Sleeping Newcomb

John W. E. Cremin*

Aix-Marseille University, CNRS, AMSE

First version: 31^{st} of March 2026
Latest Version: April 20, 2026

Abstract

I study games with self-locating uncertainty in which an agent at a single information set is uncertain of his position even *within* a given information set of a given play of the game. In such games, there is an analogy to be drawn with Newcomb's problem: in both settings, locally rational (thirder) reasoning and globally optimal (planning) reasoning can prescribe different strategies. I call this a *Newcomb tension*, and present a representation theorem: a Bayesian with commitment power and an uncommitted agent holding incorrect 'one-boxer' beliefs are behaviourally equivalent. In the single-agent case, randomisation always resolves the tension but in multi-agent games, in which planning and interim social weights diverge under some conditions, a multi-agent Newcomb tension can survive this randomisation resolution with an asymmetric awakening structure across agents. I consider the implications of this for the duplicating Sleeping Beauty problem, and a duplicating variant of the absent-minded driver.

Keywords— Absent-Minded Driver, Sleeping Beauty Problem, Newcomb's Problem, Self-Locating/Indexical Uncertainty, Imperfect Recall

JEL Codes— C72, D71, D81, D83

1 Introduction

Bored, one Sunday morning, of debating the virtues of a minimalist state, Nozick has decided to make a foray into Bayesian epistemology. He throws aside the draft of his latest letter to Rawls, heads to the blackboard, and draws a dot diagram of the duplicating Sleeping Beauty problem (Figure 1). Though aware that the controversy of the canonical variant (c.f. Appendix A.1) is over whether Beauty should believe the coin has landed heads with probability $\frac{1}{2}$ or $\frac{1}{3}$, his interest is instead peaked by a different question. On the conventional debate, the thirders seem to have probably got the better end of the stick, but despite this, and despite the equivalence between the information structures across canonical and duplicating variants, he is not sure that Beauty should *act* on the basis of those thirder beliefs here.

Being an avid reader, he is also familiar with Harsanyi (1955), and decides that when it comes to deciding the optimal action, this is a social choice problem. He assumes co-cardinal utilities,

---

*Email: `john-walter-edward.cremin@univ-amu.fr`. The project leading to this publication has received funding from the French government under the "France 2030" investment plan managed by the French National Research Agency (reference: ANR-17-EURE-0020) and from Excellence Initiative of Aix-Marseille University – A*MIDEX.

Agent & day	$\theta = H(=0)$	$\theta = T(=1)$
$n = 1, d = 1$	$\bullet$	$\bullet$
$n = 2, d = 2$		$\bullet$

Figure 1: Dot diagram for Duplicating Sleeping Beauty. Agent $n = 1$ (the original) wakes in both states; agent $n = 2$ (the clone) wakes only in $\theta = T$. Both agents share a single information set and cannot distinguish their awakenings. Upon waking up, the agent chooses a 'bet' $x \in [0, 1]$, and receives utility equal ot their Brier score $-(x - \theta)^2$. They are expected utility maximisers.

adopts a utilitarian social welfare function, and weights the original Beauty and the clone according to their probability-shares of the information set. However, here he notices an eery resemblance to his earlier Newcomb's problem (c.f. Appendix A.2). Since between Sunday, $d = 0$, and Monday, $d = 1$, Beauty's beliefs about the probabilities of the different dots change, so does her social welfare function: on Sunday, she assigns weight $\frac{1}{4}$ to the clone's interests, but upon waking up she revises this to $\frac{1}{3}$. She no longer agrees with her former self about the optimal course of action, even though at the planning stage she knew she would arrive at this point almost surely. I call this divergence a *Newcomb tension*, due to this analogy with Newcomb's problem. In Newcomb's problem, an agent who reasons causally (two-boxing) takes an action that is locally rational but globally suboptimal; the planning-optimal action (one-boxing) requires the agent to act as though their choice were correlated with states they cannot causally influence.

In both cases, the agent at the point of decision faces a tension between *interim rationality* (act on your current beliefs) and *planning rationality* (act as you would have committed to act before arriving at the information set). This Newcomb tension is not confined to the duplicating Sleeping Beauty and Newcomb problems, but can occur generally in the presence of *self-locating uncertainty*. The absent-minded driver (AMD) of Piccione and Rubinstein (1997) (c.f. Appendix A.3) exhibits a similar structure when randomisation is not permitted: the driver at the information set, reasoning with correct beliefs about the intersection at which they find themselves, does not consider the planning optimal action to be interim optimal. Beyond this, as was famously noted by Piccione and Rubinstein (1997) themselves in the special issue posing and discussing the problem, no interim optimal action can even be said to exist in that case.

The purpose of this paper is to study systematically under what conditions a Newcomb tension arises in games with self-locating uncertainty, which I define and study more generally in a companion paper: *Sleeping Beauty's Dismal Day Out*. When does the awakening structure of a game create a divergence between planning-optimal and interim-optimal actions? I restrict attention to games with a single information set, allowing multiple agents to occupy 'dots' (decision nodes) at said set, and distinguishing intra-personal self-locating uncertainty (as in the absent-minded driver) from the inter-personal variety (as in Duplicating Sleeping Beauty). After defining the Newcomb tension, I prove a representation theorem (Theorem 2.2): a committed thirder is behaviourally equivalent to an uncommitted agent holding "one-boxer" beliefs. Hence, in games exhibiting a Newcomb tension more generally, one can either incorporate some planning stage and commitment power, or equivalently one can simply impose that they act as-if they adopted said one-boxer beliefs.

The paper's main results concern when the Newcomb tension can and cannot be resolved. In the single-agent case, behavioural strategies (randomisation) always achieve this: the planning-optimal behavioural strategy is also interim-optimal (Theorem 3.1), however many states of the world exist. When a pure-strategy interim fixed point exists, it must already be planning-optimal

(Proposition 3.2), so randomisation adds nothing in such cases. In the multi-agent case, I approach the question of what action an agent should choose as a social choice problem, aggregating payoffs with a utilitarian social welfare function. The central result (Theorem 4.1) is that the Newcomb tension requires an asymmetric dot structure across agents, and generically exists in its presence. In the duplicating absent-minded driver (a variant of the AMD in which two separate individuals act at the different intersections), Proposition 4.3 provides a sharp characterisation of when and by how much the planning and interim optima diverge. I also consider the duplicating Sleeping Beauty problem, which is of interest in its own right, as my approach suggests that though it may not be optimal for her to act on thirder beliefs in this problem, they are nonetheless the correct credences to hold.

I should note that my interest in games with self-locating uncertainty is driven primarily by a desire to study models of social and individual learning involving *temporal ignorance*. My earlier article *Bot Got Your Tongue?* features this, though in that instance it is not the main focus of the paper. Though the Sleeping Beauty and absent-minded driver problems, and especially the duplicating variants, may seem arcane problems of no practical application, their purpose is to aide the study of self-locating uncertainty generally, so that it can be applied to this practical purpose.

The paper is organised as follows. Section 2 formalises single-information-set games with self-locating uncertainty, defines the planning and interim stages, introduces the Newcomb tension, and proves the one-boxer representation theorem. Section 3 presents the single-agent results: the randomisation resolution theorem, the pure-strategy fixed-point result, and multi-state examples that illustrate that the absence of a state of the world in the AMD is not necessary for my results. Section 4 extends the analysis to multiple agents, and considers Duplicating Sleeping Beauty problem, the duplicating absent-minded driver, and *Duplicating Sleeping Beauty Behind the Wheel*, a combination of the two. Section 5 reviews the related literature, and Section 6 concludes. I include appendices describing each of the Sleeping Beauty problem, Newcomb's problem, and the paradox of the absent-minded driver for readers who have yet to encounter them, and to which I have already referred in the course of this introduction. Appendix B collects omitted proofs.

2 Games with Self-Locating Uncertainty

The Newcomb tension described above arises in a class of extensive-form games in which an agent at an information set cannot determine which node they occupy, and in which this ignorance is greater than in standard games of imperfect information. I now formalise this class. The setting is deliberately simple: I consider games with a *single* information set at which a single action must be chosen, but I allow *multiple agents* to occupy nodes at that information set, and allow individual terminal histories ('plays') of the game to feature multiple nodes within it. This is the simplest environment in which self-locating uncertainty and the resulting Newcomb tension can be studied in isolation, free from complications introduced by sequential play across multiple information sets. I also assume throughout that agents have *interpersonally comparable* (co-cardinal) von Neumann-Morgenstern utility. This assumption is without loss for single-agent games, but is essential for the multi-agent welfare analysis in Section 4.

The formulation follows the *primitive extensive form* of the companion paper *Sleeping Beauty's Dismal Day Out*, itself a modification of Kreps and Wilson (1982) that drops the requirements of perfect recall, and exclusive occupancy (one agent per information set).

2.1 The Primitive Extensive Form

The physical order of play is represented by a finite set of nodes T together with a binary relation $\prec$ of *precedence* on T. Following Kreps and Wilson (1982), I require $(T, \prec)$ to form an *arborescence*: $\prec$ is a partial order, and the predecessors of each $t \in T$ are totally ordered by $\prec$. I partition T into terminal nodes Z (outcomes, having no successors), decision nodes $X = T \setminus Z$, and initial nodes Θ (states, having no predecessors). For $t \notin \Theta$, I write $p_1(t)$ for the immediate predecessor of t; the *depth* $l(t)$ is the unique non-negative integer such that $p_{l(t)}(t) \in \Theta$. For $x \in X$ I write $S(x)$ for the set of immediate successors. Choices are represented by a finite set of *actions* A and a function $\alpha : T \setminus \Theta \to A$, where $\alpha(t)$ denotes the action taken to reach t from $p_1(t)$, with α one-to-one on $S(x)$ for each $x \in X$.

