Naiveté and sophistication in dynamic inconsistency

v0.152, June 14, 2013
Preliminary version, please do not cite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 and examines concepts that are often entwined in the literature: it distinguishes 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 constructs 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 with previous approaches.

Keywords:
JEL: C70, D11, D91

☆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)

Contents

1 Introduction **3**

2 Related literature **3**

3 Basic concepts **4**
3.1 Markov decision problem . . . 4
3.2 History . . . 4
3.3 Conceptual remarks . . . 5
3.4 Intentions, beliefs, strategy . . . 7
3.5 Truncation . . . 9
3.6 Plans . . . 10
3.7 Induced strategy . . . 13

4 Utility and discounting **14**
4.1 Payoffs, horizon and utility . . . 14
4.2 "Expected utility" . . . 14

5 Naiveté **15**
5.1 The meaning of naiveté . . . 15
5.2 Properties of naive plans . . . 16

6 Sophistication **17**
6.1 The meaning of sophistication . . . 17
6.2 Properties of sophisticated plans . . . 17

7 Hybrid decision makers **18**
7.1 Hybrid optimal strategy . . . 18
7.2 A drinking problem . . . 20

8 Concluding remarks and future research **21**

Appendix **22**
A single task - procrastination and impulsiveness . . . 22
A repeated task - underinvestment and binges . . . 24
Indulgence . . . 27

References **29**

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 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).

[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.

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

In this section, we introduce our framework and notations. We provide standard definitions for the decision problem. Before introducing strategies, however, we make argue in section 3.3 for a new way of defining strategies that includes both beliefs and intentions. We then proceed by defining strategies and plans, which correspond to two different levels of analysis. Towards the end of the section, we present some results on the relationship between the concepts of consistency and stationarity.

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 $T = \{0, 1, 2, \dots\}$;
- 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;
- a payoff function $u_\omega : A_\omega \to \mathbb{R}$ that assigns a payoff to every action 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.

For simplicity, we do not allow for mixed actions in the first sections of the paper. In section **??** 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 $h = (\omega_0, a_{\omega_0}, \dots, \omega_{t-1}, a_{\omega_{t-1}}, \omega_t)$, with:

[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.

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

The length of h or current time at h is denoted by $t = t(h)$, and the function $\omega(h) = \omega_{t(h)}$ indicates the current or end state at history h. We use H to refer to the set of all histories.

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] The subset of H that consists of all histories that succeed h is denoted by $H^{\succcurlyeq h}$:

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

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 a short-run 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 in a certain state. We assume a one-to-one correspondance 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; they have preferences, and can relate to the future. As argued later, agents form 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 decision problem is fundamentally the same. Second, they suggest that the preferences of various temporal selves are similar, and a common functional form can be used to represent them (see section 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.

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 usually depends on the actions taken by other agents. In the most

[3]Obviously, $h' \preccurlyeq h \,\&\, h \preccurlyeq h' \Leftrightarrow h = h'$.

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. We keep the general assumption that past actions have no effect on the well-being generated by current and future actions. 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 one is thinking about hiring a private investigator to find our whether his wife has been cheating on him, it is the utility generated by her and his own 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.[4]

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 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 some 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 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, and they are strategically related, they ought have some connection with the future. At the least, 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 to *experienced* utility, then we just need to have very optimistic beliefs, and forming intentions are just means to specify these optimistic forecasts. As we will see

[4]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).

later, we can interpret naiveté 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 give a full description 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 history that succeeds the present:

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

Definition 4. The *beliefs* of an agent at history $\bar{h}$ assign an action to each history that succeeds the present:

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

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:

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

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 $s^{\bar{h}}$ about the future agent at 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 the agent at h will choose action a." Note that intentions and beliefs are defined at all succeeding histories, even at those that the agent does not intend to reach or believes 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,[5] 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 (or 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 $h, h' \in H^{\succcurlyeq \bar{h}}$ with $\omega(h) = \omega(h')$, we have $s^{\bar{h}}(h)_i = s^{\bar{h}}(h')_i$ (or respectively, $s^{\bar{h}}(h)_b = s^{\bar{h}}(h')_b$). 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.

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 $h'' \in H^{\succcurlyeq h} \cap H^{\succcurlyeq h'}$.[6]

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 property, whereas consistency is an extrinsic (relational) property of a strategy.

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.

[5] 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, he is aware of this impossibility, so he would not believe he will form i (Kavka, 1983).

[6] So, if two strategies are defined at histories that neither succeed nor preceed each other, then they are consistent, as there are no histories that succeed both.

