Naiveté and sophistication in dynamic inconsistency

Preliminary version, please do not cite.

Zsombor Z. Méder[a], János Flesch[b], Ronald Peeters[a,*]

[a] *Department of Economics, Maastricht University, The Netherlands*
[b] *Department of Quantitative Economics, Maastricht University, The Netherlands*

Abstract

This paper introduces a general framework for dealing with dynamic inconsistency in the context of Markov decision problems. It carefully decouples concepts that are often entwined in the literature, distinguishing between the decision maker and its various temporal agents, and between the beliefs and intentions of the agents. Classical examples on naive and sophisticated decision makers are formalized and contrasted based on this new language. Providing a unified formalism to deal with these issues allows for the introduction of a mixed type of decision-maker, who is naive in some states and sophisticated in others. Such a mixed type can be used to model situations which were inaccessible to previous approaches.

Keywords:
JEL: C70, D11, D91

Contents

☆The Netherlands Organisation for Scientific Research (NWO) is gratefully acknowledged for its support.
*Corresponding author
Email addresses: `z.meder@maastrichtuniversity.nl` (Zsombor Z. Méder), `j.flesch@maastrichtuniversity.nl` (János Flesch), `r.peeters@maastrichtuniversity.nl` (Ronald Peeters)

1. Introduction

Imagine that you are down at the pub with a few friends. You’ve just finished your second pint, and your pals are ordering a new round. You think to yourself: “well, I could deal with one or two more, but then I really should go home.” However, you are also acutely aware that after your third beer you are likely to turn into a “just-one-more-beer” drinking machine. You’ve been down that road. You don’t want to go there. So you wisely (though, one may argue, prematurely) leave your friends at the pub after just your second beer. What is happening here?

Classically, there are three main approaches to time-inconsistent preferences. The first one regards decision makers as naifs (Akerlof, 1991), the second attributes sophistication to them (Laibson, 1997; Harris and Laibson, 2004; Fischer, 2001), while the third regards even resolute behavior possible (McClennen, 1990). A common assumption behind these models is that they regard the decision maker as falling *entirely* into one of the above three categories, treating his type as an exogeneously given natural condition.[1].

A way to interpret our example above is that you are sophisticated after drenching the first two beers, but you expect to become naive later on, as you drink more. Our example shows that in certain situations, a decision maker might change his type; moreover, he might even be able to reason about and calculate with such changes. The classical perspective that assumes the independence of types is unable to capture such situations. In order to remedy this shortcoming, this paper attempts a general, formal and precise characterization of naiveté and

[1]Even the use of terminology is incoherent: sometimes it is the *agent*, sometimes the *decision maker* who is regarded to be of one type or the other

sophistication for dynamic inconsistency. The language developed here can be used to capture the essential features of any naive or sophisticated decision maker with non-transitive time preferences, i.e. whose discount function is non-exponential. Our approach aims at a more precise and clear description of the observations in this field. This will in turn allow the introduction of mixed-type decision makers such as the one in our example.

After reviewing the relevant literature, we start by building a formalism that allows for precise definitions of the two most commonly discussed types of decision makers, naifs and sophisticates. We then compare these two types, and then finally expand the framework to include decision makers with a mixed type. In the concluding section, we will point towards further extensions of the model.

2. Related literature

Naiveté and sophistication imply crucially different behaviors. Fully naive decision makers often have wrong expectations about their own future behavior, which can obviously lead to inefficient decisions. It has, however been noticed, that from a welfare perspective, sophisticates can sometimes fare worse than naifs, even though they correctly anticipate future decisions (O'Donoghue and Rabin, 1999). In a scenario that involves punishment, sophistication can actually lead to worst possible outcomes (Heidhues and Kőszegi, 2009).

Moving away from pure types, a few authors have introduced models with less-than-complete sophistication. O'Donoghue and Rabin (2001) model a decision maker who is partially naive, in the sense that he believes he is going to be present-biased, but underestimates the extent of his future present-biasedness, and show that any degree of naiveté can generate arbitrarily large losses in efficiency for a decision maker.

A more recent attempt to introduce a limited foresight and control horizon can be found in Jehiel and Lilico (2010), who find that improving the length of foresight always improves the welfare of decision makers.

3. Basic concepts

3.1. Markov decision problem

We start with a decision maker facing a finite Markov decision problem on an infinite horizon. [2]

Definition 1. A finite Markov decision problem is given by:

- the set of time periods $T = \{1, 2, \dots\}$;
- a finite set of states Ω, with $\omega_1 \in \Omega$ as the initial state;
- a finite and non-empty set of pure actions A_ω that the decision maker can choose from in state ω;
- transition probabilities $m_\omega : A_\omega \to \Delta(\Omega)$, with $m_\omega(\omega'|a_\omega)$ denoting the probability to transit from state ω to state ω' when action a_ω is chosen.

[2]The latter is a weak requirement, since it is easy to rewrite a decision problem on a finite horizon to one on an infinite horizon.

- a payoff function $u_\omega : A_\omega \to \mathbb{R}$ that assigns a payoff to every action in state ω.

For simplicity, we do not allow for mixed actions in the first sections of the paper. In Sect. 7 we return to this issue.

3.2. History

To capture all the informational aspects on which an action choice can be conditioned, we introduce the notion of a history:

Definition 2. A history h has the form $h = (\omega_1, a_{\omega_1}, \dots, \omega_{t-1}, a_{\omega_{t-1}}, \omega_t)$, with:

- $\omega_i \in \Omega$, for $i \in \{1, \dots, t\}$;
- $a_{\omega_i} \in A_{\omega_i}$, for $i \in \{1, \dots, t-1\}$;
- $m_{\omega_{i-1}}(\omega_i | a_{\omega_{i-1}}) > 0$, for $i \in \{1, \dots, t-1\}$.

The function $\omega(h) = \omega_{t(h)}$ indicates the current or end state at history h. In a similar vein, the length of h or current time at h is denoted by $t = t(h)$. The set H_t is the collection of all possible histories of length t. We use H to refer to the set of all histories: $H = \bigcup_{t \in T} H_t$.

If history h' begins with h, we say that h' *succeeds* h, or equivalently, that h *precedes* h'; and denote this with $h \preccurlyeq h'$, or equivalently, with $h' \succcurlyeq h$.[3] If either $h \preccurlyeq h'$, or $h \succcurlyeq h'$, then we say that h and h' are *compatible*. The subset of H that contains histories that succeed h is denoted by $H^{\succcurlyeq h}$:

$$H^{\succcurlyeq h} = \{h' \in H | \, h' \succcurlyeq h\}.^4$$

We will refer to $h_0 = (\omega_0)$ as "root history", and the agent at that history as the "root agent".

3.3. Conceptual remarks

Models in the dynamic inconsistency literature come in two flavors (Asheim, 2007). One class of models is called *dual-self*, and their principal goal is to deal with the conflict of a long-run and short self. The other approach, also employed by this paper, is typically referred to as dealing with *multiple selves*. Here the focus is on the incentives that change over time.

We have to make a fundamental distincion between these multiple selves, which we dub "agents", and the notion of a "decision maker". The fundamental entities in our model are *agents*: they are the ones with the ultimate power of choosing an action *and* executing it a certain state. We assume a one-to-one correspondance

[3]Obviously, $h' \preccurlyeq h \;\&\; h \preccurlyeq h' \Leftrightarrow h = h'$.
[4]With this notation, we have $H = H^{\succcurlyeq \omega_1}$.

between histories and agents, since the information available at each of these histories is different, and the action choice might be contingent on such information. We refer to agents as being "at" a history, where they are endowed with preferences, and can relate to the future. As argued later, agents are endowed with beliefs and intentions about the future.