Definition 2.1 (Single-Information-Set Game with Self-Locating Uncertainty)**.** A *single-information-set SLU game* is a collection $\Gamma = \{T, \prec; A, \alpha; N, \iota; h; u, \rho\}$ where $(T, \prec)$ and (A, α) are as above, N is a finite set of *agents*, $\iota : h \to N$ is an *agent assignment* specifying which agent occupies each node in the information set, $h \subseteq X$ is the unique information set, $u : Z \times N \to \mathbb{R}$ is a payoff function assigning a co-cardinal von Neumann-Morgenstern utility to each terminal node for each agent, and ρ is a strictly positive probability measure on Θ (the prior). The information set satisfies:

(A1) *(Action consistency)* $\quad x, x' \in h \;\Rightarrow\; \alpha(S(x)) = \alpha(S(x'))$.

The common feasible action set is $A(h) = \alpha(S(x))$ for any $x \in h$. Perfect recall is not imposed: an agent may have multiple nodes in h that lie on the same play of the game. When $|N| = 1$, the game is a *single-agent* SLU game and the agent assignment is trivial.

The decision nodes $X \setminus h$ not belonging to the information set are for *nature*: their successors are determined by ρ, not by any agent's choice. In games without a nature move, $h = X$. A *play* of the game is a maximal chain $\theta = t_0 \prec t_1 \prec \cdots \prec t_k \in Z$ with $\theta \in \Theta$. Two nodes $x, x' \in T$ are *co-reachable* if there exists a play that passes through them both.

Definition 2.2 (Self-Locating Uncertainty)**.** The information set h exhibits *self-locating uncertainty* if it contains distinct nodes that the occupying agent(s) cannot distinguish. Two forms arise:

(a) *Intra-personal self-locating uncertainty:* there exist co-reachable nodes $x, x' \in h$ with $\iota(x) = \iota(x')$. A single play of the game passes through h more than once, and the *same* agent occupies both nodes. This is the structure of the absent-minded driver, and in the language of Osborne and Rubinstein (1994), this is *absent-mindedness*, and not *forgetfulness*.

(b) *Inter-personal self-locating uncertainty:* there exist co-reachable nodes $x, x' \in h$ with $\iota(x) \neq \iota(x')$. A single play of the game passes through h more than once, and *different* agents occupy the two nodes. Each agent cannot distinguish their own node(s) from the other agents' nodes. This is the structure of Duplicating Sleeping Beauty: in the tails world, the original and the clone are visited sequentially on the same path of play, yet share an information set.

A game may exhibit both forms simultaneously.

The qualifier "co-reachable" is essential for self-locating uncertainty. In any extensive form with imperfect information, a player may have multiple nodes in the same information set that lie on different branches of the tree. Such nodes are not co-reachable: no single play visits both. This is standard imperfect information and does not constitute self-locating uncertainty. Intra-personal self-locating uncertainty arises only when a single play can pass through the agent's information set more than once, as in the absent-minded driver of Piccione and Rubinstein (1997). The perfect recall condition (KW2) of Kreps and Wilson (1982) rules out this possibility.

An *action profile* at h is a function $\mathbf{a} : h \to A(h)$ assigning an action to each node in h; when the agent plays a behavioural strategy, the realised actions at different nodes may differ (thought not if they play a pure or mixed strategy). For each state $\theta \in \Theta$ and action profile $\mathbf{a}$, let $D(\theta, \mathbf{a})$ denote the set of nodes in h that are visited on a play beginning at θ under $\mathbf{a}$ (i.e. the *dots* or *awakenings* in state θ under profile $\mathbf{a}$). Under a pure strategy in which the same action a is played at every node in h, I write $D(\theta, a)$ for this set. The game payoff in state θ under action profile $\mathbf{a}$ is $U(\mathbf{a}, \theta)$; under the pure strategy a, this is $U(a, \theta)$.

2.2 The Planning Stage

Before the game proper begins, I introduce a *planning stage* information set h_0, which precedes both the resolution of uncertainty and the main information set h. The planning stage precedes nature's choice of the state of the world: the temporal order is

$$h_0 \longrightarrow \text{Nature draws } \theta \in \Theta \text{ according to } \rho \longrightarrow h.$$

At h_0, the available actions are either $A(h_0) = \{\text{Commit}, \text{Not Commit}\}$ or $A(h_0) = \varnothing$ (the planning stage is inoperative). If $A(h_0) = \{\text{Commit}, \text{Not Commit}\}$, then choosing *Commit* binds the agent irrevocably to the 'optimal'(in the view of the planning agent) strategy at h before the state is realised, while *Not Commit* leaves the agent free to optimise at h given their beliefs upon arrival. If $A(h_0) = \varnothing$, the planning stage plays no role and the game begins directly with nature's move.

The planning stage serves two purposes. Firstly, it makes the *value of commitment* concrete: the Newcomb tension is precisely the situation in which the agent at h_0 would strictly prefer to Commit, because the planning-optimal action differs from the interim-optimal action at h. Secondly, it allows a direct representation of Newcomb's problem.

Example (Newcomb's Problem). Let $I = \{1\}$, with $A(h_0) = \{\text{Commit, Not Commit}\}$ and $A(h) = \{\text{One-box, Two-box}\}$. The state of the world $\theta \in \Theta = \{\text{Full}, \text{Empty}\}$ represents whether the opaque box contains \$1,000,000 or \$0, and is chosen by a highly reliable predictor *as a function of the action at* h_0: if the agent commits to one-boxing, $\theta = \text{Full}$ with high probability; if not, $\theta = \text{Empty}$ with high probability. The game payoff is determined by the action at h and the state θ in the usual way.

At h_0, the agent can guarantee a high expected payoff by committing to one-box. At

h, two-boxing is a dominant strategy for the interim-agent. This is the Newcomb tension: $a^{\text{plan}} = \text{One-box} \neq a^{\text{int}} = \text{Two-box}$.

This is the only example in the paper where the state of the world depends on the action at the planning stage. In all other settings, and for all the formal results and definitions that follow, the state θ is drawn independently of h_0. The Newcomb tension in self-locating uncertainty games arises not because the agent's plan causally influences the state (as the predictor does in Newcomb's problem), but because the awakening structure at h distorts the agent's beliefs about a state that is genuinely independent.

Note that this commitment/flexibility trade-off is reminiscent of but distinct from that of the principle behind mutually assured destruction (MAD). In that and similar settings, the advantage of commitment is that it allows deterrence and keeps certain decision nodes off-path. Conditional on actually arriving at such a node, planning and interim agents do not have different preferences. The Newcomb tension is about when the planning and interim agents disagree on the optimal action at *on-path* nodes.

Social welfare in multi-agent games. When $|N| > 1$, the planning agent must aggregate the payoffs of different agents. I assume that agents behind a veil of ignorance adopt a *utilitarian social welfare function* (SWF). For each agent $n \in N$, state $\theta \in \Theta$, and action profile $\mathbf{a} \in A(h)^{|h|}$, let $D(\theta, n, \mathbf{a}) \subseteq h$ denote the set of dots that agent n occupies in state θ when the actions realised along the path of play are given by $\mathbf{a}$, and let $|D(\theta, \mathbf{a})| = \sum_{n'} |D(\theta, n', \mathbf{a})|$ be the total number of dots. I assume throughout that $|D(\theta, \mathbf{a})|$ is deterministic for each state–action-profile pair $(\theta, \mathbf{a})$. In the Sleeping Beauty problem, the number of awakenings is determined by the state alone and does not depend on the action profile. In the absent-minded driver, the number of awakenings depends on the action profile, and there is no state. In both cases, $|D(\theta, \mathbf{a})|$ is a fixed (non-random) function of $(\theta, \mathbf{a})$.

The planner treats each state-action-profile pair $(\theta, \mathbf{a})$ as a compound state with probability $\rho(\theta)\ \Pr_\sigma(\mathbf{a})$, and within each such compound state distributes probability uniformly across its $|D(\theta, \mathbf{a})|$ dots. This formalises my treatment of duplicating Sleeping Beauty in the introduction. Agent n's social weight is then the total probability falling on n's dots:

$$\lambda_n^{\text{plan}}(\sigma) \;=\; \sum_{\theta \in \Theta} \sum_{\mathbf{a}} \rho(\theta)\ \Pr_\sigma(\mathbf{a})\ \frac{|D(\theta, n, \mathbf{a})|}{|D(\theta, \mathbf{a})|}.$$

These weights sum to one across all agents and reduce to $\lambda_n = 1/|N|$ when all agents share the same dots in every $(\theta, \mathbf{a})$. The planning payoff evaluates social welfare within each compound state $(\theta, \mathbf{a})$ using that compound state's social weights, then averages across compound states:

$$V^{\text{plan}}(\sigma) \;=\; \sum_{\theta \in \Theta} \sum_{\mathbf{a}} \rho(\theta)\ \Pr_\sigma(\mathbf{a}) \sum_{n \in N} \frac{|D(\theta, n, \mathbf{a})|}{|D(\theta, \mathbf{a})|}\, U_n(\mathbf{a}, \theta).$$

When $|N| = 1$ this reduces to the single-agent planning objective. When the dot structure does not depend on $\mathbf{a}$, as in the Sleeping Beauty variants, the weights factor out and V^{plan} simplifies to $\sum_n \lambda_n\, E_n(\sigma)$, where $E_n(\sigma)$ is agent n's expected payoff conditional on existing (Section 4). The key feature of multi-agent games is that the interim agent at h uses *thirder* probabilities to compute the analogous weights, which generically differ from the planner's weights; this is the source of the Newcomb tension studied in Section 4. Note also that much as the interim fixed point in games with endogenous awakening structures, such as the AMD, the planner's social weights do also vary with the behavioural strategy chosen, as this changes the probability shares of agents and thus their relative importance.

2.3 The Newcomb Tension

To define the Newcomb tension formally, we must define the beliefs and optimal strategies of agents at planning and interim stages.