To see this, take the decision problem in Figure 1. We will construct three stationary strategies s^{h_ρ}, s^{h_σ}, and s^{h_γ} such that the first and last two are consistent, but the first is not consistent with the last. Fix $h_\rho = (\rho)$, $h_\sigma = (\rho, A, \sigma)$ and $h_\gamma = (\rho, B, \gamma)$. Also, let $s^{h_\rho}(h_\rho) = (A, A)$, $s^{h_\rho}(h) = (C, C)$ if $\omega(h) = \sigma$, and $s^{h_\rho}(h) = (E, E)$ if $\omega(h) = \gamma$ for all h. Intuitively, s^{h_ρ} means: "go left and choose C, but if you ever end up at γ, choose E". Define two other strategies through $s^{h_\sigma}(h) = (C, C)$ for all $h \succcurlyeq h_\sigma$ ("do C in σ"), and $s^{h_\gamma}(h) = (F, F)$ for all $h \succcurlyeq h_\gamma$ ("do F in γ"). All of these strategies are stationary. Clearly, s^{h_ρ} and s^{h_σ} are consistent, since they both require the decision maker to choose C in state σ, and after history h_σ no state state other than σ is reachable. Next, s^{h_σ} and s^{h_γ} are consistent, since histories h_σ and h_γ neither succeed, nor precede each other. But s^{h_ρ} and s^{h_γ} 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.

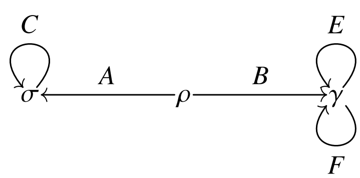


Figure 1: Stationary strategies with intransitive consistency

3.5. Truncation

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

Definition 9. Take any history $h = ((\omega_j, a_{\omega_j})_{j=1,\dots,t(h)-1}, \omega_{t(h)})$. The *truncation* operator $\dashv$ removes the first $t(h)-1$ pairs of this sequence, so $\dashv h = (\omega_{t(h)}) = (\omega(h))$. For a set of histories $H' \subseteq H$, we will refer to the set of truncated histories by $\dashv H'$.

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

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

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". Truncating 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 truncated versions are identical.

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.

Figure 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: 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 actions, i.e. $p(h)(h)_i = p(h)(h)_b$ for all h.

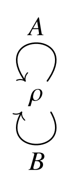


Figure 2: A basic decision problem.

We now proceed to introduce three properties of plans. Our definition of stationarity makes use of the truncation 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 $\dashv p(h) = \dashv p(h')$.

Stationarity of a plan is different from the stationarity of the strategies involved. For the decision problem on Figure 2, Table 2 offers a non-stationary plan of stationary strategies. To check for this, what needs to be verified

		$h \in H^{\succcurlyeq \bar{h}}$					
		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 Figure 2.

is that each row represents a stationary strategy. As there is only one state for this decision problem, this means that in each row, we should see the same intention-belief pair, which is indeed the case. Thus, Table 2 shows a plan of stationary strategies. The plan itself however is not stationary: by truncating the strategies in the first and second row, we get a different strategy.

On the other hand, Table 3 shows a stationary plan of non-stationary strategies. It is easy to see that the it is a plan of non-stationary strategies, as each row represent one strategy. The plan itself is stationary, which can be checked by comparing the rows. As there is only one state, we need to compare the truncations of all strategies. It should be obvious from Table **??** that by "forgetting the past", we get the same strategy in each row, namely: the agents picks action B right away, and intends to choose action A, and believes he will do so always in the future.

		$h \in H^{\succcurlyeq \bar{h}}$					
		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.

For a concrete example of a stationary plan of non-stationary 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.

Next, we define a consistent 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.

		$h \in H^{\succcurlyeq \bar{h}}$					
		ρ	$\rho A\rho$	$\rho B\rho$	$\rho A\rho A\rho$	$\rho A\rho B\rho$	$\ldots$
$\bar{h}$	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
	$\ldots$	—	—	—	—	—	$\ldots$

Table 3: Stationary plan of non-stationary strategies.

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 $h'' \in H^{\succcurlyeq h} \cap H^{\succcurlyeq h'}$, and either $h' \in H^{\succcurlyeq h}$, or $h \in H^{\succcurlyeq h'}$. Note that a consistent plan is necessarily made up by coherent strategies. This implies that a choice of an action 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:

$$p(h)(h') = p(h')(h') =\dashv p(h')(\dashv h') =\dashv p(h'')(\dashv h'') = p(h'')(h'') = p(h)(h'').$$

For the respective equations, we use, in order, consistency, definition of truncation, stationarity of the plan, definition of truncation 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 $\dashv p(h) =\dashv p(h')$. Since the strategies are stationary, their truncated versions are stationary too: $\dashv p(h) = s^h$ and $\dashv p(h') = s^{h'}$, where s^h and $s^{h'}$ are stationary. Since $p(h)$ and $p(h')$ are consistent, $\dashv p(h)$ and $\dashv p(h')$ are also consistent, so $s^h = s^{h'}$. So $\dashv p(h) =\dashv p(h')$. □

Based on the two theorems above, one might expect that a stationary plan of stationary strategies would be consistent. However, this is not necessarily so, as can be seen from the following example. Consider the Markov decision problem on Figure 3. For all $h \succcurlyeq \bar{h}$, let:

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

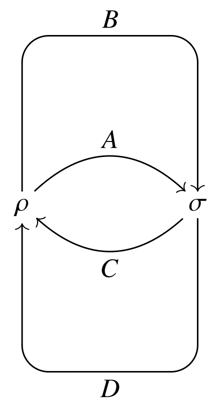


Figure 3: Stationary plan of stationary strategies is not necessarily consistent.