To refer to the unity and shared aspects of all temporal selves, we use the notion of *decision maker*. These shared aspects are threefold: first, they refer to the fact that the the decision problem is fundamentally the same. Second, they suggest that the preferences of various temporal selves are similar in a way, and a common functional form can be used to represent them (see Sect. 4). Third, it alludes to the elusive issue of *personal identity*, the fact that the connections between various temporal selves is substantively more intimate than the similarities between distinct individuals. We use the notion of a decision maker as comprising these three aspects, but carrying no normative weight: no preferences can be attributed to the decision maker, and it is not in conflict with any of its particular manifestations. Note that while we now build on the unity of temporal selves to outline our interpretation of the "decision maker", we have already implicitly used this term when we introduced Markov decision problems at the beginning of the section.

Returning our focus to agents, we note that although they are temporally distinct, they are clearly related strategically, since the well-being of each agent might depend on the actions taken by other agents. In the most general case, the well-being of each agent can be decomposed into three components: well-being generated by past actions, current actions and future actions.

Past actions are encoded in the history. In fact, our framework implies that there can be no differential effect of past actions on the well-being generated by current and future actions, i.e. for no two future series of actions and states would the effect of history be different for the current agent. In other words, we stick to the notion of "bygones are bygones". This might seem an obvious remark, but one could imagine otherwise. For example, if I am thinking about hiring a private investigator to find our whether my wife has been cheating on me, it's the utility generated by my (and, of course, her) past actions (e.g. not paying sufficient attention to her) that is potentially being re-evaluated.

As for the present, it is ordinarily assumed that the agent has full control over at least his current action. We stick to this assumption in the current paper.[5]

The third determinant of well-being can be the future actions of the decision maker. Using the distinction between experienced and decision utility (Kahneman et al., 1997), we can delimit two senses of "expected utility from taking action a". Let us disregard the immediate payoff for taking the action, and consider only future payoff. In one sense, the phrase could mean "experienced utility from expectation" i.e. utility that *is actually experienced* by the agent due to the fact that he is expecting to gain a certain stream of future payoffs. Think of a student that decides to study for an upcoming exam instead of watching his favorite TV show. He might, in fact, *already* enjoy the benefits of the decision to study (he is already less anxious for the exam, maybe he relishes

[5]To see that this choice is not so obvious, see Jehiel and Lilico (2010). Similarly, Elster's interpretation of the Ulysses story is an example of a model where conrol over current action is essentially eliminated (Elster, 1999).

the idea that he is doing “the right thing” etc.). The other sense of “expected utility” could be rendered as “the expected present value of various streams of payoffs” that the agent’s current decision can lead to. In this sense, the agent is either *unaware* of these, or, albeit he is able to calculate with them, he is not experiencing any of it.

This distinction translates well into the question of who the subject experiencing expected utility is. It is either the agent, or the decision maker. If agents are the subjects of experiencing utility, then the future utility is, in one sense, already present. In fact, in this case, it doesn’t really matter what future agents will *actually* do; all that is important is what the current agent believes they are going to do. In this case, it is irrelevant for the student whether he is really going to do study, or he just believes he will. Under this interpretation, expected utility is something attributable to an *agent*.

If, however, the real subject of experiencing utility is the decision maker, then the issue is what *really* happens, not what one believes will. In this case, the scope of experienced utility for the agents is limited to immediate payoffs. Instead, we should then talk about the expected utility of the *decision maker* at some history. In this case, since agents are expediters of the decision maker’s interests, any they are strategically related, they ought have some connection with the future. Minimally, agents should form intentions on future actions. Intentions on how to act in various future eventualities give practical reasons for taking this or that particular action. We assign these intentions to an agent at a particular history.

One can regard “expected utility” both ways. If beliefs lead *experienced* utility, then we just need to have very optimistic beliefs, and forming intentions are just means to specifying these optimistic forecasts. As we will see later, we can interpret naivetè this way. However, if beliefs are only relevant for *decision* utility, and conversely, if only the actual sequences of actions generate “experienced” utility, then one perhaps should attempt to have more realistic beliefs about his own future behavior, and shape his intentions accordingly; this is exactly what sophisticates do. As an example, think of our initial example. Obviously, a decison maker’s sober, tipsy and drunk selves will not only form different beliefs and intentions about the future, but will have a different propensity for wishful thinking.

In this paper we first attempt to model to give as a full description as possible for the two most prevalent decision maker types (naifs and sophisticates) before introducing mixed types. Therefore, we define a strategy as having three components: the current action, the intended future actions, and the belief on what future agents will in fact do. There is no special reason for assuming that the latter two coincide for future actions, although with our definitions, they will coincide for pure, but not for mixed decision maker types. To simplify notation, we reduce this triadic framework to just intentions and beliefs, and assume that for the current action, these two have to coincide: no agent can be wrong about which action he takes, and each agent takes the action that he intends to.

3.4. Intentions, beliefs, strategy

The basic building blocks of our model are all functions from the set of histories that succeed the agent to the set of available actions at those histories.

Definition 3. The *intentions* of an agent at history $\bar{h}$ assign an intended action to each present and future history:

$$i^{\bar{h}} : h \in H^{\succeq \bar{h}} \mapsto A_{\omega(h)}.$$

Definition 4. The *beliefs* of an agent at history $\bar{h}$ assign an action to each present and future history, specifying which action the agent at $\bar{h}$ beliefs the future agent at that history will choose:

$$b^{\bar{h}} : h \in H^{\succeq \bar{h}} \mapsto A_{\omega(h)}.$$

Definition 5. A *strategy* of an agent at history $\bar{h}$ is a pair of intentions and beliefs for that agent, with the added property that the belief and intention for the current action coincide:

$$s^{\bar{h}} =_{\text{def}} (i^{\bar{h}}, b^{\bar{h}}), \text{with } i^{\bar{h}}(\bar{h}) = b^{\bar{h}}(\bar{h}).$$

The set of all strategies for this agent is denoted by $S^{\bar{h}}$.

For an agent at $\bar{h}$, $s^{\bar{h}}(h)_i$ refers to the intention, while $s^{\bar{h}}(h)_b$ refers to the belief component of the strategy $s^{\bar{h}}$. For example, $s^{\bar{h}}(h)_b = a$ should be read as such: "The agent at $\bar{h}$ who holds strategy $s^{\bar{h}}$ believes he will choose action a at history h." Note that intentions and beliefs are defined at all succeeding histories, even at those that the agent does not intend or believe will be reached.

This approach brings us very close to a full epistemic characterization of intra-personal decision making. The full epistemic framework should include not only beliefs about the actions of future agents, but also beliefs about future agent's beliefs about future agent's beliefs etc. Moreover, it should also include beliefs about intentions, beliefs about beliefs about intentions etc. It is more controversial whether it should include intentions about intentions, intentions about beliefs,[6] intentions about beliefs about intentions, or any sequence of intention- and belief-operators, for that matter. Our current goal is just to provide an adequate characterization of naiveté and sophistication, and we can avoid delving into such details.

We can now define two properties of strategies, stationarity and coherence, as well as a relation over the set of strategies, consistency.

Definition 6. The intentions (beliefs) of an agent at $\bar{h}$ are *stationary*, whenever the intended (believed) actions depend only on the end-state. Formally, $s_i^{\bar{h}}$ (or $s_b^{\bar{h}}$) is called stationary if, for all h and h' with $\omega(h) = \omega(h')$, we have $s^{\bar{h}}(h)_i = s^{\bar{h}}(h')_i$ (or $s^{\bar{h}}(h)_b = s^{\bar{h}}(h')_b$) . A strategy $s^{\bar{h}}$ is stationary if both its constituent intentions $s_i^{\bar{h}}$ and beliefs $s_b^{\bar{h}}$ are stationary.