Definition 2.3 (Planning-Optimal Strategy)**.** A *behavioural strategy* $\sigma \in \Delta(A(h))$ specifies the probability $\sigma(a)$ with which the agent independently plays action a at each visit to h. The *planning-optimal strategy* maximises the ex ante expected game payoff:

$$\sigma^* \;=\; \arg\max_{\sigma \in \Delta(A(h))} V^{\text{plan}}(\sigma), \qquad V^{\text{plan}}(\sigma) \;=\; \sum_{\theta \in \Theta} \sum_{\mathbf{a}} \rho(\theta) \; \Pr_{\sigma}(\mathbf{a})\, U(\mathbf{a}, \theta).$$

When I restrict attention to pure strategies, I write $a^{\text{plan}} = \arg\max_{a \in A(h)} V^{\text{plan}}(a)$ with $V^{\text{plan}}(a) = \sum_{\theta} \rho(\theta)\, U(a, \theta)$.

I refer to the Bayesian beliefs an agent would form ex-interim as 'thirder' beliefs, in reference to the Sleeping Beauty problem. One can still find epistemologists who argue the halfer position, but the literature has largely settled on the thirder stance, and indeed the absent-minded driver literature in game theory implicitly assumes this position too. Ross (2010) states generalised halfer and thirder principles that show how to follow halfer and thirder reasoning in more general situations of state-dependent awakening, though he does not consider games or situations with endogenous awakening. The following definition of thirder beliefs follows his, with agents forming beliefs *in a given equilibrium* σ by assuming that all agents (and temporal parts of agents) act according to the equilibrium strategy-profile.

Definition 2.4 (Thirder Beliefs)**.** If strategy σ is being played, a *thirder* assigns to each pair $(\theta, \mathbf{a})$ a probability proportional to the prior times the probability of the action profile times the number of awakenings:

$$\pi(\theta, \mathbf{a}) \;=\; \frac{\rho(\theta)\; \Pr_{\sigma}(\mathbf{a})\, |D(\theta, \mathbf{a})|}{\displaystyle\sum_{\theta', \mathbf{a}'} \rho(\theta')\; \Pr_{\sigma}(\mathbf{a}')\, |D(\theta', \mathbf{a}')|}.$$

The marginal over states is $\pi(\theta) = \sum_{\mathbf{a}} \pi(\theta, \mathbf{a})$. When $|D(\theta, \mathbf{a})|$ does not depend on $\mathbf{a}$ (as in Sleeping Beauty), this reduces to $\pi(\theta) = \rho(\theta)|D(\theta)|/C$ with $C = \sum_{\theta'} \rho(\theta')|D(\theta')|$. The agent forms beliefs over dots visited conditional on a given state-action-profile pair $(\theta, \mathbf{a})$ by distributing the probability uniformly over all these dots.

Definition 2.5 (Interim-Optimal Strategy)**.** The *interim-optimal strategy* for a thirder is the behavioural strategy $\sigma^{\text{int}} \in \Delta(A(h))$ that maximises the expected continuation payoff at h under thirder beliefs. The agent at a node in h chooses their action a' at the current node to maximise:

$$V^{\text{int}}(a') \;=\; \sum_{\theta,\,\mathbf{a}} \pi(\theta,\mathbf{a})\,\frac{1}{|D(\theta,\mathbf{a})|} \sum_{d \in D(\theta,\mathbf{a})} v_d(a',\,\mathbf{a}_{-d},\,\theta),$$

where $v_d(a', \mathbf{a}_{-d}, \theta)$ is the continuation payoff at dot d when the agent plays a' at d, the actions at all other dots are as in $\mathbf{a}_{-d}$, and the state is θ. Since the thirder beliefs depend on the strategy σ through $\Pr_\sigma(\mathbf{a})$, the interim-optimal strategy is a fixed point: σ^{int} is interim-optimal if, when all temporal parts play σ^{int}, each finds σ^{int} optimal given the resulting beliefs. Concretely, σ^{int} places positive weight only on actions a' that maximise V^{int}. When the game payoff depends only on the pure strategy played at every node (i.e. $v_d(a', \mathbf{a}_{-d}, \theta) = U(a', \theta)$ for all d and $\mathbf{a}_{-d}$), this reduces to $V^{\text{int}}(a') = \sum_\theta \pi(\theta)\, U(a', \theta)$.

Having defined these objects, I can now formally define the concept of a Newcomb tension:

Definition 2.6 (Newcomb Tension)**.** A single-information-set SLU game Γ exhibits a *Newcomb tension* if $\sigma^* \neq \sigma^{\text{int}}$: the planning-optimal and interim-optimal strategies diverge. The agent at the information set, reasoning correctly given their epistemically justified beliefs, chooses a strategy that is suboptimal from the planning perspective. The *value of commitment* is $V^{\text{plan}}(\sigma^*) - V^{\text{plan}}(\sigma^{\text{int}}) > 0$.

Note that the Newcomb tension can only arise when $|D(\theta, \mathbf{a})|$ varies across states or action profiles: if the number of awakenings is constant across all $(\theta, \mathbf{a})$, then $\pi(\cdot \mid \mathbf{a})$ is proportional to ρ and the two objectives have the same maximiser.

2.4 One-Boxer Beliefs: A Representation Theorem

I now formalise the relationship between commitment and beliefs. The central observation is that a Bayesian (thirder) agent who can commit to the planning-optimal action is behaviourally indistinguishable from an agent without commitment who holds different beliefs. I call these beliefs *one-boxer beliefs*, by analogy with one-boxing in Newcomb's problem.

To state the definition, I use the *continuation payoffs* already introduced in the interim-optimal strategy above. Recall that $v_d(a', \mathbf{a}_{-d}, \theta)$ denotes the payoff realised from dot d onwards when the agent plays a' at d, the actions at all other dots are $\mathbf{a}_{-d}$, and the state is θ. The continuation payoffs are *dot-independent* if $v_d(a', \mathbf{a}_{-d}, \theta)$ does not depend on d or on $\mathbf{a}_{-d}$ for any given a' and θ. In many games of interest, including all Sleeping Beauty variants, the game payoff depends on the action and the state but not on which dot the agent occupies, so $v_d(a', \mathbf{a}_{-d}, \theta) = U(a', \theta)$ for all d and $\mathbf{a}_{-d}$, and the continuation payoffs are trivially dot-independent. In other games, notably the absent-minded driver, the payoff depends on *where* in the game tree the action is taken: exiting at the first intersection yields payoff 0, while exiting at the second yields payoff 4.

Definition 2.7 (One-Boxer Beliefs)**.** A joint probability distribution P^{OB} over triples $(\theta, \mathbf{a}, d)$ with $d \in D(\theta, \mathbf{a})$ gives the *one-boxer beliefs for the planning-optimal strategy* σ^* if an interim agent holding P^{OB} and choosing their action finds σ^* optimal. Concretely, define the interim expected continuation payoff of playing a' at the current dot:

$$V^{\mathrm{OB}}(a') \;=\; \sum_{\theta, \mathbf{a}, d} P^{\mathrm{OB}}(\theta, \mathbf{a}, d)\, v_d(a', \,\mathbf{a}_{-d}, \,\theta).$$

Then P^{OB} is one-boxer if:

(a) for every a' in the support of σ^*, $V^{\mathrm{OB}}(a') \geq V^{\mathrm{OB}}(a'')$ for all $a'' \in A(h)$; and

(b) if σ^* is non-degenerate, then $V^{\mathrm{OB}}(a') = V^{\mathrm{OB}}(a'')$ for all a', a'' in the support of σ^* (indifference).

I write $P^{\mathrm{OB}}(\theta) = \sum_{\mathbf{a},d} P^{\mathrm{OB}}(\theta, \mathbf{a}, d)$ for the marginal over states.

This next definition simply notes that the thirder beliefs I have already defined lead the agent to select an interim-optimal strategy, and thus can also be called *two-boxer* beliefs to complete the Newcomb analogy that motivates the definition of *one-boxer* beliefs. In games with a Newcomb tension, interim EU-maximisation under two-boxer beliefs selects a strategy that is not planning-optimal.

Definition 2.8 (Two-Boxer (Bayesian) Beliefs)**.** The *two-boxer* or *Bayesian* beliefs at h are the thirder beliefs $\pi(\theta, \mathbf{a})$ defined in Definition 2.4.

Proposition 2.1 (Non-uniqueness of One-Boxer Beliefs)**.** One-boxer beliefs are not unique, in general many conditionals over dots will do.

(i) **(Universal marginal.)** Suppose continuation payoffs are dot-independent: $v_d(a', \mathbf{a}_{-d}, \theta) = v(a', \theta)$ for all d, $\mathbf{a}_{-d}$, and θ. Then the prior marginal $P^{\mathrm{OB}}(\theta) = \rho(\theta)$, combined with *any* conditional over action profiles and dots, constitutes one-boxer beliefs for every dot-independent payoff specification v.

(ii) **(Halfer conditionals need not suffice.)** When continuation payoffs are dot-dependent, the conditional beliefs $P^{\mathrm{OB}}(d \mid \theta, \mathbf{a})$ matter, and the uniform ("halfer") conditional $P^{\mathrm{OB}}(d \mid \theta, \mathbf{a}) = 1/|D(\theta, \mathbf{a})|$ need not be one-boxer. In the absent-minded driver (restricting to pure strategies for illustration), the pure strategy Continue yields two dots $D = \{X, Y\}$. Halfer beliefs assign $P^{\mathrm{OB}}(X) = P^{\mathrm{OB}}(Y) = \frac{1}{2}$, but one-boxer beliefs require $P^{\mathrm{OB}}(X) \geq \frac{3}{4}$.

Proof. See Appendix B.1. □

Theorem 2.2 (Representation Theorem)**.** A Bayesian agent who holds thirder beliefs π at the information set but has commitment power is behaviourally equivalent to an interim expected-utility maximiser who holds one-boxer beliefs P^{OB} and has no commitment power.