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 Definition 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.

$$U^h(u) = u_{t(h)} + \beta \sum_{t=t(h)+1}^{\infty} \delta^{t-t(h)}\, u_t,$$

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 section3.3, 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.

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

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

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:

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

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:

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

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 section1 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 Figure 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

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

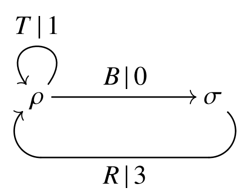
A plan $\tilde{p}$ is naive if the strategy $\tilde{p}(h)$ is naively optimal at each history h, .

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 naively 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



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

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	$\boxed{TT}$	BB	BB	BB	BB	...
	$\rho T\sigma$	—	$\boxed{TT}$	—	BB	BB	...
	$\rho B\sigma$	—	—	$\boxed{TT}$	—	—	...
	$\rho T\rho T\rho$	—	—	—	$\boxed{BB}$	—	...
	$\rho T\rho B\sigma$	—	—	—	—	$\boxed{TT}$	...
	...	—	—	—	—	—	...

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

Strotz (1956) shows that only when the discount function is exponential (i.e. $\beta = 1$) does the decision maker possess a consistent naive plan for all decision problems. For $\beta \in (0, 1)$, there are Markov decision problems for which there is no consistent naive plan. Figure 5.2 shows a decision problem which generates dynamic inconsistency for $\beta = 0.5$.

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 truncated 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 truncated strategy set and the utility function defined at those histories are identical, and therefore so are the set of truncated 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 :

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

and

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

where the notation $\hat{s}^{h'}[a : h']$ denotes the strategy where the action taken at h' is replaced with action a in strategy $\hat{s}^{h'}$.

A plan $\hat{p}$ is sophisticated if the strategy $\hat{p}(h)$ is sophisticatedly optimal at each history 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.

Theorem 4. *For any Markov decision problem, there exists a consistent sophisticated plan.*

6.2. Properties of sophisticated plans

. . . ...

We address the question of welfare comparision between sophistication and naivetĂŠ in the Appendix. The main finding is that for a set of decision problems, especially those with a one-off task, there is no difference between the induced strategies, and thus, sophisticates and naifs earn the same payoff. For recurring decisions (e.g. problems of investment), however, sophistication can outperform naivetĂŠ. Another example shows that in some situations, sophistication can lead to a *decrease* in induced utility compared to naivetĂŠ. Overall, although sophistication can save the decision maker from becoming inconsistent, its welfare implications are not clear-cut.

7. Hybrid decision makers

7.1. Hybrid optimal strategy

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

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

- the set of time periods $\{0, 1, 2, \dots\}$;
- a finite set of states 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 $\theta_0 \in \Theta$; the type component is denoted by as $x(\theta)$, while the state is denoted by$\omega(\theta)$;
- a finite and non-empty set of pure actions A_ω that the decision maker can choose from in state ω;
- a payoff function $u_\omega : A_\omega \to \mathbb{R}$ that assigns a payoff to every action in state ω;
- transition probabilities $m_\theta : A_{\omega(\theta)} \to \Delta(\Theta)$, with $m_\theta(\theta' | a_{\omega(\theta)})$ denoting the probability to transit from the state-type pair θ to the state-type pair θ' when action $a_{\omega(\theta)}$ is chosen.

This definitions keeps the Markovian properties of the original model, and adds a specification of naiveté or sophistication to each state. This requires a re-definition of history:

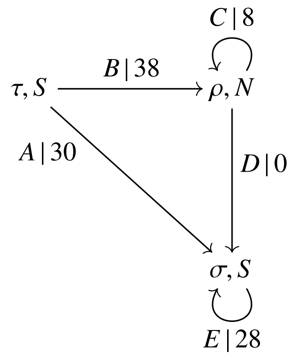
Definition 21. A *type-dependent history h* has the form $h = (\theta_0, a_{\omega(\theta_0)}, \dots, \theta_{t-1}, a_{\omega(\theta_{t-1})}, \theta_t)$, with:

- $\theta_i \in \Theta$, for $i \in \{0, 1, \dots, t\}$;
- $a_{\omega(\theta_i)} \in A_{\omega_i}$, for $i \in \{0, 1, \dots, t-1\}$;
- $m_{\theta_i}(\theta_{i+1} | a_{\omega(\theta_i)}) > 0$, for $i \in \{0, 1, \dots, t-1\}$.