For example, if each day of the week can be modeled as a single state, the strategy of an agent who intends and believes eating in a restaurant every second Saturday, but staying home on every other one is not stationary.

[6]The toxin puzzle seems to indicate that there are some scenarios in which an agent might have an intention to form a future intentions i, but would be unable to ever form i. Also, this impossibility is open to him, so he would not believe he will form i (Kavka, 1983).

Definition 7. An agent at $\bar{h}$ is said to hold a *coherent* strategy, if, for all future histories, his intention and beliefs about future actions coincide. Formally, a strategy $s^{\bar{h}} = (i^{\bar{h}}, b^{\bar{h}})$ of an agent at $\bar{h}$ is coherent if $i^{\bar{h}}(h) = b^{\bar{h}}(h)$ for all $h \in H^{\succcurlyeq \bar{h}}$.

For example, a strategy of an agent who intends to stop drinking, but believes he will be unable to do so is not coherent.

Definition 8. The strategies of two agents at h and h' are said to be *consistent* if they assign the same intentions and beliefs to each history that succeeds both agents, i.e. s^h and $s^{h'}$ are consistent, if $s^h(h'') = s^{h'}(h'')$ for all $h'' \in H^{\succcurlyeq h} \cap H^{\succcurlyeq h'}$.[7]

For example, a strategy formulated yesterday which intended eating apples for today as desert, and a strategy formulated today that intends eating cookies instead are not consistent.

Whereas coherence concerns the relationship between the intentions and beliefs of the same strategy, i.e. belonging to one agent, consistency compares strategies of two distinct agents. In other words, coherence is an intrinsic, whereas consistency is an extrinsic (relational) property of a strategy – and thus, of an agent.

A natural question is whether consistency of strategies is transitive. If s^h and $s^{h'}$ are consistent, and $s^{h'}$ and $s^{h''}$ are also consistent, for some $h \preccurlyeq h' \preccurlyeq h''$, then s^h and $s^{h''}$ are also consistent. However, consistency is not transitive in general – it is not even transitive within the set of stationary strategies. To see this, take the decision problem in Fig. 1. We will construct three stationary strategies s^h, $s^{h'}$, and $s^{h''}$ such that the first and last two are consistent, but the first is not consistent with the last. Fix $h = (\rho)$, $h' = (\rho, A, \sigma)$ and $h'' = (\rho, B, \gamma)$. Also, let $s^h(h) = (A, A)$, $s^h(h''') = (C, C)$ if $\omega(h''') = \sigma$, and $s^h(h''') = (E, E)$ if $\omega(h''') = \gamma$. Intuitively, s^h means: "go left and choose C, but if you ever end up at γ, choose E". Define two other strategies through $s^{h'}(h''') = (C, C)$ for all h''' ("do C once in σ"), and $s^{h''}(h''') = (F, F)$ for all h''' ("do F once in γ"). All of these strategies are stationary. Clearly, s^h and $s^{h'}$ are consistent, since they both require the decision maker to choose C in state σ, and after history h', no state other state than σ is reachable. Next, $s^{h'}$ and $s^{h''}$ are consistent, since histories h' and h'' are not compatible. But s^h and $s^{h''}$ are not consistent, as they assign different actions to the state γ. This shows that consistency of strategies is not transitive on the set of stationary strategies.

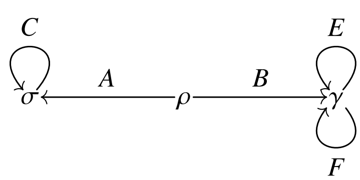


Figure 1: Stationary strategies with intransitive consistency

[7]So, if strategies are defined at histories that are not compatible, the respective strategies are consistent (since there are *no* histories that succeed both).

3.5. Rootcutting

The following definitions of "rootcutting", although they look like technicalities, are necessary for the definition of a stationary plan. Rootcutting formalizes the idea of "bygones are bygones", and chips away everything from the history but the final state.

Definition 9. Take any history $h = ((\omega_j, a_{\omega_j})_{j=1,\ldots,t-1}, \omega_{t(h)})$. The *rootcutting* operator $|_k$, defined for any $k \leq t(h)$, removes the first $k-1$ pairs of this sequence, so $h|_k = ((\omega_j, a_{\omega_j})_{j=k,\ldots,t-1}, \omega_t)$. Specifically, $h|_{t(h)} = (\omega_{t(h)}) = (\omega(h))$, and $h|_1 = h$. For a set of histories $H' \subseteq H$, $H'|_k$ will refer to the set of rootcut histories.

A rootcut strategy applies this idea to strategies: it is defined at an agent that forgot some its past, and on future histories obviously do not include the descriptions of the forgotten segments of the past anymore (i.e. on histories in $H^{\succeq \bar{h}}|_k$).

Definition 10. For any strategy $s^{\bar{h}}$, the *rootcut* strategy $s^{\bar{h}}|_k : h \in H^{\succeq \bar{h}}|_k \mapsto A_{\omega(h)}$ denotes the function for which $s^{\bar{h}}|_k(h|_k) = s^{\bar{h}}(h)$, for all $h \in H^{\succeq \bar{h}}$.

To see this definition at work, think of a resolve to stop smoking on the first day of the next month. Take an agent that makes this resolve on July 25th, fails, but makes the resolve again on August 25th. If we interpret the resolve as a strategy, it is easy to see that they are not consistent: they prescribe different smoking behavior for instance, for August 28th – the first strategy forbids it, while the second allows it. However, there is an intuitive sense in which they are very similar. Indeed, they map them into the same resolve that uses indexicals instead of precise dates: "I can smoke for one more week, and then I will stop". Rootcutting the present history achieves this role by getting rid of the past. In our example, the original strategies are not identical or consistent; but their rootcut strategies are identical.[8]

3.6. Plans

In this subsection, we will move from the agent level to the level of the decision maker. Since there is no *a priori* reason for the agents to have consistent strategies, different agents can form different intentions and entertain different beliefs about any certain future agent. To have an "external" overview of all agents, we introduce the concept of a plan. In our terminology, a plan is an auxiliary tool for examining the strategies of all possible agents, and not something that is intentionally put together by the decision maker. Whereas each agent chooses a strategy, the decision maker does not choose a plan. Instead, a plan contains a full description of the intentions and beliefs in all contingencies, i.e. at all histories.

Definition 11. A *plan* is a function $p : h \in H \mapsto S^h$. Intuitively, a plan assigns a strategy to each agent.

[8]One can also think of the equivalence of rootcut strategies as a weaker form of consistency of stationary strategies: when two rootcut strategies are identical, they specify the same beliefs and actions to histories which can be reached via the same sequence of states and actions.

Fig. 2 shows an extremely simple decision problem, for which an example of a plan is represented in Table 1. Each entry is a pair of A's and B's, an intended action and a belief about an action. Each row corresponds to a strategy for an agent at $\bar{h}$, defining an intention and a belief for each history that succeeds $\bar{h}$. For example, in our table, the entry AB for row $\bar{h} = \rho A \rho$ and column $h = \rho A \rho A \rho$ should be interpreted as such: the agent at $\rho A \rho$ intends to choose action A at history $\rho A \rho A \rho$, while believing the agent at $\rho A \rho A \rho$ will, in fact, choose action B. The whole plan thus specifies the intentions and beliefs of all agents over all other (present and future) agents. Our definition of a strategy ensures that the "diagonal" of the table contains identical acions, i.e. $p(h)(h)_i = p(h)(h)_b$ for all h.