Proof. The result is true by definition: one-boxer beliefs are defined (Definition 2.7) as beliefs under which interim optimisation reproduces the planning-optimal strategy. The committed Bayesian plays σ^* at the planning stage. The uncommitted one-boxer, holding P^{OB} and maximising $V^{\text{OB}}(a')$, finds σ^* optimal by the one-boxer conditions (a) and (b). □

The committed thirder plays σ^* because they are bound to the plan; the uncommitted one-boxer plays σ^* because their beliefs P^{OB} make it interim-optimal. When the planning-optimal strategy is a non-degenerate mixture, the one-boxer is indifferent between the actions in its support: the belief revision from π to P^{OB} equalises the expected continuation payoffs that the thirder's beliefs would have ranked differently. Given this result, if we want to model agents with commitment power in a game, we do not need to formalise a planning stage and the ability to commit. Instead, we can use this *as-if* representation and just assume they hold one-boxer beliefs instead.

3 Single Agent Results

I now characterise when the Newcomb tension arises in single-information-set SLU games, when it can be resolved, and how the halfer–thirder debate relates to commitment. Throughout, I write Θ for the set of states (initial nodes), $D(\theta, \mathbf{a})$ for the set of dots at h in state θ under action profile $\mathbf{a}$, and $U(\mathbf{a}, \theta)$ for the game payoff. When the number of dots does not depend on the action profile (as in Sleeping Beauty), I write simply $D(\theta)$.

3.1 Behavioural Strategies Resolve the Newcomb Tension

The Newcomb tension described in Section 2 arises when the agent is restricted to pure strategies. I now show that when the agent can *randomise*, adopting a behavioural strategy $\sigma \in \Delta(A(h))$, the tension vanishes entirely, regardless of the number of states.

Theorem 3.1 (Randomisation Resolution)**.** In any single-agent SLU game with a single information set h, the planning-optimal behavioural strategy σ^* is also interim-optimal: an agent waking at an unknown dot in h, taking all other temporal parts' strategy as σ^* and maximising their own continuation payoff, finds σ^* optimal. This holds for any number of states $|\Theta| \geq 1$.

Proof. See Appendix B.2. □

This result generalises the classical resolution of the absent-minded driver problem. In the standard AMD, which has a single state ($|\Theta| = 1$), the planning-optimal behavioural strategy $q^* = 2/3$ is also the interim fixed point. Theorem 3.1 shows that this coincidence does not depend on their being only one state of the world, or on the specific properties of the AMD: for *any* single-agent, single-information-set SLU game, per-dot continuation-payoff reasoning recovers the planning optimum, regardless of how many states the game has or whether payoff-irrelevant dots are present.

A notable feature of the AMD is that the Newcomb tension under pure strategies is not merely a divergence between planning and interim optima: the pure-strategy interim fixed point does not even *exist*, as I explain in A.3. It is thanks to the introduction of randomisation that the interim fixed point can exist, and as Theorem 3.1 shows, this fixed point necessarily coincides with the planning optimum. In this sense, randomisation is a stone that kills two birds: it creates the interim fixed point and resolves the Newcomb tension in a single step. The following result formalises this observation: whenever a pure-strategy interim fixed point exists, it must be planning-optimal even without randomisation.

Proposition 3.2 (Pure-Strategy Fixed Points Resolve the Newcomb Tension)**.** In any single-agent SLU game with a single information set h, if a pure strategy $a \in A(h)$ is an interim fixed point (i.e. if the interim agent, believing the other temporal parts all play a, also finds a optimal) then a is planning-optimal among pure strategies: $a = a^{\text{plan}}$.

Proof. See Appendix B.3. □

3.1.1 Multi-State Example

Consider a game with two equiprobable states, $\theta \in \{0, 1\}$ with $\rho(0) = \rho(1) = 1/2$. In state $\theta = 0$, the game is the standard AMD: the agent traverses intersections X_0 and Y_0, with payoffs 0 (Exit at X_0), 4 (Exit at Y_0), and 1 (Continue past Y_0). In state $\theta = 1$, an additional intersection Z_1 is inserted between X_1 and Y_1, with Exit at Z_1 leading to its own terminal node with payoff 3. Exit at Y_1 gives 5 (rather than 4), and the remaining payoffs are unchanged. See Figure 2.

Under behavioural strategy σ, the state payoffs are

$$V_0(\sigma) = 4\sigma - 3\sigma^2,$$
$$V_1(\sigma) = 3\sigma + 2\sigma^2 - 4\sigma^3,$$

where V_1 reflects the three-intersection chain: Exit at X_1 gives 0, Exit at Z_1 (reached with probability σ) gives 3, Exit at Y_1 (reached with probability σ^2) gives 5, and Continue past Y_1 (probability σ^3) gives 1.

Planning-optimal strategy. The planning payoff is

$$V^{\text{plan}}(\sigma) \;=\; \tfrac{1}{2}(4\sigma - 3\sigma^2) + \tfrac{1}{2}(3\sigma + 2\sigma^2 - 4\sigma^3) \;=\; \frac{7\sigma - \sigma^2 - 4\sigma^3}{2}.$$

Setting $dV^{\text{plan}}/d\sigma = (7 - 2\sigma - 12\sigma^2)/2 = 0$ gives $12\sigma^2 + 2\sigma - 7 = 0$, hence

$$\sigma^* \;=\; \frac{-1 + \sqrt{85}}{12} \;\approx\; 0.685.$$

Interim fixed point. The normalising constant is $C(\sigma) = \frac{1}{2}(1+\sigma) + \frac{1}{2}(1+\sigma+\sigma^2) = 1+\sigma+\frac{1}{2}\sigma^2$. The continuation-payoff derivatives at each dot are:

$$g_{X_0} = 4 - 3\sigma, \qquad g_{Y_0} = -3,$$
$$g_{X_1} = 3 + 2\sigma - 4\sigma^2, \qquad g_{Z_1} = 2 - 4\sigma,$$
$$g_{Y_1} = -4.$$

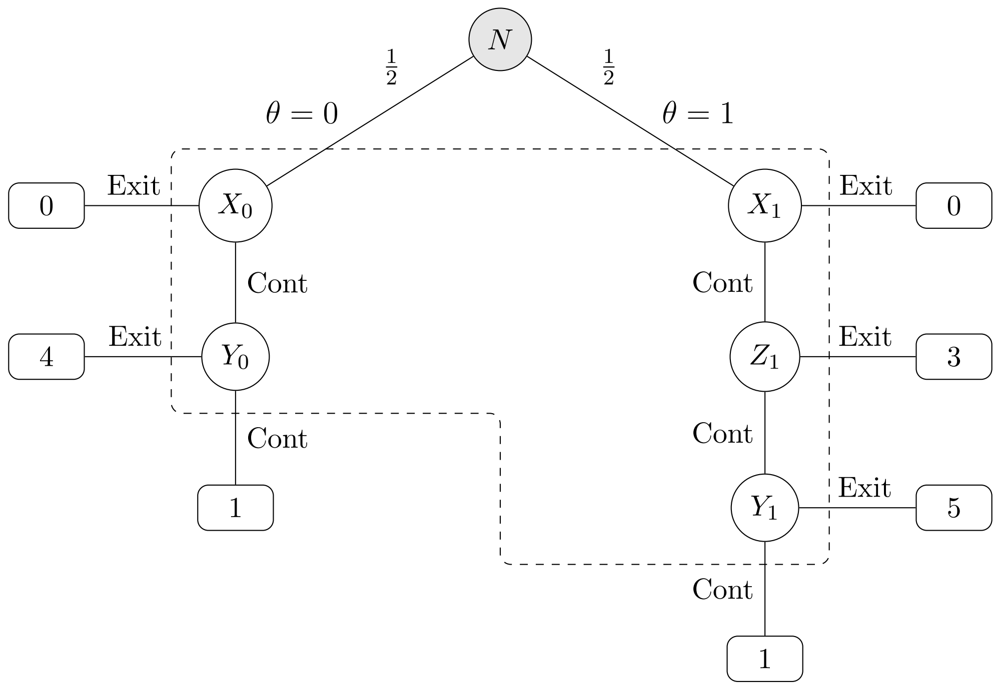


Figure 2: Sleeping Beauty Behind the Wheel. In $\theta = 0$, the standard AMD with intersections X_0 and Y_0. In $\theta = 1$, an additional intersection Z_1 is inserted, with Exit at Z_1 leading to its own terminal node with payoff 3. Exit at Y_1 pays 5 (versus 4 at Y_0). All five decision nodes lie in a single information set (dashed box).

The interim first-order condition is

$$\text{FOC}^{\text{int}}(\sigma) = \frac{7 - 2\sigma - 12\sigma^2}{2\,C(\sigma)} = 0 \qquad \Longrightarrow \qquad \hat{\sigma} = \frac{-1 + \sqrt{85}}{12} \approx 0.685.$$

The planning optimum and the interim fixed point coincide, as guaranteed by Theorem 3.1. The additional intersection Z_1 in $\theta = 1$ inflates the number of awakenings in that state (three dots versus two in $\theta = 0$), causing the thirder to overweight $\theta = 1$ — exactly analogous to the Monday-in-Tails awakening in Sleeping Beauty. Yet this belief distortion does not create a Newcomb tension: the per-dot continuation-payoff reasoning of Theorem 3.1 absorbs it.

4 Multiple Agent Results

I now turn to games in which multiple agents occupy dots at the same information set. I assume co-cardinal utility throughout. For each state of the world, $\theta \in \Theta$, agent $n \in N$, and action profile $\mathbf{a}$, let $D(\theta, n, \mathbf{a}) \subseteq h$ denote the dots that agent n visits in state θ under action profile $\mathbf{a}$, and I write $D(\theta, \mathbf{a}) = \bigcup_{n \in N} D(\theta, n, \mathbf{a})$ for the full set of dots visited. Since the agent assignment ι assigns each dot to exactly one agent, the sets $D(\theta, n, \mathbf{a})$ are disjoint across agents, and $|D(\theta, \mathbf{a})| = \sum_n |D(\theta, n, \mathbf{a})|$. I use $h_n(\theta) = \bigcup_{\mathbf{a}} D(\theta, n, \mathbf{a})$ for the set of all dots reachable by agent n in state θ, and say that agent n *exists* in state θ whenever $h_n(\theta) \neq \varnothing$.