Extending the previous notation (c.f. $t(h)$ and $\omega(h)$), $x(h)$ will refer to the current type. We keep the association between histories and agents - each history now corresponds to an agent, and it also includes the agent's type.

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

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



$\beta = 0.5, \delta = 0.5$

Figure 6: A drinking problem.

- for $x(h) = S$:

$$\breve{s}^h(h')_b \in (\arg\max_{s \in S^{h'}} U_i^{h'}(s))(h'), \text{ for all } h' \in H^{\succsim h} \text{ with } x(h') = N;$$

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

and

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

A plan $\breve{p}$ is type-dependently optimal, if $\breve{p}(h)$ is a type-dependent optimal strategy for all h.

This definition, albeit lengthy, captures our basic intuitions for the two types. If the current agent at h is naive, then, he is effectively unable to reason about future agents, and his intentions determine his beliefs, and thus, his whole strategy in the standard way. However, if the current agent at h is sophisticated, he is able to reason about future agents in the following way: if a future agent at h' is naive, then the agent at h believes he will act in such a way, maximizing his expected utility based on intentions. If, on the other hand, a future agent at h' is sophisticated, then the agent at h believes he will act in a sophisticated way, being able to reason about future agents just as well as h himself does. Finally, a sophisticated agent at h intends to choose in a sophisticated manner at all nodes, giving a best response to the choices of future agents - doing otherwise would require him to entertain different beliefs.[7]

7.2. *A drinking problem*

We now return to the problem raised in the introduction. The decision maker is sitting at the pub, having finished his second pint (in state τ), and is of type S (sophisticated). He can either go home directly, by choosing action A, transiting to state σ, where he doesn't need to make any more decisions. Alternatively, he can drink "one more pint" by choosing action B. This, however, transitions him to a "drunken" state ρ, where he becomes type N (naive). In this drunken state, he can choose between drinking one more beer by choosing C (thus, maintaing his drunkedness and returning to ρ), or going home to state σ by choosing D.

First, let us introduce some notation. Let $h = (\tau)$ be the root agent, and let $h_{drink} = (\tau, B, \rho)$ be the agent who drank "one more pint". Our goal is to construct $\breve{s}^h$, the type-dependent optimal strategy for the root agent. Since $x(h) = S$, we will have to deal with the interesting case, that of a sophisticated root agent.

We start the analysis of this situation by focusing on the state ρ. As $x(h') = N$ if $\omega(h') = \rho$, we get $\breve{s}^h(h')_b \in \arg\max_{s \in S^{h'}} U_i^{h'}(s)(h')$ according to the definition; the root agent believes that an agent at history h' with a current state ρ would be naive. What is the naive choice in state ρ? It is easy to see [8] that the optimal naive strategy is $(C,C)(D,D)(D,D)\ldots$, i.e. drink one more pint, and then go home (and later, if you find yourself in the pub, go home immediately). Thus, $\breve{s}^h(h')_b = C$ whenever $\omega(h') = \rho$. The root agent thus believes he would continue drinking once he becomes naive.

What about the intentions of the root agent about the agent at ρ? He believes that at history h', the continuation actions will be C. At history h', he could only choose between C and D. Going for C yields $8+0.5\cdot0.5\cdot\frac{8}{1-0.5} = 12$, whereas picking D gives $0 + 0.5 \cdot 0.5 \cdot \frac{28}{1-0.5} = 14$.Therefore, $\breve{s}^h(h')_i = D$ whenever $\omega(h') = \rho$. Together with the result of the previous paragraph, we get that $\breve{s}^h(h') = (D,C)$ whenever $\omega(h') = \rho$. We see that the intentions and the beliefs of the root agent do not match: he would like to pick D at every h', but correctly anticipates that he will be unable to do so, and would actually choose C. The sober, sophisticated root agent realizes that if he drinks just one more beer, he will end up drinking a lot more than he actually wishes for.

So what should the root agent choose at h? He can pick A, going home directly, earning him $30 + 0.5 \cdot 0.5 \cdot \frac{28}{1-0.5} = 44$. Or, he can pick B, drink one more beer, and end up in the pub. This would earn him $38+0.5\cdot0.5\cdot\frac{8}{1-0.5} = 42$. Going home seems best. Thus, $\breve{s}^h(h) = (A,A)$. The type-dependent optimal strategy is thus:

$$
\begin{aligned}
\breve{s}^h(h) &= (A,A), \\
\breve{s}^h(h') &= (C,D), \quad \text{whenever } \omega(h') = \rho, \\
\breve{s}^h(h') &= (E,E), \quad \text{whenever } \omega(h') = \sigma.
\end{aligned}
$$

[7] It could be argued that the agent at h intends something else at h', namely, to give a best response to the choices of future agents *based on the incentives at h, not the ones at h'!* However, taking this aspect into consideration would complicate our formalism, changing little in the upcoming examples.