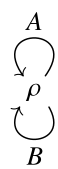


Figure 2: A basic decision problem

		$h \in H^{\succcurlyeq \bar{h}}$					
		ρ	$\rho A \rho$	$\rho B \rho$	$\rho A \rho A \rho$	$\rho A \rho B \rho$	...
$\bar{h}$	ρ	BB	AB	AB	AA	BB	...
	$\rho A \rho$	—	BB	—	AB	BA	...
	$\rho B \rho$	—	—	AA	—	—	...
	$\rho A \rho A \rho$	—	—	—	BB	—	...
	$\rho A \rho B \rho$	—	—	—	—	AA	...
	...	—	—	—	—	—	...

Table 1: An example of a plan for the decision problem in Fig. 2

We now proceed to introduce three properties of plans. Our definition of stationarity makes use of the root-cutting operator, defined above.

Definition 12. A plan p is said to be *stationary*, if only the end-state matters when assigning strategies to histories, i.e. if, for any histories h and h', if $\omega(h) = \omega(h')$, then $p(h)|_{t(h)} = p(h')|_{t(h')}$.

Stationarity of a plan is different from the stationarity of the strategies involved. For the decision problem on Fig. 2, Table 2 offers a non-stationary plan of stationary strategies, while Table 3 shows a stationary plan of non-stationary strategies.

For a concrete example of a stationary plan (of nonstationary strategies), think of the decision maker who, waking up every day, decides to take *just one more* shot of heroin, and intends (and believes) to quit the next day. However, if this decision maker ever chooses to quit immediately, his plan would not be stationary anymore.

		$h \in H^{\succcurlyeq \bar{h}}$					
		ρ	$\rho A \rho$	$\rho B \rho$	$\rho A \rho A \rho$	$\rho A \rho B \rho$	...
$\bar{h}$	ρ	AA	AA	AA	AA	AA	AA
	$\rho A \rho$	—	BB	—	BB	BB	BB
	$\rho B \rho$	—	—	AA	—	—	AA
	$\rho A \rho A \rho$	—	—	—	AA	—	AA
	$\rho A \rho B \rho$	—	—	—	—	BB	BB
	...	—	—	—	—	—	...

Table 2: Non-stationary plan of stationary strategies

		$h \in H^{\succcurlyeq \bar{h}}$					
		ρ	$\rho A \rho$	$\rho B \rho$	$\rho A \rho A \rho$	$\rho A \rho B \rho$	...
$\bar{h}$	ρ	BB	AA	AA	AA	AA	AA
	$\rho A \rho$	—	BB	—	AA	AA	AA
	$\rho B \rho$	—	—	BB	—	—	AA
	$\rho A \rho A \rho$	—	—	—	BB	—	AA
	$\rho A \rho B \rho$	—	—	—	—	BB	AA
	...	—	—	—	—	—	...

Table 3: Stationary plan of non-stationary strategies

Next, we define consistency of a plan. The intuitive idea is that a plan is consistent if no deviation can be expected from previous intentions and beliefs.

Definition 13. A plan p is said to be *consistent* if the strategies $p(h)$ and $p(h')$ assigned to any two histories h and h' are consistent.

Consistency is a very strong notion: a consistent decision maker would never change his mind about any action choice, whenever (at whichever history) he is contemplating it. An example would be the heroin user who goes cold turkey immediately and definitely, never ever thinking to restart his substance use.

According to this definition, if p is a consistent plan, then we get $p(h)(h'') = p(h')(h'')$ whenever h and h' are compatible, and $h'' \succcurlyeq h$, and $h'' \succcurlyeq h'$. Note that a consistent plan is necessarily made up by coherent strategies. Thus, an action choice for all histories uniquely determines a consistent plan. Similarly, a choice of an action for all states uniquely determines a consistent plan of stationary strategies.

Theorem 1. *A consistent, stationary plan consists of stationary strategies.*

Proof. Take any histories h, h' and h'' for which $\omega(h') = \omega(h'')$. We have to show that $p(h)(h') = p(h)(h'')$. For this, see that:

$$p(h)(h') = p(h')(h') = p(h')|_{t(h')}(h'|_{t(h')}) = p(h'')|_{t(h'')}(h''|_{t(h'')}) = p(h'')(h'') = p(h)(h'').$$

For the respective equations, we use, in order, consistency, definition of rootcutting, stationarity of the plan, definition of rootcutting and consistency again. □

Theorem 2. *A consistent plan of stationary strategies is a stationary plan.*

Proof. Take two histories h and h', with $\omega(h) = \omega(h')$. We have to show that $p(h)|_{t(h)} = p(h')|_{t(h')}$. Since the strategies are stationary, their rootcut versions are stationary, too: $p(h)|_{t(h)} = s^h$ and $p(h')|_{t(h')} = s^{h'}$, where s^h and $s^{h'}$ are stationary. Since $p(h)$ and $p(h')$ are consistent, $p(h)|_{t(h)}$ and $p(h')|_{t(h')}$ are also consistent, so $s^h = s^{h'}$. So $p(h)|_{t(h)} = p(h')|_{t(h')}$. □

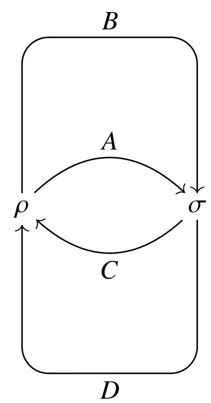


Figure 3: Stationary plan of stationary strategies is not consistent

Based on the two theorems above, one might expect that a stationary plan of stationary strategies would be consistent. However, this is not so, as can be seen from the following example. Consider the Markov decision problem on Fig. 3. Let:

$$s_1^{\bar{h}}(h) = \begin{cases} (A, A) \text{ if } \omega(h) = \rho, \\ (C, C) \text{ if } \omega(h) = \sigma, \end{cases} \quad \text{and} \quad s_2^{\bar{h}}(h) = \begin{cases} (B, B) \text{ if } \omega(h) = \rho, \\ (D, D) \text{ if } \omega(h) = \sigma. \end{cases}$$

Now, let us define a plan p, so that:

$$p(h) = \begin{cases} s_1^{\bar{h}} \text{ if } \omega(\bar{h}) = \rho; \\ s_2^{\bar{h}} \text{ if } \omega(\bar{h}) = \sigma. \end{cases}$$

This is obviously a plan of stationary strategies. It also is a stationary plan, since only the end-state matters in assigning a strategy to a history, according to the definition. However, it is not a consistent plan, since:

$$p(\rho)(\rho A\sigma) = s_1^{\rho}(\rho A\sigma) = (A, A) \neq (B, B) = s_2^{\rho A\sigma}(\rho A\sigma) = p(\rho A\sigma)(\rho A\sigma).$$

3.7. *Induced strategy*

We assume that, since each agent has control over his current action (and only that), the actual actions executed by each agent h can be obtained from plan p by looking at $p(h)(h)$, in other words, from the diagonal of the plan.

Definition 14. The *induced strategy* of a plan p specifies the actual actions chosen by each agent:

$$\Lambda(p) : h \in H \mapsto A_{\omega(h)}, \text{ given by } \Lambda(p)(h) = p(h)(h).$$

It is handy to define for some plan p, and an agent at $\bar{h}$, the induced strategy for the (present and) future:

$$\Lambda^{\succeq \bar{h}}(p) : h \in H^{\succeq \bar{h}} \mapsto A_{\omega(h)}, \text{ given by } \Lambda^{\succeq \bar{h}}(p)(h) = p(h)(h);$$
$$\Lambda^{\succ \bar{h}}(p) : h \in H^{\succeq \bar{h}} \setminus \{\bar{h}\} \mapsto A_{\omega(h)}, \text{ given by } \Lambda^{\succ \bar{h}}(p)(h) = p(h)(h).$$

The induced strategy of the plan represented in Table 1 is $\Lambda(p)(\rho) = (BB)$, $\Lambda(p)(\rho A \rho) = (BB)$, $\Lambda(p)(\rho B \rho) = (AA)$ etc. Alternatively, we can write out $\Lambda(p)$ in a more simple form, as $(BB)(BB)(AA)(BB)(AA)\ldots$.

