About coherence theorems

May 7, 2023

This document is in draft form. While it is meant to be readable and sufficiently clear as is, it does not reflect a complete or polished work.

1 Context

Consider the following claim: *Under very little general modelisation assumption, it is in the best interests of any agent playing a game to act as a bayesian utility maximizer. If it is not doing so, the agent could become strictly better by becoming a utility maximizer.*

It seems to me that this claim, or variations of it, is either alluded to or taken for granted in a lot of the literature on rational decision and game theory. Unless I am wrong this is notably the core point of [1].

I found some result that somewhat reflect this claim. But I found nothing that either proved something satisfyingly close or even made a precise corresponding claim. I suspect this is because I did not find it, rather than because it does not exist. But in any case I figure it is worthwhile to propose a mathematical formulation and study it for a bit.

In this document I will:

1. Offer a general model of game/action for a single player with perfect rule knowledge and perfect memory
2. Provide a conjecture that reflect the claim above, using the model as framework
3. Show this conjecture is false
4. List some variations of the conjecture that could be considered

2 Model

I will now give a model of action from an agent's perspective. The goal here is to provide a context to discuss the assumptions of game theory. Hence I want a model as unoriginal and generic as possible, with very few baked assumptions.

It will describe a single player game with discrete unbounded time. The choice of unbounded time is made out of care for generality; bounded game are simply a particular case of the model for unbounded games.

In keeping with game theory the rules of the game are perfectly known. All uncertainty, including about "how the world works" and about the other agents, is encapsulated in the set Ω of possible state. There "is" a true state w in Ω and all uncertainty is about which one it is. If the result of the game depends on random events, their result in also contained in advance in w. This causes no loss of generality.

Definition 1 (Game)**.** *A game $G = (\Omega, A, O, \preceq, O_d, I)$ is defined by*

Ω	*The set of all states. Possibly infinite.*
A	*The set of all actions. Always finite.*
O	*The set of outcomes.*
$\preceq$	*A partial ordering on O.*
$O_d : \Omega \times A^\omega \to O$	*An outcome function.*
$I : \Omega \times A^\star \to 2^\Omega$	*An information function.*

With:

$$I(w,u) = \{w' \in \Omega \mid I(w',u) = I(w,u)\} \qquad \forall (w,u) \in \Omega \times A^\star$$
$$I(w,u_1u_2) \subseteq I(w,u_1) \qquad \forall (w,u_1,u_2) \in \Omega \times A^\star \times A^\star$$

Definition 2. *A strategy is a function $s : 2^\Omega \times A^\star \to A$. We write S the set of all strategies.*

If we want to describe a "game run" (what happens when someone plays the game once), we need to know the game G, the initial state w, and the strategy s.
We start at turn 1. At each turn n we write u_n the series of actions already played at the begining of the turn. u_1 is empty and u_n is always of size $n-1$.

At each turn the player receives information. The information received at turn n is given by $I(w,u_n)$ (it depends only on the initial state and what was played before). Then the player chooses an action according to its strategy. This is given by $s(I(w,u_n),u_n)$ (it depends only on the information received and on the previously played actions, which the player remembers). This concludes turn n and allows the beginning of turn $n+1$, and so on and so forth.

Repeating this process infinitely many times yields an infinite series of actions $u_\infty \in A^\omega$. The outcome is then given by $O_d(w,u_\infty)$. The set of outcome is ordered by $\prec$. Not all two outcomes are comparable, but some are preferred to others. If $o_1 \prec o_2$ holds then o_2 is preferred to o_1.

We now fix a game $G = (\Omega, A, O, \preceq, O_d, I)$.