[8] Obviously, the decision problem reduced to the states ρ and σ is that of a single task, discussed in section 7.1. Since $8 < (1-0.5)+0.5\cdot28 = 14$, a non-present-biased agent would choose to go home. However, because $8 > (1-0.5\cdot0.5)\cdot0+0.5\cdot0.5\cdot28 = 7$, a present-biased agent prefers postponing C.

Note that a fully naive root agent would expect that he can resist the temptation of additional beers, and would expect a utility of $38 + 0.5 \cdot 0.5 \cdot 0 + 0.5 \cdot 0.5^2 \cdot \frac{28}{1-0.5} = 45$. The sophisticated root agent realizes that this is unachievable, as the incentives *and the type* of the decision makers changes by transiting to ρ.

In the drinking problem, the sober, sophisticated root agent avoids becoming naive, and thus is better off. Since we have seen in section 7.3 that sometimes a sophisticated decision maker is worse off than a naive one, a decision problem in which a sophisticated root agent makes a choice that turns him into a naive, instead of keeping his sophistication can be easily constructed. It thus seems that a sophisticated agent might sometimes prefer self-deception over introspection.

By what means such self-deception might be effectively achieved, or whether it can be achieved intentionally at all is, of course, a difficult problem. The main message of this model is that self-deception has its virtues, which might, in itself, challenge ethical arguments on the inherent immorality of self-deception.[9]

8. Concluding remarks and future research

This paper attempts to play a foundational role for future discourse in multi-self models of dynamic inconsistency. It attempts to establish that the basic epistemic concepts to be considered are beliefs and intentions, and the main levels of analysis should be those of strategies and plans. We would now like to outline some directions for follow-up research in this area.

An obvious limitation of the current framework is that it only allows for pure actions (except for the remark in section 7.2). This limitation is introduced to ease the presentation, and the technical adaptations required for dealing with mixed action can be accomplished rather straightforwardly. Mixed actions should play a particularly important role when moving from decision-theoretic models to a game setting.

The second self-imposed limitation was the use of quasi-hyperbolic discounting for the utility functions. Again, we can relax this assumption, and our existence results would carry over to more general utility functions, if the limit of the expected utility of each strategy exists, and the utility function is continuous at infinity (Fudenberg and Levine , 2006). Quasi-hyperbolic discounting was sufficient to illustrate our main points, however.

Whereas our focus was the two most common types of decision makers facing dynamic inconsistency, naifs and sophisticates, there have been arguments in the literature for taking seriously other types as well. In particular, McClennen (1990) argues for the possibility of resolute decision making. In fact, resoluteness can very easily be incorporated in our framework. The most interesting applications of this idea would be in our model of hybrid decision makers, where an agent sometimes might be able to exercise resoluteness, sometimes would become naive (due to cognitive limitations), and sometimes would have resort to sophistation.

The horizon of sophisticated decision makers requires further investigation. In case decision makers have only a finite horizon, reasoning about future agents can be based on of two assumption: either the *length*, or the *endpoint* of the horizon of that future agent is the same as that of the current agent. In the former case, we are

[9]For an overview on the ethical problem of self-deception, see Deweese-Boyd (2012)

talking about a *moving*, in the latter, about a *fixed* horizon. The implications of these two assumptions on the optimal strategies (generated, for instance, through a backward induction reasoning) are not yet understood.

Finally, the most interesting implications for the framework presented above will be for game theory. How can game players reason about the intentions and beliefs of other players, as well as their types? How can strategies exploit the naiveté (or sophistication) of others? What kind of equilibria are generated when (hybrid) players are pitted against each other? We hope that through this paper, we have broken the ground for such questions.

Appendix

A single task - procrastination and impulsiveness

The following problem models a situation where the decision maker can perform a single task once. 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 Figure 7).

$$\begin{tikzcd} \rho \arrow[loop above, "A\,|\,a"] \arrow[rr, "B\,|\,b"] & & \sigma \arrow[loop above, "C\,|\,c"] \end{tikzcd}$$

Figure 7: Single task problems.