The planning payoff is a social welfare function with halfer-derived social weights. As in Section 2, the planner treats each state-action-profile pair $(\theta, \mathbf{a})$ as a compound state with probability

$\rho(\theta)\ \Pr_\sigma(\mathbf{a})$, distributes probability uniformly across the $|D(\theta,\mathbf{a})|$ dots in each compound state, and evaluates social welfare accordingly:

$$V^{\text{plan}}(\sigma) \;=\; \sum_{\theta\in\Theta}\sum_{\mathbf{a}}\rho(\theta)\ \Pr_\sigma(\mathbf{a})\sum_{n\in N}\frac{|D(\theta,n,\mathbf{a})|}{|D(\theta,\mathbf{a})|}\,U_n(\mathbf{a},\theta).$$

The *social weight* of agent n is the total probability falling on n's dots:

$$\lambda_n(\sigma) \;=\; \sum_{\theta\in\Theta}\sum_{\mathbf{a}}\rho(\theta)\ \Pr_\sigma(\mathbf{a})\ \frac{|D(\theta,n,\mathbf{a})|}{|D(\theta,\mathbf{a})|}.$$

When the dot structure does not depend on the action profile (as in the Sleeping Beauty variants, where agent existence is determined by the state alone) the per-compound-state weights $|D(\theta,n,\mathbf{a})|/|D(\theta,\mathbf{a})|$ do not vary with $\mathbf{a}$, and V^{plan} simplifies to $\sum_n \lambda_n\, E_n(\sigma)$, where

$$E_n(\sigma) \;=\; \frac{\sum_{\theta:\,h_n(\theta)\neq\varnothing}\rho(\theta)\,U_n(\sigma,\theta)}{\sum_{\theta:\,h_n(\theta)\neq\varnothing}\rho(\theta)}$$

is agent n's expected payoff conditional on existing. In games where the dot structure depends on the action profile, such as the duplicating absent-minded driver, this factorisation does not hold, and the full compound-state formula must be used.

The Newcomb tension arises because the planner and the interim agent assign different social weights to the agents. Both condition on the prior ρ over states of the world and the induced distribution over action profiles $\mathbf{a}$, and both distribute probability uniformly over the dots visited in each state-action-profile pair $(\theta,\mathbf{a})$. The difference is that the thirder (Definition 2.4) adjusts the probability of each $(\theta,\mathbf{a})$ proportionally to the number of dots $|D(\theta,\mathbf{a})|$, while the planner does not. This means the thirder assigns larger social weight to agents who appear in state-action-profile pairs with more awakenings, while the planner does not make this adjustment. Since the number of dots varies across state-action-profile pairs whenever the dot structure is asymmetric, the two sets of social weights generically disagree, and the planning-optimal and interim-optimal actions diverge.

Theorem 4.1 (Multi-Agent Newcomb Tension)**.** Consider a multi-agent SLU game with co-cardinal utility and a single information set h.

(i) *(Symmetric dot structure ⇒ no tension.)* If $|D(\theta,n,\mathbf{a})| = |D(\theta,n',\mathbf{a})|$ for all $n,n'\in N$, $\theta\in\Theta$, and $\mathbf{a}$ (every agent visits the same number of dots in every state-action-profile pair) then the equation

$$\frac{dV^{\text{plan}}}{d\sigma} = \frac{C(\sigma)}{|N|}\cdot \text{FOC}^{\text{int}}(\sigma)$$

holds, and per-dot continuation-payoff reasoning of Theorem 3.1 recovers the planning optimum. There is no Newcomb tension.

(ii) *(Asymmetric dot structure generically ⇒ tension.)* If $|D(\theta_0,n_0,\mathbf{a}_0)| \neq |D(\theta_0,n_1,\mathbf{a}_0)|$ for some $(\theta_0,\mathbf{a}_0)$ and agents n_0,n_1 (the dot structure is asymmetric across agents) then the above equation generically fails to hold, and the planning-optimal and interim-

optimal strategies diverge (Version (a) and Duplicating Sleeping Beauty Behind the Wheel illustrate this case).

Proof. See Appendix B.4. □

Part (i) shows that payoff disagreement alone does not create a Newcomb tension: even if $u_{n_1} \neq u_{n_2}$, the proportional planning-interim equation holds as long as every agent visits the same number of dots in every state-action-profile pair. Part (ii) shows that when the dot structure *is* asymmetric, the tension generically arises.

4.1 Duplicating Sleeping Beauty

Have established this general result on Newcomb tensions in multi-agent games, I now discuss some important special cases. Firstly, this allows us to comment on the duplicating Sleeping Beauty problem, as we began to do in the Introduction:

Proposition 4.2 (Newcomb Tension in Duplicating Sleeping Beauty)**.** In Duplicating Sleeping Beauty, there is a Newcomb tension between the planning-optimal and interim-optimal reports, and randomisation does nothing to resolve it. Denoting their strictly proper scoring rule[1] $S(x, \theta)$, then we have:

(i) *(Planning stage.)* The planner assigns halfer social weights: the original agent accounts for all awakenings in Heads and half in Tails, giving $\lambda_{\text{orig}} = 3/4$; the clone accounts for the remaining half of Tails only, giving $\lambda_{\text{clone}} = 1/4$. The original's expected score, using the prior conditional on existing, is $E_{\text{orig}}(x) = \frac{1}{2}\, S(x, H) + \frac{1}{2}\, S(x, T)$; the clone, existing only in Tails, has $E_{\text{clone}}(x) = S(x, T)$. The planning-optimal report maximises

$$V^{\text{plan}}(x) \;=\; \tfrac{3}{4}\, E_{\text{orig}}(x) \;+\; \tfrac{1}{4}\, E_{\text{clone}}(x) \;=\; \tfrac{3}{8}\, S(x, H) \;+\; \tfrac{5}{8}\, S(x, T),$$

yielding a report x^{plan} that places more weight on Heads than on Tails.

(ii) *(Interim stage.)* An agent who wakes and does not know which agent they are uses thirder probabilities, assigning social weight $\lambda^{\text{int}}_{\text{orig}} = 2/3$ and $\lambda^{\text{int}}_{\text{clone}} = 1/3$. Each agent's conditional expected score is the same as in part (i). The interim-optimal report maximises

$$V^{\text{int}}(x) \;=\; \tfrac{2}{3}\, E_{\text{orig}}(x) \;+\; \tfrac{1}{3}\, E_{\text{clone}}(x) \;=\; \tfrac{1}{3}\, S(x, H) \;+\; \tfrac{2}{3}\, S(x, T),$$

yielding $x^{\text{int}} \neq x^{\text{plan}}$ whenever S is strictly proper.

(iii) *(Randomisation is ineffective.)* Since there is only one payoff-relevant dot per agent, the game has no coordination structure: each agent would ideally take the same action at every visit. Mixing over reports cannot improve the planning payoff, and the Newcomb tension survives it.

Proof. See Appendix B.5. □

[1] A scoring rule is *strictly proper* if the expected score is uniquely maximised when the report equals the agent's true belief. E.g. the Brier score of the introduction.

Kierland and Monton (2005) discuss the canonical and duplicating Sleeping Beauty Problems where Beauty is attempting to maximise a Brier score, and argue that halfer beliefs are justified in the latter problem since they minimise the expected average error. Conversely, in the standard problem one should be a thirder, they believe, since the solution $\frac{1}{3}$ minimises total expected error. In their view, the two positions simply amount to pursuing different epistemic goals, where the one is more reasonable in the canonical problem, and the other in the duplicating variant. The stance I argue here is of course different, I take for granted that the thirder position is the epistemically correct answer in both versions, but approach the question of what bet one should choose as a social choice problem.[2] With this social choice approach, we would need to assign a weight of 1 to the original Beauty to justify a choice of $\frac{1}{2}$. Janda (2024) argues instead simply that we can justify *different* answers to the canonical and duplicating problems because they should be modelled as different games, one with absent-mindedness and the other forgetfulness. Epistemically, I disagree, but agree that agents trying to maximise expected utility should make different bets in each case, and of course am arguing that whereas the one problem exhibits a Newcomb tension, the other does not. I discuss Janda's paper at greater length in *Sleeping Beauty's Dismal Day Out*, where I discuss how one should model games with self-locating uncertainty in general.

4.2 Duplicating AMD

The absent-minded driver normally has no state of the world: the game tree is deterministic, with dots X and Y at a single information set. In the *duplicating* version, a clone (Player 2) is created when Player 1 continues at X: the red dot on the Continue branch marks the clone's creation, and Player 2 occupies Y. I adopt the convention that red dots may only precede action nodes (never terminal nodes), and that a terminal node carries a payoff for the clone only if the clone has been created somewhere on its branch. Under this convention, Exit at X occurs before the clone exists, so its payoff is a scalar for Player 1 alone (see Figure 3). Since Exit at X precedes the clone's creation, $D(n{=}1) = \{X\}$ and $D(n{=}2) = \{Y\}$: Player 1 occupies only X, while the clone occupies only Y. The normalising constant is

$$C(\sigma) = (1-\sigma)\cdot 1 + \sigma \cdot 2 = 1 + \sigma.$$

Version (a): identical payoffs. Here $u_1 = u_2$ at every terminal where both players are defined: Exit at Y gives $(4, 4)$ and Continue past Y gives $(1, 1)$. Under behavioural strategy σ, the social weights are $\lambda_1(\sigma) = 1 - \sigma/2$ and $\lambda_2(\sigma) = \sigma/2$: under Exit at X (probability $1-\sigma$), only Player 1 is present; under Continue (probability σ), both players have one dot each, giving equal weight $1/2$. Since $u_1 = u_2$ at every terminal, the per-compound-state social welfare reduces to the common payoff regardless of the weights, and the planning payoff is simply

$$V^{\text{plan}}(\sigma) = u_1(\sigma) = 4\sigma - 3\sigma^2, \qquad \sigma^* = 2/3.$$

The thirder beliefs are

$$P(n{=}1, X) = \frac{1}{1+\sigma}, \qquad P(n{=}2, Y) = \frac{\sigma}{1+\sigma}.$$

[2]To my knowledge, this is novel in the literature on Sleeping Beauty.