Definition 3 (Convenience function)**.** *$eval(s)$ is the function that associates to each hidden state the corresponding outcome. So $eval(s)(w)$ is the outcome obtained with strategy s on true state w.*

The two following conditions, taken from definition 1, might need some justification.

$$I(w,u) = \{w' \in \Omega \mid I(w',u) = I(w,u)\} \qquad \forall (w,u) \in \Omega \times A^\star \tag{1}$$
$$I(w,u_1u_2) \subseteq I(w,u_1) \qquad \forall (w,u_1,u_2) \in \Omega \times A^\star \times A^\star \tag{2}$$

They both seek to express the fact that the agent has perfect knowledge of the rules of the game and a perfect memory. It would have been possible for the information received by the agent to take some arbitrary value in a set of possible signals. But given that the only information that matters is the set of possible states compatibles with the information the agent received, we directly consider that to be the agent's information in the formalism.

Equation 1 signals that in all the states the agent considers possible, the agent's knowledge is equal to the one it presently has. Otherwise it could reason along the lines of "if I were in state w' I would think X, but I am not currently thinking X. Some I must not be in state w'."

Equation 2 signals that the agent can only ever eliminate possibilities, not add new ones. This is expected, considering that the output of I is the set of states that are compatible with what the agent knows. If a state was incompatible at some point, then it is forevermore true that it is incompatible with history (which the agent remembers perfectly).

3 Results, conjecture and counterexample

3.1 Domination and maximization

Definition 4 (Domination)**.** *A strategy s dominates a strategy t if and only if:*

$$\forall w \in \Omega,\ eval(t)(w) \preceq eval(s)(w)$$
$$\exists w \in \Omega,\ eval(t)(w) \prec eval(s)(w)$$

Definition 5 (Utility Maximisation)**.**
A strategy s is a utility maximization of probability P on Ω and utility U on O if and only if:

- *the utility satisfies $\preceq$: $\forall x, y.$ if $O_d(x) \prec O_d(y)$ then $U(x) < U(y)$.*
- *P is a probability function on Ω, taking 2^Ω as set of events.*
- *$s \in argmax_{t \in S} E_P[U(eval(t))]$.* [1]

[1] An implicit requirement this creates is that $E_P[U(\text{eval}(t))]$ be well defined for all t.

3.2 Lemmas

The "game player" we model is supposed to have a perfect memory. Hence if for two states w and w' there is a point in a game run at which they lead to different states of knowledge than the agent cannot consider them both possible afterward. At least one is incompatible with the past of the agent's knowledge. This was the reason for condition 2, lemma 1 prove an alternative mathematical formulation.

Lemma 1. *For all $w, w' \in \Omega$ and $u_1, u_2 \in A^\star$. If $I(w, u_1u_2) = I(w', u_1u_2)$ then $I(w, u_1) = I(w', u_1)$.*

Our definition of utility maximization only assumes that the strategy gives the best expected utility on the whole. It does not assume that each individual move is played through utility maximization. Indeed this is unnecessary as this property is a consequence of our definition. Lemma 2 shows this.

Lemma 2. *If s is a utility maximization of parameters P and U then $\forall w \in \Omega$, $\forall n \in \mathbb{N}$.*
If u_n is the series of actions taken up to step n according to s and $R : S \times A^\star \to \Omega \to O$ is the extension of eval that completes an evaluation for a given prefix.
Then $s(I(w, u_n), u_n) = argmax_{x\in A} E_{P|I(w,u_n)}[U(R(s, u_n a))]$.

Remark 1. *The converse is not true, which is a quirk of infinitely long games.*

3.3 Conjecture and counterexample

Conjecture 1. *Any strategy s that is not dominated is a utility maximization for some probability P on Ω and some utility U on O.*

Counterexample:
Take $\Omega = \mathbb{N}$ and $O = \{-1, 0, +1\}$, with $-1 \prec 0 \prec +1$. We write w the (unknown) hidden state ($w \in \Omega$). The game is as follows:
On the first turn you decide either STOP or INCREMENT. If you decide STOP you drop out of the game and get outcome 0 in all cases. Otherwise on each turn you have two actions: STOP or INCREMENT. Once you choose STOP the game ends (all future actions are irrelevant) and your total choice number T is equal to the number of times you picked INCREMENT. If you never pick STOP, T is arbitrarily set to 0. If $T = w$ you lose and get -1, otherwise you win and get $+1$.