We will work with the assumption that whenever the agent is indifferent between choosing A and B, he will execute the task. Checking first whether the agent has incentives to execute the task in the *next* period, we get that he would execute it if $a \leq (1-\delta)b + \delta c$, and would not do it otherwise. This yields two cases:

1. $a \leq (1-\delta)b + \delta c$.
   The agent would execute the task in the next period, but will he execute it immediately? If he does, he gains $b + \beta\delta\frac{c}{1-\delta}$; if he doesn't, he gains $a + \beta\delta b + \beta\delta^2\frac{c}{1-\delta}$. This again, generates two cases, where after some algebra we get:
   (a) $a \leq b(1-\beta\delta) + \beta\delta c$. [10]
   Here, executing the task is the optimal immediate choice, as well as the optimal choice for all future periods. Both naively and sophisticatedly optimal strategies prescribe choosing B at all histories where the end-state is ρ; there is no dynamic inconsistency.
   (b) $a > b(1-\beta\delta) + \beta\delta c$. [11]
   In this case, executing the task is optimal in the next period, but not immediately. This is the case of procrastination of a single task. A naively optimal strategy prescribes postponing B one period by choosing A. A naive plan would involve choosing such a strategy at any history. This implies that in the induced strategy, task B is in fact never executed, as each naive agent keeps postponing B by

[10] e.g. $a = 4, b = 0, c = 20, \beta = \delta = \frac{1}{2}$.
[11] e.g. $a = 4, b = 0, c = 10, \beta = \delta = \frac{1}{2}$.

one period. The utility of a naive agent is thus: $U_{naive} = a + \frac{\beta\delta}{1-\delta}a$. To find sophisticatedly optimal strategies, first assume that the sophisticated agent believes that no future agent will choose B. Then, by choosing A, his expected utility is:

$$\frac{1}{1-\delta}a > \frac{1-\beta\delta}{1-\delta}b + \frac{\beta\delta}{1-\delta}c > \frac{1-\delta}{1-\delta}b + \frac{\beta\delta}{1-\delta}c = b + \frac{\beta\delta}{1-\delta}c,$$

which is what he could gain by choosing B instead. In this case, it is optimal for the agent to choose A. Suppose on the other hand that the agent beliefs that for some number of periods t, he will choose A, followed by B (strategy $\bar{s}$). We will show that choosing A for one *less* period ($\bar{s}$') generates more utility:

$$U_b^h(\bar{s}') = a + \frac{\beta\delta}{1-\delta}(1-\delta^{t-2})a + \beta\delta^{t-1}b + \beta\delta^t\frac{1}{1-\delta}c \geq a + \frac{\beta\delta}{1-\delta}(1-\delta^{t-1})a + \beta\delta^t b + \beta\delta^{t+1}\frac{1}{1-\delta}c = U_b^h(\bar{s}) \Leftrightarrow$$

$$\beta\delta\delta^{t-2}(\delta-1)\frac{1}{1-\delta}a - \beta\delta\delta^{t-2}(\delta-1)b - \beta\delta\delta^{t-2}(\delta-1)\frac{\delta}{1-\delta}c \geq 0 \Leftrightarrow$$

$$\frac{1}{1-\delta}a - b - \frac{\delta}{1-\delta}c \geq 0,$$

which is true, as we are in a subbranch of case 1. This means that any sophisticatedly optimal strategy that prescribes choosing B in the future should prescribe choosing it in the next period. Therefore, we only have three remaining candidates for a sophisticatedly optimal strategy: never choosing B, always choosing B, and choosing B starting from the next period, but choosing A in this period. However, we know that choosing B when B is played in the next period is not optimal - the agent would like to postpone. On the other hand, choosing B starting from the next period, but choosing A in this period cannot be a sophisticatedly optimal strategy, as it prescribes a suboptimal response (choosing B) for the next agent (who believes all future agents will choose B, so should choose A himself). Therefore, the only sophisticatedly optimal strategy is never choosing B. The expected utility of a sophisticated agent is the same as that of a naive one $U_{soph} = a + \frac{\beta\delta}{1-\delta}a = U_{naive}$.

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

   This condition means that the agent has no incentives to execute the task in the next (and subsequent) periods. Thus, his expected utility is $b + \beta\delta\frac{c}{1-\delta}$ if he decides to perform it immediately, while just $a + \beta\delta\frac{a}{1-\delta}$ if he does not do so (thus not performing the task at all). The two sub-cases are now:

   (a) $(1-\delta+\beta\delta)a \leq (1-\delta)b + \beta\delta c$. [12]

   Here, an agent would not prefer executing the task in later periods, but has the incentives to do it right away. An irreversible, impulsive decision would be an example of this configuration. A naively optimal strategy simply prescribes choosing B right away, generating the naive agent a utility of

[12] e.g. $a = -4, b = 0, c = -10, \beta = \delta = \frac{1}{2}$.

$U_{naive} = b + \beta\delta\frac{c}{1-\delta}$.

The sophisticatedly optimal strategy depends on the beliefs of the agent.

Case 1: the root agent believes that all future agents will not take the B. 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 choosing B is optimal.

Case 2: the sophisticated agent believes that some future agent at h will take B. Then $t = t(h)$ is the earliest future period in which B will be taken, according to the strategy $\bar{s}$. We get that:

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

Taking B right away is thus optimal for the sophisticated agent even when he believes some future agent will execute B. Therefore, for any beliefs about future agents, a sophisticated agent's best response is choosing B. This means that $U_{soph} = U_{naive}$, and the sophisticated agent cannot avoid the problem of impulsivity. In either case, if the decision maker later reflects on this decision from a non-present-biased perspective, he will regret taking B.