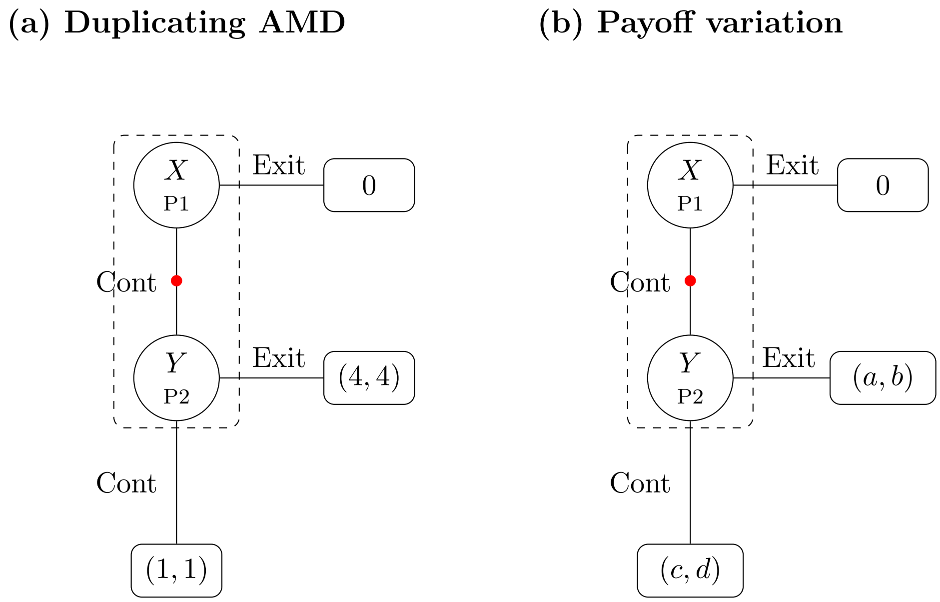


Figure 3: Duplicating AMD. A red dot on the Continue branch marks the creation of the clone (Player 2). Since red dots may only precede action nodes, Exit at X occurs before the clone exists and carries a scalar payoff. In (a), the remaining payoffs are identical across players. In (b), Exit at Y gives (a, b) and Continue past Y gives (c, d), allowing payoffs to diverge across players.

The continuation-payoff derivatives at each dot are $g_X = 4 - 3\sigma$ (Player 1's gain from continuing at X: expected payoff $4(1-\sigma)+\sigma$ versus 0) and $g_Y = 1-4 = -3$ (Player 2's gain from continuing at Y). The interim first-order condition is

$$\mathrm{FOC}^{\mathrm{int}}(\sigma) = \frac{1 \cdot (4 - 3\sigma) + \sigma \cdot (-3)}{1+\sigma} = \frac{4 - 6\sigma}{1+\sigma} = 0 \qquad \Longrightarrow \qquad \hat{\sigma} = \frac{2}{3}.$$

The planning optimum and the interim fixed point coincide: $\sigma^* = \hat{\sigma} = 2/3$. Proportionality holds: $dV^{\mathrm{plan}}/d\sigma = 4 - 6\sigma = C \cdot \mathrm{FOC}^{\mathrm{int}}$. Since $u_1 = u_2$, the compound-state social welfare is the common payoff in every action profile, and the planner's problem is equivalent to a single-agent optimisation. The duplication does not by itself create a Newcomb tension; for that, the players' payoffs must diverge.

Version (b): general clone payoffs. Now suppose exiting at Y gives (a, b) and continuing past Y gives (c, d) with $a, b, c, d > 0$, while exiting at X remains a scalar 0. The dot structure is unchanged: $D(n{=}1) = \{X\}$, $D(n{=}2) = \{Y\}$, and $C(\sigma) = 1 + \sigma$. Since the dot structure depends on the action profile, the full compound-state formula is required. The three action profiles and their contributions to V^{plan} are:

Profile	Prob	Dots	Weights (w_1, w_2)	Payoffs (U_1, U_2)
Exit at X	$1-\sigma$	1 (P1 only)	(1, 0)	(0, −)
Cont–Exit	$\sigma(1-\sigma)$	2	(1/2, 1/2)	$(a,\ b)$
Cont–Cont	σ^2	2	(1/2, 1/2)	$(c,\ d)$

The planning payoff is therefore

$$V^{\text{plan}}(\sigma) = \sigma(1-\sigma)\,\frac{a+b}{2} + \sigma^2\,\frac{c+d}{2} = \frac{(a{+}b)\,\sigma - (a{+}b{-}c{-}d)\,\sigma^2}{2},$$

which is quadratic in σ, with

$$\sigma^* = \frac{a+b}{2(a+b-c-d)}.$$

The continuation-payoff derivatives are $g_X = a - (a-c)\sigma$ (Player 1's gain from continuing at X) and $g_Y = d - b$ (Player 2's gain from continuing at Y). The interim first-order condition gives

$$C(\sigma)\cdot \text{FOC}^{\text{int}}(\sigma) = a - (a-c+b-d)\sigma,$$

so $\hat{\sigma} = a/(a+b-c-d)$. The no-tension condition $\sigma^* = \hat{\sigma}$ requires $(a+b)/2 = a$, i.e. $b = a$. Note that $b = a$ does not require $c = d$: the tension vanishes whenever the two agents' payoffs coincide at the Exit terminal (Y), even if they differ at the Continue terminal. The game reduces to Version (a) only when $b = a$ and $c = d$. For $b \neq a$:

Proposition 4.3 (Persistent Tension)**.** Suppose 'Exit at X' gives a scalar payoff of 0 for Player 1, 'Exit at Y' gives (a, b), and 'Continue past Y' gives (c, d) with $a, b, c, d > 0$ and $a+b > c+d$. The Newcomb tension is present if and only if $b \neq a$, with

$$\sigma^* - \hat{\sigma} = \frac{b-a}{2(a+b-c-d)}.$$

The planner favours more continuation than the interim agent when $b > a$ (the clone benefits more from reaching Y than the original) and less when $b < a$.

Proof. $\sigma^* = (a+b)/[2(a+b-c-d)]$ and $\hat{\sigma} = a/(a+b-c-d)$. Subtracting: $\sigma^* - \hat{\sigma} = (a+b-2a)/[2(a+b-c-d)] = (b-a)/[2(a+b-c-d)]$. □

The tension arises from the interaction between the dot asymmetry and the payoff asymmetry. Under the compound-state formula, 'Exit at X' contributes only Player 1's payoff (with full weight), while the Continue profiles split weights equally between the players. When $b > a$, the planner's equal weighting of the clone's higher payoff at Y makes continuation more attractive than the interim agent perceives; when $b < a$, the reverse holds. At $b = a$ the tension vanishes, even if $c \neq d$: the payoff asymmetry at the Continue terminal does not matter because the relevant wedge is driven by the Exit-at-Y payoffs, where the dot asymmetry bites. When additionally $c = d$, the game reduces to Version (a).

Duplicating Sleeping Beauty Behind the Wheel. The same construction applies to the multi-state game of Figure 2. A clone (Player 2) is created on the first Continue branch in each state: in $\theta = 0$, the clone appears at Y_0; in $\theta = 1$, at Z_1. Exit at X_0 or X_1 occurs before the clone exists, so the payoff is a scalar for Player 1 alone. All subsequent payoffs are vectors with equal components. See Figure 4.

Since $u_1 = u_2$ at every terminal where both are defined, the planning payoff is the same as in the non-duplicating game:

$$V^{\text{plan}}(\sigma) = \frac{7\sigma - \sigma^2 - 4\sigma^3}{2}, \qquad \sigma^* = \frac{-1+\sqrt{85}}{12} \approx 0.685.$$

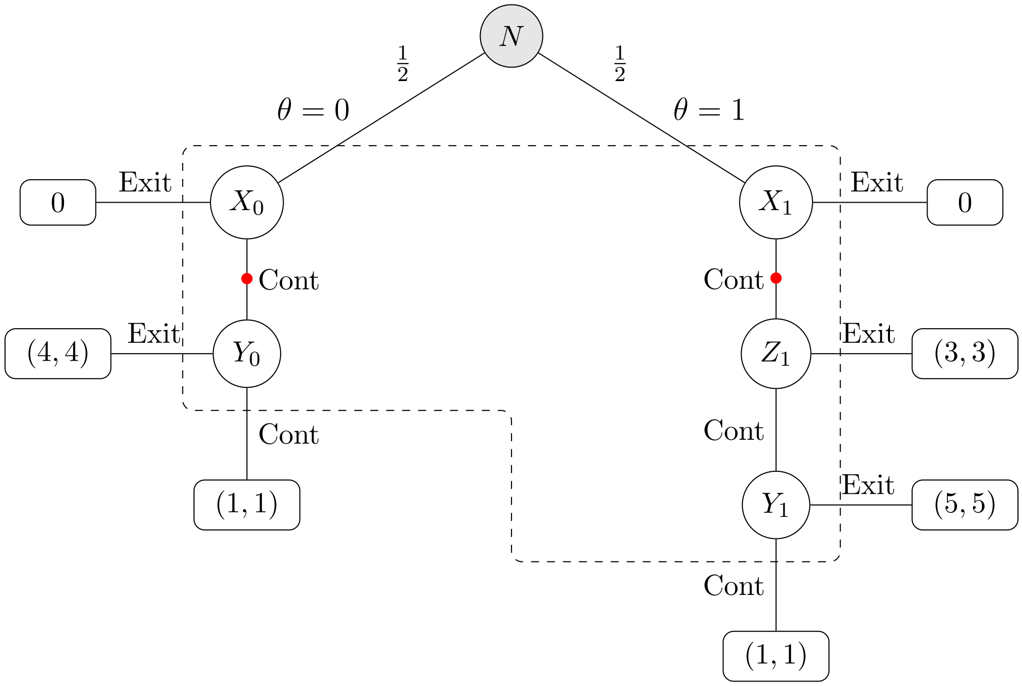


Figure 4: Duplicating Sleeping Beauty Behind the Wheel. Red dots on the first Continue branch of each state mark the creation of the clone. Exit at X_0 or X_1 gives a scalar payoff (Player 1 only); all subsequent payoffs are vectors (u_1, u_2) with equal components. All five decision nodes lie in a single information set.