This is all equivalent to you either refusing to bet (get outcome 0) or betting and choosing a number (any integer). When betting, if the integer you picked is equal to the hidden state you lose and get -1; otherwise you win and get $+1$.

Remark 2. *Nothing says how bad it is to get -1 or how good it is to get $+1$. Maybe -1 stands for "lose everything and die" and $+1$ for "get a cookie". The only constraint defined by the game is that getting $+1$ is better than getting 0, which is better than -1.*

In this game the strategy "refuse to play", (which I will write s) is not dominated. I will show it is not a utility maximization.

Proof. Let's assume that it is for some probability function P and utility U. We write $a = U(+1) - U(0)$ and $b = U(0) - U(-1)$. We only know that they are strictly positive. Because Ω is infinite there are values $n \in \Omega$ with arbitrarily small values for $P(\{n\})$. So there is some $n \in \Omega$ such that $P(\{n\})b < (1 - P(\{n\}))a$. The expected difference between s and the strategy s_n:"pick number n" is $(1 - P(\{n\}))a - P(\{n\})b$.
This is positive so s_n has a better expected utility than s. qed. □

Hence conjecture 1 is false.

4 Variations

A number of variations of the initial conjecture might be considered. Some of these are important because the initial conjecture's goal is to translate in a formal setting an initially general and vague claim. The variation of subsection 4.1, which moves from infinite games to finite games, is certainly the most obviously important one.

4.1 Known bound

There is a constant $k \in \mathbb{N}$ such that no action taken after step k has an impact on the outcome. More formally, $\forall u_1 \in A^k,\ \forall u_2, u_3 \in A^\omega,\ \forall w \in \Omega, O_d(w, u_1u_2) = O_d(w, u_1u_3)$. I do not know whether the conjecture holds in this case.

4.2 Unknown bound

restriction Only a finite number of actions will be taken in account, depending on $w \in \Omega$. Arbitrarily large bounds are possible. Main counterexample still applies.

4.3 Weaker domination

Weakening Replace domination with weaker domination, making it a stronger restriction that a strategy not be dominated.

Definition 6. *A strategy s weakly dominates a strategy t iff:*

$$eval(s)(w) \nprec eval(t)(w)\ \forall w \in \Omega$$
$$eval(t)(w) \prec eval(s)(w)\ \forall w \in \Omega$$

Main counterexample still applies.

4.4 More general utility

Allow, in the definition of U the case where $O_d(x) \prec O_d(y)$ and $U(x) = U(y)$. The question becomes trivial; set all utilities as equal and then all strategies are utility maximizations.

5 Notations

A^ω is the set of all infinite series of elements in A.
$A^\star$ is the set of all finite series of elements of A.
2^Ω is the set of all subsets of Ω.
$E_P[X]$ is the expected value of random variable X according to the probability law P.
If u_1 and u_2 are elements of $A^\star$ we write u_1u_2 the series obtained by concatenation of u_2 after u_1.

6 Related works

This section is for now wholly incomplete.

The post "There are no coherence theorems" [1] was the spark that decided me to think about this seriously. It seeks a theorem similar to the informal conjecture of section 1 and finds none.

Savage's theorem is somewhat similar to the conjecture at hand but makes too many assumptions, including some assumption of a probabilistic framework to begin with.

Wald's complete class theorem [2] only applies to a single step of a game and includes no modelisation of updates.

References

[1] EJT (pseudonyme). *There are no coherence theorems.* 2023. URL: https://www.lesswrong.com/posts/yCuzmCsE86BTu9PfA/there-are-no-coherence-theorems.

[2] Abraham Wald. "An Essentially Complete Class of Admissible Decision Functions". In: *The Annals of Mathematical Statistics* 18.4 (1947), pp. 549–555. DOI: 10.1214/aoms/1177730345. URL: https://doi.org/10.1214/aoms/1177730345.