the same game matrix, and these utilities are aggregated in some fashion. According to standard game-theoretical analysis, the only equilibrium in the finitely-repeated PD is the mutual repeated defect strategy. However, this conclusion is not supported by experiments [1]. Our opinion is that the difference stems from the difference in utilities in theory and practice. Consider the material payoff matrix in Figure 6. When this game is played twice, the second stage is played in the context of the play of the first stage. Therefore there are actually four types of games that could be played in the second stage. Thus, the two-stage PD can be depicted as in Figure 7.

The different possible plays in the first stage may evoke different emotional states for the player in the second stage, hence their utilities from a particular play in the second stage can be different, depending on the particular emotional state. For example, consider the two-stage game with utility payoffs that is depicted in Figure 8.


Figure 6: One-Shot Prisoner's Dilemma


Figure 7: Two-Stage Prisoner's Dilemma

If in the first stage the row player defects and the column player cooperates, then the row player may have some guilt feelings about it, while the column player may feel revengeful. Thus, the utility payoffs to the row player, from mutual cooperation in the second stage, after a play of $(D, C)$ and a after a play of $(C, D)$ in the first stage may be quite different. Similarly, the utility payoffs to the row player, from $(D, C)$ in the second stage, after a play of $(D, C)$ and a after a play of $(C, D)$ in the first stage may be quite different. In Figure 8 we indicated some possibilities of utility values in the second stage, depending on the play during the first stage. As indicated, there are multiple equilibria, and mutual cooperation in both rounds is one them. Unlike the standard model, there are even equilibria in mixed strategies. We do not claim that this example explains the finitely-repeated PD but rather that the assignment of utility values is far more complicated than contemplated in he standard theory.


Figure 8: Two-Stage Prisoner's Dilemma with utility payoffs

## References

[1] J. Anderoni and J. H. Miller, "Rational cooperation in the finitely repeated Prisoner's Dilemma: Experimental evidence," The Economic Journal 103 (1993) 570-585.
[2] L. E. Dubins and L. J. Savage, How th gamble if you must: Inequalities for Stochastic Processes, McGrawHill, 1965.
[3] N.Megiddo, "Formation of Preferences and Strategic Analysis: Can they be De-Coupled?" 9/24/2003. Presented at IMGTA'2001- XIV Italian Meeting on Game Theory and Applications. http://theory.stanford.edu/~megiddo/pdf/ver3.092403.pdf
[4] J. von Neumann and O. Morgenstern, Theory of Games and Economic Behavior, Princeton University Press, 1944, second ed. 1947, third ed. 1953


[^0]:    *IBM Almaden Research Center, San Jose, CA 95120. email: megiddo@us.ibm.com
    ${ }^{1}$ For example, a player may regret having given a gift to another player.

[^1]:    ${ }^{2}$ For simplicity, assume for a moment that the players has to decide in advance on how many bets to take.