4. Utility and discounting

4.1. *Payoffs, horizon and utility*

The term *payoff*, introduced in Def. 1, refers to the immediate gains or losses resulting from an action. Formally, a payoff gained in period t is denoted by u_t. Payoffs are fully determined by the decision problem, the state and the action taken. However, time preference implies that identical payoffs might be regarded differently by various agents, based on the temporal distance between the agent and the payoff in regard.

We assume that for each agent, they integrate present and future payoffs in such a way that their preferences over outcomes respect First and Second Order Separability. The former implies "there is no interaction between the effects of payoffs of different periods", while the latter "isolates the impact of time into decision weights" (Lapied and Renault, 2012). Then, if agents are impatient, it is possible to represent their preferences with a linearly separable utility function. In this paper, we will use a linearly separable utility function of a particular form, namely, quasi-hyperbolic discounting.

$$U^h(u) = u_{\underline{t}} + \beta \sum_{t=\underline{t}}^{\infty} \delta^{t-\underline{t}} u_t,$$

with $0 \leq \beta \leq 1$ and $0 < \delta < 1$. We will point to the possibility of relaxing this discount function in the concluding section.

4.2. "Expected utility"

In Sect.3.3, we distinguished between two senses of the term "expected utility": utility *actually experienced* from expecting a future payoff stream, or "expected utility" as simply *a means of calculating* with various future courses of action. This distinction is formally nailed down and further refined by the following definitions.

Each agent attempts to control the future stream of payoffs, by deciding not only a current action, but devising an entire course of action for future eventualities. Whenever an agent is contemplating possible courses of action, he is calculating with how he *intends* to play in the future.

Definition 15. The *expected utility based on intentions* of playing strategy s^h for an agent at h is:

$$U_i^h\left(s^h\right) = \mathbb{E}\left[s_i^h\right]\left(U^h(u)\right).$$

On the other hand, whenever an agent is reflecting on how much utility he can reasonably expect, he will calculate his utility based on his beliefs.

Definition 16. The *expected utility based on beliefs* of playing strategy s^h for an agent at h is:

$$U_b^h\left(s^h\right) = \mathbb{E}\left[s_b^h\right]\left(U^h(u)\right).$$

Neither his intentions, nor his beliefs determine the *real* utility of an agent. When calculating his real (expected) utility, the sole thing that matters is which actions future agents will actually implement under various eventualities. The definition of induced strategy captures just this, and can thus be used to define *real expected utility*:

Definition 17. Given a plan p, the (ex post) *realized expected utility* of an agent at h is:

$$U_r^h(p) = \mathbb{E}\left[\Lambda^{\succcurlyeq h}(p)\right]\left(U^h(u)\right).$$

Notice that traditionally, the above three meanings of the term "expected utility" coincide. The reason is that where dynamic inconsistency does not pose a problem, intentions and beliefs on future actions coincide; moreover, the decision maker always executes the intentions of past agents.

5. Naiveté

5.1. The meaning of naiveté

Naiveté has been characterized in several ways in the literature. All of the following are valid descriptions of naif agents:

A. Naifs believe that their preferences won't change (whereas they do).

B. Naifs believe that they won't adopt new strategies (but they will).

C. Naifs believe that – while their preferences might change – they can commit to a strategy chosen at this moment (albeit they can't).

It is not even clear whether naiveté is a property of the decision maker or that of an agent. In this and the following section, we will define naiveté and sophistication for agents, but will assume that the decision maker is always naive (or sophisticated). We will return to the issue presented in Sect.1 after analysing and contrasting these base cases.

The common aspect of the characterizations above is that something is wrong with the beliefs held by the agent. We would like to suggest that these troubles arise from the way the naif determines its beliefs: particularly, that for a naive agent, his current preferences determine his intentions, which in turn determine his beliefs on future actions. Thus, it does not matter whether the agent holds an explicit belief on the lack of change in his preferences (as in case A), or whether he beliefs he will simply fail to act on such changes (case B), or that he has strong beliefs in his own will- or pre-commitment power (case C). The essential features of naiveté are the directions of determination seen on Fig. 4. All the above cases are described by this model.

Definition 18. A strategy $\tilde{s}^h$ of an agent at h is *naively optimal*, if it maximizes expected utility based on intentions on the complete strategy space and it is coherent :

$$\tilde{s}^h \in \arg\max_{s \in S^h} U_i^h(s),$$

and

$$\tilde{s}^h(h')_b = \tilde{s}^h(h')_i, \text{ for all } h' \in H^{\succcurlyeq h}.$$

A plan $\tilde{p}$ is naive if at each history h, the strategy $\tilde{p}(h)$ is naively optimal.

Thus, naiveté is primarily a property of a strategy (or an agent holding that strategy). We talk about a naive decision maker if the plan describing him is naive. Note that we require $\tilde{s}^h$ to be optimal at all histories, not only on the set of states that are reachable as the induced strategy is executed.

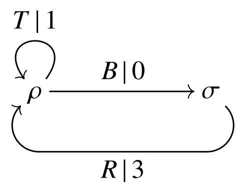
Figure 4: The forming of intentions and beliefs of a naif.

Preferences → intentions → beliefs.

5.2. Properties of naive plans

For $\beta \in (0, 1)$, there are Markov decision problems for which there is no consistent naive plan. Fig. 5.2 shows a decision problem which generates dynamic inconsistency for $\beta = 0.5$.

Strotz (1956) shows that, whenever that only when the discount function is exponential does the decision maker have a consistent naive plan for all decision problems.



$\beta = 0.75 \;\; \delta = 0.75$

Figure 5: Dynamic inconsistency in naiveté

		$h \in H^{\succcurlyeq \bar{h}}$					
		ρ	$\rho T \rho$	$\rho B \sigma$	$\rho T \rho T \rho$	$\rho T \rho B \sigma$	...
$\bar{h}$	ρ	[TT]	BB	BB	BB	BB	...
	$\rho T \sigma$	—	[TT]	—	BB	BB	...
	$\rho B \sigma$	—	—	[TT]	—	—	...
	$\rho T \rho T \rho$	—	—	—	[BB]	—	...
	$\rho T \rho B \sigma$	—	—	—	—	[TT]	...
	...	—	—	—	—	—	...

Table 4: Naive plan for the decision problem on Fig. 5.2

Theorem 3. *For any decision problem, there exists a stationary naive plan.*

Proof. For each h, the set $\max_{s \in S^h} U^h_{T,i}(s)$ is non-empty, because S^h is non-empty and closed for pointwise limits. Therefore, the set of strategies where the maximum is, in fact, reached is non-empty. But note that the optimality condition in the definition of naively optimal strategies only determines the intention-component of strategies, thus, beliefs can be constructed freely. This ensures that we can choose an optimal naive strategy at each h.