(b) $(1-\delta+\beta\delta)a > (1-\delta)b + \beta\delta c$. [13]

Now, picking A and staying in ρ is always optimal both for naifs and sophisticates, and dynamic inconsistency does not arise.

This shows that for single task problems, there is no difference in the behavior of sophisticated and naive agents for the problems of procrastination and impulsiveness, which might be somewhat surprising. At this point, it is not clear whether sophistication brings any welfare benefits at all.

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, 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 one period after doing the task. Figure 8 presents the parametrized version of the repeated task problem.

To simplify the analysis, we will restrict our attention to strategies where the action assinged to a history h can be conditioned only on the current time and state, i.e. Markovian strategies. This will allow for a very compact representation of coherent strategies, namely, as a sequence of $A-s$ and $B-s$. For an agent at $\bar{h}$ The

[13] e.g. $a = 4, b = 0, c = 0, \beta = \delta = \frac{1}{2}$.

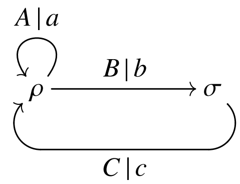


Figure 8: Repeated task problems.

k^{th} element in such a sequence specifies what intention-belief pair the agent assigns to period $k-1$. For instance, when we write $s^{\bar{h}} = ABAA\ldots$ it can be read as: "the agent at $\bar{h}$ (in state ρ) intends and believes A for his current time, intends and believes action B if he is in state ρ" etc. We also assume that whenever the agent is indifferent between the two actions available in state ρ, he will execute the task B. Starting with the *next* period, the agent has incentives to execute B whenever $a + \delta a \leq b + \delta c$. We get the following cases

1. $a + \delta a \leq b + \delta c$.
   The agent will prefer executing B from the second period onwards. Whether he prefers to do it right away depends on the relationship between $a + \beta\delta a$ and $b + \beta\delta c$. The two cases are:
   (a) $a + \beta\delta a \leq b + \beta\delta c$. [14]
      Here B is optimal whenever the state is ρ, regardless whether the agent is naive or sophisticated. Dynamic inconsistency does not arise.
   (b) $a + \beta\delta a > b + \beta\delta c$. [15]
      . . .
2. $a + \delta a > b + \delta c$.
   In this case, the agent will prefer picking A from the second period onwards. He will prefer to execute B immediately whenever $a + \beta\delta a \leq b + \beta\delta c$.
   (a) $a + \beta\delta a \leq b + \beta\delta c$. [16] Here, the agent has incentives to choosing B once immediately, but never choose it thereafter. A shopping binge or the consumption of an addictive substance might serve as examples for such a decision problem: present-biasedness induces the enjoyable action B, for which the costs c only need to payed later. A naive agent thus plans to choose B at the outset, but refrain from choosing it again later. Obviously, the plan generated by such strategies will be inconsistent, and the naive agent will stay addicted or keep on binging, executing B whenever in state ρ. The induced payoff for this naive plan will be $U_{naive} = b + \beta\delta c + \beta\delta^2 \frac{b+\delta c}{1-\delta^2}$.
      For finding sophisticated plans, first consider an agent at history h who believes the agent in the next period in state A would choose A. This implies that regardless of what the agent at h chooses, he will be in state ρ two periods from now. As his payoff from that period is fixed - as he will not condition

[14] e.g. $a = 0, b = 0, c = 4, \beta = \delta = \frac{1}{2}$.
[15] e.g. $a = 4, b = 0, c = 18, \beta = \delta = \frac{1}{2}$.
[16] e.g. $a = -4, b = 0, c = -16, \beta = \delta = \frac{1}{2}$.

his choices on the history, per our assumption - he only needs to contrast the payoffs gained in this and the next period. By choosing A, he would gain $a + \beta\delta a$; by choosing B, he would gain $b + \beta\delta c$. With our assumption on the parameters, it is optimal for the agent at h to choose B. This means that in a sophisticated strategy, an action A is always preceded by an action B.

Now, take an agent at a history $\bar{h}$ that chooses action A at that history. There are two possibilities: either his strategy does not assign action A to any future history where the current state is ρ, or it assigns A to at least one. Suppose the first possibility; then, the strategy of the agent assigns B to all future histories where the current state is ρ. He can then choose between adopting strategy $s_A^{\bar{h}} = ABBBBB\ldots$ and strategy $s_B^{\bar{h}} = BBBBBB\ldots$. His expected utility for these is:

$$\begin{aligned} U(s_B^{\bar{h}}) = \quad & b + \beta\delta c + \beta\delta^2 \tfrac{\beta+\delta c}{1-\delta^2} \geq a + \beta\delta a + \beta\delta^2 \tfrac{\beta+\delta c}{1-\delta^2} - \beta\delta \tfrac{\beta+\delta c}{1-\delta^2} = \\ = \quad & \beta\delta \tfrac{a+\delta a-(b+\delta c)}{1+\delta} + a + \beta\delta \tfrac{\beta+\delta c}{1-\delta^2} > a + \beta\delta \tfrac{\beta+\delta c}{1-\delta^2} = U(s_A^{\bar{h}}) \end{aligned}$$

Thus, if the agent beliefs that all future agents will choose B, he should choose B himself. Therefore, one sophisticatedly optimal strategy -based on the belief that all future agents will choose B - is playing and intending to play B in all periods.