Interim fixed point. Under the corrected dot structure, Player 1 occupies one dot per state (X_0 in $\theta = 0$, X_1 in $\theta = 1$) and Player 2 occupies the remaining dots (Y_0 in $\theta = 0$; Z_1 and Y_1 in $\theta = 1$). The normalising constant is

$$C(\sigma) = \tfrac{1}{2}(1+\sigma) + \tfrac{1}{2}(1+\sigma+\sigma^2) = 1 + \sigma + \tfrac{1}{2}\sigma^2.$$

The continuation-payoff derivatives at each dot are the same as in the non-duplicating game: $g_{X_0} = 4 - 3\sigma$, $g_{Y_0} = -3$, $g_{X_1} = 3 + 2\sigma - 4\sigma^2$, $g_{Z_1} = 2 - 4\sigma$, $g_{Y_1} = -4$. Since $u_1 = u_2$, each agent's continuation-payoff derivative at their dot is the same as the single agent's would be, and the interim FOC sums over exactly the same set of dot–derivative products:

$$\begin{aligned} C \cdot \mathrm{FOC}^{\mathrm{int}} &= \tfrac{1}{2}\big[1 \cdot (4 - 3\sigma) + \sigma(-3)\big] + \tfrac{1}{2}\big[1 \cdot (3 + 2\sigma - 4\sigma^2) + \sigma(2 - 4\sigma) + \sigma^2(-4)\big] \\ &= \tfrac{1}{2}(4 - 6\sigma) + \tfrac{1}{2}(3 + 4\sigma - 12\sigma^2) \\ &= \frac{7 - 2\sigma - 12\sigma^2}{2} = \frac{dV^{\mathrm{plan}}}{d\sigma}. \end{aligned}$$

The proportionality holds, and the interim fixed point coincides with the planning optimum: $\hat{\sigma} = \sigma^* \approx 0.685$. As in Version (a) of the duplicating AMD, the duplication alone does not create a Newcomb tension when the players' payoffs are identical. Tension would arise if the payoffs diverged across players at some terminal, by the same mechanism as Proposition 4.3.

5 Literature Review

This paper touches on a few different areas of existing literature. First of all, it adopts the notation of Kreps and Wilson (1982). However, the companion article *Sleeping Beauty's Dismal Day Out* does more to develop Kreps and Wilson's article than this one, since in that article I consider how one should model games with self-locating uncertainty in general, and to what extent one can extend their sequential equilibrium existence result to them. In this paper, beyond adopting their notation, I do not consider similar questions or extend their results.

Beyond this, however, this article *does* contribute the to literature on the Sleeping Beauty problem, and specifically the section of that literature that discusses whether the credences of an awakened agent should be the same in the canonical and duplicating problems. I have already discussed this literature (specifically Kierland and Monton (2005); Janda (2024)) in Section 4.1, but note here that in my consideration of the novel game of the *duplicating absent-minded driver*, I extend the debate to general games with inter-personal self-locating uncertainty, and take a social choice approach to the problem. This literature has not discussed the formation of self-locating beliefs in games with interpersonal self-locating uncertainty with endogenous awakenings. To the best of my knowledge, each of these points is original to this paper.

Concerning the absent-minded driver, summarised in Appendix A.3, the main collection of articles within economic theory is that of the original special issue of *Games and Economic Behavior* on Imperfect Recall (Volume 20, Issue 1, July 1997). As suggested by Aumann et al. (1997), I take the *modified multi-selves approach* of Piccione and Rubinstein (1997) to be the straightforwardly correct approach in such games. An agent always considers changing the action they control in a given awakening assuming they will follow the candidate strategy in all other awakenings. My Theorem 3.1 builds on their analysis directly, and generalises the fact, which they of course observe themselves, that with behavioural strategies the 'paradox of the absent-minded driver' (what I call a Newcomb tension) disappears. I show that in single-information set games with only one agent, introducing multiple states does not defeat this resolution. In the multi-agent *duplicating absent-minded driver*, this then changes, and the paradox can defeat even behavioural strategies whenever there is payoff divergence.

Beyond the economic literature, two articles in the philosophical literature are worth discussing here in relation to the absent-minded driver: Spohn (2025) and Schwarz (2015). Spohn's article argues, as per its title: "*The Puzzle of the Absent-Minded Driver is Not About Absent-Mindedness, but about Indexical Belief.*" In particular, I consider the topic of games with self-locating uncertainty interesting in so far as they can be used to study learning games with temporal ignorance (my own article *Bot Got Your Tongue?*, is itself a model of social learning in the presence of some self-locating uncertainty, though in the context of that paper this is a minor point). In this sense, therefore, I completely agree with Spohn. His observing that the proportionality principle he discusses (the 'R-principle') also applies in cases of mere ignorance concerning one's own temporal location points in this direction. However, as is clear in the contrast between the duplicating and canonical absent-minded driver problems, I think that different forms of indexical uncertainty are importantly different, and whether or not agents behave *as-if* they are forming beliefs according to his R-principle will depend on whether or not they have commitment power.

Moving on to Schwarz (2015), this article discusses the absent-minded driver and the Sleeping Beauty problem, as well as mentioning Newcomb's problem and the importance of the distinction between causal and evidential decision theory. This is the only other article of which I am aware that draws a connection between all three problems. He considers how the agent should behave as an adherent of causal and evidential decision theory. Unlike my approach here, he models the agent as forming beliefs of the 'states' (First & Continue_2, First & Leave_2, Second), and considers the implications of all four combinations of halfer/thirder and evidential/causal decision theory. A central finding of his is that under some of these four combinations, randomisation (the 'useless coins' of the title) helps the driver, while under others it does not. My Theorem 3.1 provides a contrast with this, as in my setting randomisation always helps (in single agent games) by removing the Newcomb tension. On the other hand, my approach sidesteps the EDT/CDT axis entirely; I assume causal decision theory throughout, as is standard in game theory, whereas for Schwarz this distinction drives half the analysis. Finally, Schwarz treats the AMD as a single-agent problem throughout. The multi-agent setting of my duplicating AMD, in which the Newcomb tension can survive behavioural strategies whenever payoffs diverge across agents (Proposition 4.3), lies outside the scope of his framework.

6 Conclusion

This paper has studied the Newcomb tension in games with self-locating uncertainty, restricting attention to games with a single information set. I have drawn an analogy between these games and Newcomb's problem: in both settings, an agent at the point of decision faces a conflict between locally rational reasoning and globally optimal planning for on-path decision nodes. In SLU games, the locally rational agent uses thirder beliefs (the epistemically correct self-locating credences in my view and the view of most epistemologists) while the planner evaluates strategies ex ante. The one-boxer belief representation theorem (Theorem 2.2) establishes that we can simply assume agents act with one-boxer beliefs in place of assuming commitment power: a committed Bayesian who holds thirder beliefs and one who holds one-boxer beliefs are behaviourally equivalent. The one-boxer beliefs that restore planning optimality may coincide with halfer beliefs, but are instrumentally motivated; they are the beliefs an agent would need to hold in order for interim reasoning to reproduce the planning-optimal strategy without commitment power.

In the single-agent case, this divergence is always resolved by randomisation. Behavioural strategies do kill two birds with one stone: they create the interim fixed point, which may not exist under pure strategies (as in the absent-minded driver), and they ensure that this fixed point coincides with the planning optimum (Theorem 3.1). This holds for any game corresponding to Definition 2.1 with a single agent, regardless of the number of states of the world. On the other hand, when a pure-strategy interim fixed point does exist, it must already be planning-optimal (Proposition 3.2), and randomisation adds nothing. This is illustrated by the Newcomb tension in Duplicating Sleeping Beauty (Proposition 4.2).

In multi-agent games, I have approached the question of what bet an awakened agent should make as a social choice problem, aggregating the payoffs of different agents with a utilitarian social welfare function. This framework treats each state-action-profile pair $(\theta, \mathbf{a})$ as a compound state, distributes probability uniformly across the 'dots' or awakenings within each, and evaluates welfare accordingly. To my knowledge, this social choice approach is novel in the literature on Sleeping Beauty. With multiple agents, the Newcomb tension is generically present with an asymmetric awakening structure across agents (Theorem 4.1). Awakening asymmetry creates a wedge between the planner's halfer-derived social weights and the interim agent's thirder-derived weights.