Now, to guarentee that the generated plan is stationary, we need to choose the same rootcut strategy for each set of histories where the end-state is identical. However, this is always possible, since whenever the final state is identical for two histories, both the rootcut strategy set and the utility function defined at those histories are identical, and therefore so are the set of rootcut optimal strategies. □

6. Sophistication

6.1. The meaning of sophistication

Definition 19. A strategy $\hat{s}^h$ of an agent at h is *sophisticatedly optimal*, if it maximizes expeced utility based on beliefs at all present and future histories and it is coherent :

$$\hat{s}^h(h')_b \in \arg\max_{a \in A_{\omega(h')}} U^{h'}_b\left(a, \hat{s}^h\right), \text{ for all } h' \in H^{\succcurlyeq h}.$$

and

$$\hat{s}^h(h')_i = \hat{s}^h(h')_b, \text{ for all } h' \in H^{\succcurlyeq h}.$$

The strategy of a sophisticated agent is made up of intentions to choose the actions that he believes future agents will choose, and acts optimally on these beliefs in the present.

6.2. Properties of sophisticated plans

A sophisticated plan is not necessarily consistent.

7. Comparing naiveté and sophistication

In this section, we proceed to examining some classical problems of dynamic inconsistency. The basis of comparison of naive and sophisticated plans is their performance in these problems. Since in our framework each agent occupies the same normative position, for welfare evaluations, we use the perspective of the root agent. Albeit it can be argued that this choice is arbitrary, it is nevertheless still preferable to any externally imposed welfare evaluation like the "long-run self": while the preferences of the starting agent are still the preferences of *some* decision making body, the long-run self is entirely external to the entire problem. In the following analysis, whenever we refer to the utility generated by the decision maker, it refers to the utility gained by the root agent.

To show the rather strong links between various decision problems, we present them as instances of parametrized problems.

7.1. A single task - procrastination and impulsiveness

The following problem models a situation when the decision maker can perform a single task once. The task can be pleasant or unpleasant. The state space contains only two elements: in state ρ the task hasn't been chosen (yet), while in state σ it has already been performed (see Fig. 6).

The specification of a strategy is only interesting for histories h with $\omega(h) = \rho$. A strategy for an agent at $\bar{h}$ then specifies for all future histories $H^{\succcurlyeq \bar{h}}$ when to take action B, if it hasn't been taken yet. Therefore, a strategy can be represented by a function, $f^{\bar{h}} : H^{\succcurlyeq \bar{h}} \to \{A, B\}$, where $f^{\bar{h}}(h)$ specifies whether B should be chosen at h, for those histories where it hasn't been chosen yet. A plan is a collection of such functions for all $\bar{h}$.

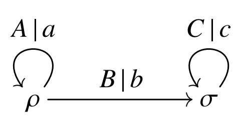


Figure 6: Single task problems.

The Bellman equation for state ρ is:

$$U_\rho^\beta = \max\left(a + \beta\delta U_\rho^1, b + \beta\delta\frac{c}{1-\delta}\right)$$

We get that whenever A – staying in ρ – is optimal if

$$(1 - \delta + \beta\delta)a > (1 - \delta)b + \beta\delta c$$

. Therefore, (assuming that, in case of a tie, B is chosen) with quasi-hyperbolic discounting, we get four possibilities:[9]

1. $a > (1 - \delta)b + \delta c$ and $(1 - \delta + \beta\delta)a > (1 - \delta)b + \beta\delta c$.
   For this case, A is optimal for both the present and the future, and all strategies of all agents intend and expect (and thus choose) A at all nodes. Dynamic inconsistency does not arise, and thus, the optimal plan for both naifs and sophisticates is $p = \{f^{\bar{h}} | \bar{h} \in H, f^{\bar{h}}(h) = A, \ \forall h \in H^{\succcurlyeq \bar{h}}\}$.
2. $a > (1 - \delta)b + \delta c$ and $(1 - \delta + \beta\delta)a \leq (1 - \delta)b + \beta\delta c$.
   In this scenario, an agent with no present-biasedness (i.e., $\beta = 1$) would prefer not to take action B, but a present-biased one will prefer taking it. An impulsive (or hot-headed), irreversible decision would be an example of this configuration.
3. $a \leq (1 - \delta)b + \delta c$ and $a > (1 - \beta\delta)b + \beta\delta c$.
   This is the dual of case 2: a present-biased agent would like to postpone, but eventually execute B, while one with no present-biasedness would opt for it right away. This is the case of procrastination of a single task.
4. $a \leq (1 - \delta)b + \delta c$ and $a \leq (1 - \beta\delta)b + \beta\delta c$.
   Here B is optimal, regardless whether the agent is present-biased or not. Therefore, B is immediately chosen, and the decision maker transitions to σ right away. There is no dynamic inconsistency, and optimal plan for both naifs and sophisticates is $p = \{f^{\bar{h}} | \bar{h} \in H, f^{\bar{h}}(h) = B, \ \forall h \in H^{\succcurlyeq \bar{h}}\}$.

Since case 1 and 4 do not give rise to dynamic inconsistency, we continue with the analysis of cases 2 and 3.

For the problem of impulsivity (case 2), the optimal naive plan is $\tilde{p} = \{f^{\bar{h}} | \bar{h} \in H, f^{\bar{h}}(\bar{h}) = B, f^{\bar{h}}(h) = A, \forall h \in H^{\succ \bar{h}}\}$. According to the induced strategy, the naive decision maker transitions to σ immediately. His induced utility is:

$$\tilde{U} = b + \frac{\beta\delta}{1 - \delta}c.$$

If it ever occurs to the naive decision maker to reconsider the situation, he will regret taking B.

The behavior of the sophisticated decision maker depends on his beliefs. There are two possibilities: either the root agent believes that all future agents will not take the B, or he believes that some future agent will take B.

[9]The following bindings for a, b and c, with $\beta = \delta = \frac{1}{2}$ show that each of this scenarios is, indeed, possible.

1. $a = 4, b = 0, c = 0$
2. $a = -4, b = 0, c = -10$
3. $a = 4, b = 0, c = 10$
4. $a = 4, b = 0, c = 20$

In the first case, since

$$b + \beta\delta\frac{c}{1-\delta} = \frac{1}{1-\delta}((1-\delta)b + \beta\delta c) \geq \frac{1}{1-\delta}(1-\delta+\beta\delta)a = (1+\frac{\beta\delta}{1-\delta})a,$$

we get that B is optimal. Therefore, in this case, the sophisticated decision maker will choose B immediately.

For the second case, take an agent at h with $\omega(h) = \rho$, and suppose he believes that t is the earliest future period in which B will be taken, according to the continuation strategy $\bar{s}$. For the utility of taking A and B, given this continuation strategy, we find that:

$$\begin{aligned} U^h(B, \bar{s}) &= b + \frac{\beta\delta}{1-\delta}c \geq \frac{1-\delta+\beta\delta}{1-\delta}a = a + \frac{\beta\delta}{1-\delta}(a - \delta^{t-1}a + \delta^{t-1}a) \\ &> a + \frac{\beta\delta}{1-\delta}((1-\delta^{t-1})a + (1-\delta)\delta^{t-1}b + \delta^t c) \\ &= U^h(A, \bar{s}) \end{aligned}$$

Therefore, for any belief on future actions, a sophisticated agent's best response is choosing B.

This shows that there is no difference in the behavior of sophisticated and naive agents for problems of impulsiveness: the only optimal sophisticated plan is identical to the only optimal naive plan. The name "impulsiveness" is more suitable than one might perhaps expect, as even sophisticates are not spared of its powers.

In the problem of procrastination (case 3), the optimal naive strategy at σ would prescribe to postpone B by one period, and then execute it. More precisely, the optimal naive plan is $p = \{f^{\bar{h}} | \bar{h} \in H, f^{\bar{h}}(\bar{h}) = A, f^{\bar{h}}(h) = B, \forall h \in H^{>\bar{h}}\}$. This implies that in the induced strategy, task B is in fact never executed, as the naive agent keeps postponing B, end ends up in ρ after an arbitrary length of time. The utility of a naive decision maker is:

$$\tilde{U} = a + \frac{\beta\delta}{1-\delta}a.$$

7.2. A repeated task: underinvestment and binges

The following decision problem is very similar to that in the previous one in that there are two states, with only one of them requiring a decision, i.e. the decision maker again has one substantial decision in front of him, whether to perform a task or not. The main difference is that the task can be executed repeatedly, as the decision maker returns to the initial node each time after doing the task. Fig. 7 presents the parametrized version of the repeated task problem.