The second possibility is that the agent at $\bar{h}$ believes that at least one future agent will choose A. Denote the earliest such period by $k + 1$. The strategy of the agent at $\bar{h}$ can then be either $s_A^{\bar{h}} = ABB\ldots BBA\ldots$ or $s_B^{\bar{h}} = BBB\ldots BBA\ldots$, where the first and last represented action are seperated by k periods of choosing B. It is clear that after period $k + 1$, the agent will definitely by in state ρ - therefore, the payoffs from that point on can be ignored for the purposes of choosing an action for the agent at $\bar{h}$. We will show that $U(s_B^{\bar{h}}) > U(s_A^{\bar{h}})$, so that choosing B will be optimal. For simplicity, we will assume k is even; the proof for k odd is analogous.

$$\begin{aligned} U(s_B^{\bar{h}}) \quad &> \quad U(s_A^{\bar{h}}) \Leftrightarrow \\ b + \beta\delta c + \beta\delta^2(b+\delta c)\frac{1-\delta^k}{1-\delta^2} \geq a + \beta\delta a + \beta\delta^2(b+\delta c)\frac{1-\delta^k}{1-\delta^2} \quad &> \quad a + \beta\delta(b+\delta c)\frac{1-\delta^k}{1-\delta^2} + \beta\delta^{k+1}a \Leftrightarrow \\ \beta\delta(a - (b+\delta c)\frac{1-\delta^k}{1-\delta^2}(1-\delta)) \quad &> \quad \beta\delta\delta^k a \Leftrightarrow \\ a - (b+\delta c)\frac{1-\delta^k}{1+\delta}) > a - (a+\delta a)\frac{1-\delta^k}{1+\delta}) = \delta^k a \quad &= \quad \delta^k a \end{aligned}$$

This shows that on the belief that some future agent will choose A, a sophisticated agent should choose B, which implies that no sophisticated strategy can include an action A. As we have already shown that $BBBBB\ldots$ is a sophisticatedly optimal strategy, it is clear that sophisticates cannot outperform naives in the binge/addiction problem.

(b) $a + \beta\delta a > b + \beta\delta c$. [17]

With this parameter configuration, picking A is always optimal for both naifs and sophisticates, and there is no issue of dynamic inconsistency.

1. $(1 + \delta)a \leq b + \delta c$ 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.

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 1action spaces 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 Figure 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 Figure 9, "D" and "C" stand for "delay" and "consume". We will ignore specifying the choice of an action for histories h with $\omega(h) = \sigma$.

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:

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

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

- $U^{h_0}(s_0) = 4$;

[17] e.g. $a = 4, b = 0, c = 0, \beta = \delta = \frac{1}{2}$.

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

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

- $U^{h_1}(s_1) = 12$;
- $U^{h_1}(s_2) = 0 + \left(\frac{1}{2}\right)^2 \cdot 40 = 10$;
- $U^{h_1}(s_3) = 0 + 0 + \left(\frac{1}{2}\right)^3 \cdot 144 = 18$.

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 himself. 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 $U^{h_1}(s_1) = 12 > 10 = U^{h_1}(s_2)$, 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.

$$
\begin{tikzcd}
\rho_0 \arrow[r, "D_0\,|\,0"] \arrow[drr, "C_0\,|\,4"'] & \rho_1 \arrow[r, "D_1\,|\,0"] \arrow[dr, "C_1\,|\,12"] & \rho_2 \arrow[r, "D_3\,|\,0"] \arrow[d, "C_2\,|\,40"] & \rho_3 \arrow[dl, "C_3\,|\,144"] \\
 & & \sigma \arrow[loop below, "R\,|\,0"] &
\end{tikzcd}
$$

$$\beta = 0.5, \delta = 0.5$$

Figure 9: A problem of indulgence.

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.

Bach, C. and Heilmann, K.: "Agent connectedness and backward induction." International Game Theory Review 13.2 (2011): 195–208.

Deweese-Boyd, Ian: "Self-Deception." The Stanford Encyclopedia of Philosophy (Spring 2012 Edition), Edward N. Zalta (ed.), Retrieved from `http://plato.stanford.edu/archives/spr2012/entries/self-deception`.

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.