As in the previous section, we will consider only a reduced set of strategies, those that assign the same action to all histories of equal length. We will also only consider coherent strategies, as it will allow us to specify only one component of the intention-belief pair. Moreover, since at σ there is no real choice for the decision maker, we will avoid specifying C. For example, $s^{\bar{h}} = s(t) = ABBABBA\ldots$ should be read as: "if the decision maker is in state ρ at history $\bar{h}$ with $t(\bar{h}) = t$, then he intends (and believes) choosing A; then, if he is in state ρ at a history h with $t(h) = t(\bar{h}) + 1$, then he intends choosing B" etc.

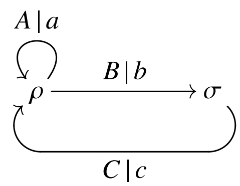


Figure 7: Repeated task problems

We start the analysis of this problem with the Bellman equation for state ρ:

$$U_\rho^\beta = \max\left(a + \beta\delta U_\rho^1, b + \beta\delta c + \beta\delta^2 U_\rho^1\right).$$

Again, this leads to four possibilities:

1. $a - b > \beta\delta(c - a)$.
   In this case, choosing A is optimal for all periods. Thus, for naifs and sophisticates, regardless whether they are present-biased or not, the only optimal strategy picks A at all nodes for both the intention and belief component. The issue of inconsistency does not arise.
2. $\beta\delta(c - a) \geq a - b > \delta(c - a)$.
   With this setup, a non-present-biased (naive) decsision maker will pick A, but a present-biased one will choose B. We can interpret the $B - C$ pair as a repeatedly arising "binge" choice, such as alcohol consumption or overspending, that brings benefits in the short run (B), but then leads to a backlash (C). The question is whether the decision maker can avoid taking B, and how many times.
   The naively optimal strategy for an individual with $\beta < 1$ is $\tilde{s}^h = \tilde{s}(t(h)) = BAAAA\ldots$ for all agents h. Thus, the naive plan is $\tilde{p}(h) = BAAAA\ldots$ for all h, which is not consistent. The induced strategy $\Lambda(p) = BBBBB\ldots$, and thus, the naive decision maker will always choose to binge, which leads to an induced payoff of:
   $$U_r^{h_0}(\tilde{p}) = b + \beta\delta\frac{c + \delta b}{1 - \delta^2}.$$
3. $(1 + \delta)a \leq b + \delta c$ and $a > (1 - \beta\delta + \beta\delta^2)b + \beta\delta(1 - \delta + \delta^2)c$.
   A naive optimal plan prescribes choosing A in this period, and B starting in the second period, whenever the state is σ. However, the naive agent ends up revising his strategy every turn, and executing A forever. Most classic examples of dynamic inconsistency are illustrations of this situation, e.g. where B would mean not repeating an addictive but harmful action; or a beneficial but costly action that brings benefits on the long run like physical exercise.
4. $(1 + \delta)a \leq b + \delta c$ and $a \leq (1 - \beta\delta + \beta\delta^2)b + \beta\delta(1 - \delta + \delta^2)c$.
   Here B is optimal whenever the state is $\rho 0$, regardless whether the agent is present-biased or not. Therefore, B is chosen in every second period, and the optimal naive plan is again consistent.

The following bindings show that each of this scenarios is, indeed, possible.

1. $a = 4, b = 0, c = 0$
2. $a = -4, b = 0, c = -16$
3. $a = 4, b = 0, c = 18$
4. $a = 0, b = 0, c = 4$

7.3. Indulgence

In the previous two decision problems, the utility induced by a sophisticated plan was never strictly lower than that induced by a naive plan. Is this always the case? In this section, we review a decision problem introduced by O'Donoghue and Rabin (1999). Since the state and action space are much larger then for the previous problems, instead of analyzing the fully parametrized version of the decision problem, we will focus on the specification on Fig. 9.

The problem is that of performing a single task in either period $0, 1, 2$ or 3 by a present-biased decision maker with $\beta = 0.5$ and $\delta = 0.5$. Think of consuming a bottle of valuable wine that gains in taste for up to three years, but then becomes undrinkable. Will one indulge in drinking it right away, is waiting for full maturity an option? On Fig. 9, "D" and "C" stand for "delay" and "consume". We will ignore specifying R in our strategies, since it doesn't affect the final utilities.

The root agent is $h_0 = (\rho_0)$. As we will deal with naifs and sophisticates, we can reduce our investigation to coherent strategies. Then, a strategy is for the root agent is a four-touple, that specifies whether to consume if the wine hasn't been consumed yet. Since only the first choice of C matters for induced utility, we get four relevant classes of strategies for the root agent:

- $S_0 = ((C_0, C_0), \dots);$
- $S_1 = ((D_0, D_0), (C_1, C_1), \dots);$
- $S_2 = ((D_0, D_0), (D_1, D_1), (C_2, C_2), \dots);$
- $S_3 = ((D_0, D_0), (D_1, D_1), (D_2, D_2), (C_3, C_3)).$

Taking any $s_i \in S_i$, and calculating payoffs from the perspective of h_0, we find that:

- $U^{h_0}(s_0) = 4;$
- $U^{h_0}(s_1) = 0 + \left(\frac{1}{2}\right)^2 \cdot 12 = 3;$
- $U^{h_0}(s_2) = 0 + 0 + \left(\frac{1}{2}\right)^3 \cdot 40 = 5;$
- $U^{h_0}(s_3) = 0 + 0 + 0 + \left(\frac{1}{2}\right)^4 \cdot 144 = 9.$

Executing similar calculations for the agent at $h_1 = (\rho_0, D_0, \rho_1)$:

- $U^{h_1}(s_1) = 12;$

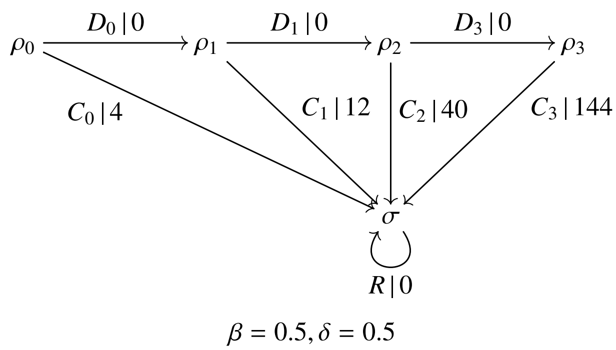


Figure 8: A problem of indulgence

- $U^{h_1}(s_2) = 0 + \left(\frac{1}{2}\right)^2 \cdot 40 = 10;$
- $U^{h_1}(s_3) = 0 + 0 + \left(\frac{1}{2}\right)^3 \cdot 144 = 18.$

Finally, for $h_2 = (\rho_0, D_0, \rho_1, D_1, \rho_2)$

- $U^{h_2}(s_2) = 40 = 10;$
- $U^{h_2}(s_3) = 0 + \left(\frac{1}{2}\right)^2 \cdot 144 = 36.$

The naively optimal strategy at h_0 and h_1 is thus choosing s_3, and at h_2 is picking some s_2. Therefore, for the naive plan $\tilde(p)$, we get $p(h_0) = p(h_1) = s_3$, and $p(h_2) = s_2 \in S_2$. The naive decision maker believes and intends postponing consumption right until the end, but in the last decision period, he indulges himselves. He thus waits two periods, and this generates a total utility of $U^(h_0)(s_2) = 5$ for the root agent.

To calculate the sophisticated plan, we need to resort to backward induction. The sophisticated agent at h_2 chooses s_2, for the same reasons as the naive agent, since there are no more decisions to make afterwards. Therefore, the sophisticated agent at h_1 can only consider two possibilities: s_1 or s_2. Since $U^{h_1}(s1) = 12 > 10 = U^{h_1}(s_2)$, he chooses to hasten the indulgence, since he believes (correctly) that he is unable to hold on the end anyway. Following the same reasoning, the root agent can choose between s_0 and s_1, and opts for s_0, consuming immediately. For the sophisticated plan $\hat(p)$, we get $\hat(p)(h_0) = s_0$, $\hat(p)(h_1) = s_1$ and $\hat(p)(h_2) = s_2$, which generates a total utility of $U^{h_0}(s_0) = 4$ for the root agent.

This shows that there are decision problems in which a fully sophisticated decision maker is worse off than a fully naive one. We empasize that this is not a new result, but shows that our framework can easily be put to use to make such welfare comparisons between agents of various types.

8. Hybrid decision makers

We now extend our model to capture mixed types. Our type space just includes naifs and sophisticates.

Definition 20. A Markov decision problem with agent types is made up of:

- the set of time periods $\{1, 2, \ldots\}$;
- a finite set of states Ω;
- a finite and non-empty set of pure actions A_ω that the decision maker can choose from in state ω;
- the type space $X = \{N, S\}$; we denote the state-type space by $\Theta = \Omega \times \{N, S\}$, and a state-type pair by θ; the initial state-type pair is $\theta_1 \in \Theta$; the type component is referred to as $x(\theta)$, while the state component as $\omega(\theta)$
- transition probabilities $m_\theta : A_{\omega(\theta)} \to \Delta(\Theta)$, with $m_\theta(\theta'|a_\omega)$ denoting the probability to transit from the state-type pair θ to the state-type pair θ' when action a_ω is chosen.
- a payoff function $u_\omega : A_\omega \to \mathbb{R}$ that assigns a payoff to every action in state ω;

Our definitions keeps the Markovian properties of the original model, and adds a specification of naiveté or sophistication. This calls for the inclusion of states into histories:

Definition 21. A type-dependent history h has the form $h = (\theta_1, a_{\omega_1}, \ldots, \theta_{t-1}, a_{\omega_{t-1}}, \theta_t)$, with:

- $\theta_i \in \Theta$, for $i \in \{1, \ldots, t\}$;
- ω_i is the state component of θ_i;
- $a_{\omega_i} \in A_{\omega_i}$, for $i \in \{1, \ldots, t-1\}$;
- $m_{\theta_{i-1}}(\theta_i|a_{\omega_{i-1}}) > 0$, for $i \in \{1, \ldots, t-1\}$.

Extending the previous notation (c.f. $t(h)$ and $\omega(h)$), $x(h)$ will refer to the current type.

Definition 22. A type-dependent optimal strategy $\check{s}^h$ at history h, for a Markov decision problem with agent types has the following properties:

- for $x(h) = N$:
$$\check{s}^h \in \arg\max_{s \in S^h} U_i^h(s),$$
and
$$\check{s}^h(h')_b = \check{s}^h(h')_i, \text{ for all } h' \in H^{\succcurlyeq h}.$$
- for $x(h) = S$:
$$\check{s}^h(h')_b \in \arg\max_{s \in S^{h'}} U_i^{h'}(s), \text{ for all } h' \in H^{\succcurlyeq h} \text{ with } x(h') = N;$$

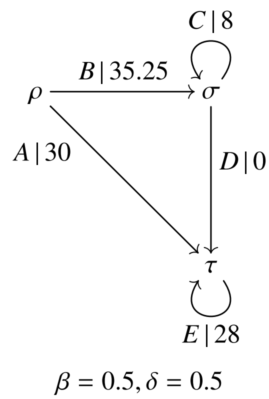


Figure 9: A drinking problem with mixed types

$$\breve{s}^h(h')_b \in \arg\max_{a \in A_{\omega(h')}} U_b^{h'}\left(\breve{s}^{h'}[a : h']\right)^{10}, \text{ for all } h' \in H^{\succsim h} \text{ with } x(h') = S\,;$$

and

$$\breve{s}^h(h')_i \in \arg\max_{a \in A_{\omega(h')}} U_b^{h'}\left(\breve{s}^{h'}[a : h']\right), \text{ for all } h' \in H^{\succsim h}.$$

9. Concluding remarks

Although our paper only allows for pure actions, the framework can be straightforwardly extended to include mixed actions. Similarly, our use of quasi-hyperbolic discounting is merely for purposes of presentation, and we could very well use more general discount functions, as long as the maximum expected utility converges to zero as we go further into the future.

A future area for research would be expanding our approach to deal with multiplayer settings, where players would be able to reason about and exploit the dynamic inconsistencies of the other players in a non-cooperative game-theoretic framework.

References

Akerlof, G.: "Procrastination and obedience." The American Economic Review 81.2 (1991): 1–19.

Ali, S. N.: "Learning self-control." The Quarterly Journal of Economics 126.2 (2011): 857–893.

Asheim, G. B.: "Procrastination, partial naivete, and behavioral welfare analysis." No. 2007, 02. Memorandum, Department of Economics, University of Oslo, 2007.

[10]The notation $\breve{s}^{h'}[a : h']$ should be interpreted as replacing in strategy $\breve{s}^{h'}$ the action taken at h' with some a.

Elster, Jon. Ulysses and the sirens: Studies in rationality and irrationality. Cambridge: Cambridge University Press, 1979.

Fischer, C.: "Read this paper even later: Procrastination with time-inconsistent preferences." Journal of Economic Behavior & Organization 46.3 (2001): 249–269.

Fudenberg, D., and D. K. Levine. "A dual-self model of impulse control." The American Economic Review (2006): 1449–1476.

Harris, C., and D. Laibson. "Dynamic choices of hyperbolic consumers." Econometrica 69.4 (2004): 935–957.

Heidhues, P., and B. Kőszegi. "Futile attempts at self-control." Journal of the European Economic Association 7.2-3 (2009): 423–434.

Jehiel, P., and A. Lilico. "Smoking today and stopping tomorrow: a limited foresight perspective." CESifo Economic Studies 56.2 (2010): 141–164.

Kahneman, D., P. P. Wakker, and R. Sarin. "Back to Bentham? Explorations of experienced utility." The Quarterly Journal of Economics 112.2 (1997): 375-406.

Kavka, G. S.: "The toxin puzzle." Analysis 43.1 (1983): 33–36.

Laibson, D. "Golden eggs and hyperbolic discounting." The Quarterly Journal of Economics 112.2 (1997): 443–478.

Lapied, A., and O. Renault: "A subjective discounted utility model." Economics Bulletin, 32.2 (2012): 1171–1179.

McClennen, E. F.: Rationality and dynamic choice. Cambridge: Cambridge University Press.

Méder, Z. Z., J. Flesch, and R. Peeters: "Optimal choice for finite and infinite horizons." Operations Research Letters, 40.6 (2012): 469–474.

O'Donoghue, T., and M. Rabin: "Doing it now or later." American Economic Review (1999): 103–124.

O'Donoghue, T., and M. Rabin: "Choice and Procrastination." The Quarterly Journal of Economics, 16.1 (2001): 121–160.

Strotz, R. H.: "Myopia and inconsistency in dynamic utility maximization." The Review of Economic Studies, 23.3 (1956): 165–180.