A Background

A.1 Sleeping Beauty

> *Sleeping Beauty is in a windowless room on Sunday ($t = 0$) with a group of scientists, who explain to her that she will undergo the following experiment. The scientists shall toss a fair coin out of sight; if it lands heads, they shall put Beauty to sleep and awaken her on Monday ($t = 1$), before putting her to sleep again shortly afterwards for the rest of the experiment. If the coin lands tails, however, they shall also awaken her a second time on Tuesday ($t = 2$), before putting her to sleep for the rest of the experiment. Each time they put her to sleep they do so with a memory-wiping drug such that each awakening is indistinguishable to her. Upon every awakening they ask Beauty with what probability she believes that the coin landed heads. What should she say?* (Elga, 2000)

The two main schools of thought are *halfer* and *thirder*. Halfers, e.g. Lewis (2001), argue that upon awakening Sleeping Beauty has learnt no new information (she was certain to wake up whatever the state of the world) and so her credence in Heads should remain 1/2. Elga and his fellow thirders, however, argue that since each awakening is indistinguishable, she should assign equal probability to each of the three possible awakenings: {(Monday, Heads), (Monday, Tails), (Tuesday,Tails)}. Hence $\mathbb{P}(\theta = H) = 1/3$.

By referring to them as simply 'halfer' and 'thirder' positions, one risks giving the impression of two monolithic tribes which brook no internal dissent. Conversely, a dizzying array of varying arguments (often varying very subtly) have been offered in defence of each position, and the literature is enormous. Winkler (2017) offers a great review for those interested in investigating it. I myself am a thirder, in both this and the duplicating variant of the problem I discuss in the introduction (where in the event that the coin lands tails, we clone Sleeping Beauty and wake the clone up instead).

A.2 Newcomb's Problem

> *You walk into a room. Upon doing so, you are greeted by a friendly, super-intelligent robot, whose honesty you consider beyond dispute. He explains to you that you have a choice, and gestures towards two boxes on the desk. One of them is transparent, and you can see it contains \$1000. The other contains either a gazillion trillion dollars, or nothing at all. Your choice is between either taking only the mystery box (one-boxing), or both of them (two-boxing). The final crucial detail is that the mystery box contains the fortune if the robot predicted you would one-box, and nothing if he predicted you would two-box. He has an accuracy of near 100%. What do you do?* (Nozick, 1969)

Nozick (1969) introduced this problem and attributed it to Newcomb, famously noting that: "*To almost everyone, it is perfectly clear and obvious what should be done. The difficulty is that these people seem to divide almost evenly on the problem, with large numbers thinking that the opposing half is just being silly.*" One could say a similar thing of the debate over the Sleeping Beauty problem, though Nozick himself alas did not live to witness it.

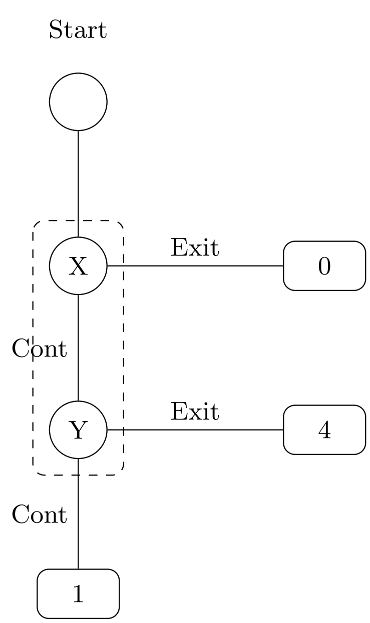
On the one hand, one could argue that the correct solution is clearly to two-box. The contents of the box are already fixed when one enters the room, so it is a strictly dominant strategy to take both boxes whatever the mystery box contains: only a muppet of the highest order would do otherwise. This position follows from *causal decision theory*, which is implicitly assumed in game theory, for example.

On the other, *evidential decision theory* asks: if you were to learn of your own decision after the fact, what would make you happier? To learn you had one-boxed or two-boxed? Since the robot is so accurate, whatever you did he likely predicted it. Hence, if you one-boxed you are likely rich and should be thrilled, but if you two-boxed you have covered this month's rent. Whilst this is not nothing, it is a real shame given the alternative. If you, like me, would two-box, you can nonetheless observe that the foolish one-boxes will fare much better. If before walking into the room you learnt of the scheme ahead of time, and could somehow commit yourself to one-boxing, it would clearly be a good idea to do so.

This is the phenomenon I call a *Newcomb tension*: an agent who would commit to one action, if given the chance, finds that their locally rational reasoning at the point of decision leads them to a different action. The value of commitment is positive, and the interim-rational agent cannot capture it.

A.3 The Absent-Minded Driver

The absent-minded driver (AMD), introduced by Piccione and Rubinstein (1997), is a single-player game with imperfect recall (specifically, *absent-mindedness* and not *forgetfullness*) in which the player cannot distinguish two decision nodes that lie on the same path of play. The game tree is as follows.



A driver, having foolishly decided to down several pints of lager before entering his vehicle, starts at intersection X and must choose between Exit and Continue. If he exits at X, the game ends with payoff 0. If he continues, he arrives at intersection Y, where he faces the same choice. Crucially, however, he cannot distinguish Y from X: in his drunken stupor they seem identical, and he both can never remember passing an intersection, and knows this fact upon arriving at any intersection (he is a logically omniscient drunk...). Formally, the intersections X and Y form a single information set h. If the driver exits at Y, the payoff is 4; if he continues past Y, the payoff is 1.

Since the driver cannot distinguish X from Y, any (pure) strategy must prescribe the same action at both intersections. Exit yields a game payoff of $U(\text{Exit}) = 0$ (the driver exits immediately at X), while Continue yields $U(\text{Continue}) = 1$ (the driver continues at both intersections). The planning-optimal pure action is therefore Continue.

However, the continuation payoffs are 'dot-dependent' (in the main body of the text I refer to different 'awakenings' or nodes within one branch of a play of the game as 'dots'): $v_X(\text{Exit}) = 0$, $v_Y(\text{Exit}) = 4$, $v_X(\text{Continue}) = 1$, $v_Y(\text{Continue}) = 1$. Under halfer beliefs $P(X) = P(Y) = 1/2$, the interim expected payoff from Exit is $\frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 4 = 2$, exceeding the expected payoff from Continue of 1. The fact that each pure strategy implies a belief such that the opposite strategy is optimal is the paradox of the absent-minded driver. Piccione and Rubinstein (1997) discuss different approaches to reasoning in this model but note that under the *modified multi-selves approach*, adopting a behavioural strategy of continuing with probability $2/3$ resolves the problem.

In addition to providing an interesting early example of what I call a *Newcomb tension*, this game also serves as an example of why sequential equilibrium may not exist in games of imperfect recall (Kreps and Wilson, 1982). In standard games, an equilibrium is defined in mixed strategies, which randomise which pure strategy is played at each information set, but do not independently randomise each time one arrives at that information set in the same play of the game. As per the famous Kuhn's Theorem (Kuhn, 1953), in games of perfect recall this distinction is unimportant.

B Omitted Proofs

B.1 Proof of Proposition 2.1

Proof. **Part (i).** When continuation payoffs are dot-independent, $v_d(a', \mathbf{a}_{-d}, \theta) = v(a', \theta)$ for all d and $\mathbf{a}_{-d}$, so the interim objective reduces to

$$V^{\text{OB}}(a') \;=\; \sum_{\theta,\,\mathbf{a},\,d} P^{\text{OB}}(\theta, \mathbf{a}, d)\, v(a', \theta) \;=\; \sum_{\theta} P^{\text{OB}}(\theta)\, v(a', \theta),$$

since the beliefs over action profiles and dots sum out. Taking $P^{\text{OB}}(\theta) = \rho(\theta)$ makes the interim and planning objectives identical.

Part (ii). In the absent-minded driver, there is a single state of the world. Under the pure strategy Continue, the action profile is $\mathbf{a} = (C, C)$ and $D(\theta, \mathbf{a}) = \{X, Y\}$. The continuation payoffs are $v_X(\text{Exit}, \mathbf{a}_{-X}, \theta) = 0$, $v_Y(\text{Exit}, \mathbf{a}_{-Y}, \theta) = 4$, $v_X(\text{Continue}, \mathbf{a}_{-X}, \theta) = 1$, $v_Y(\text{Continue}, \mathbf{a}_{-Y}, \theta) = 1$. The planning-optimal pure action is Continue ($U(\text{Continue}) = 1 > 0 = U(\text{Exit})$). Under halfer beliefs $P(X) = P(Y) = \frac{1}{2}$:

$$V^{\text{OB}}(\text{Exit}) = \tfrac{1}{2} \cdot 0 + \tfrac{1}{2} \cdot 4 = 2 > 1 = \tfrac{1}{2} \cdot 1 + \tfrac{1}{2} \cdot 1 = V^{\text{OB}}(\text{Continue}),$$

so the halfer chooses Exit, which is not planning-optimal. For one-boxer beliefs, I require $P(X) \cdot 0 + (1 - P(X)) \cdot 4 \leq P(X) \cdot 1 + (1 - P(X)) \cdot 1$, yielding $P(X) \geq \frac{3}{4}$. □

B.2 Proof of Theorem 3.1

Proof. The argument proceeds in five steps.

Step 1: All interim agents form the same belief. Since the game has a single information set h, every temporal part conditions on the same event (being at h) and hence forms the same interim belief over state-dot pairs. Under behavioural strategy σ, this belief is $P(\theta, d) = \rho(\theta)\, \sigma_d(\theta, \sigma)/C(\sigma)$, where $\sigma_d(\theta, \sigma)$ is the reach probability of dot d in state θ and $C(\sigma) = \sum_{\theta',d'} \rho(\theta')\, \sigma_{d'}(\theta', \sigma) > 0$.

Step 2: Existence of an interim best response. Fix a behavioural strategy σ used by all other temporal parts. The interim agent at a random dot chooses an action $a \in A(h)$ to maximise the expected continuation payoff $\sum_{\theta,d} P(\theta, d)\, v_d(a, \sigma_{-d}, \theta)$. Since $A(h)$ is finite, a maximiser exists. Moreover, since this is a single-agent game, all temporal parts share the same utility over terminal nodes, so every temporal part agrees on which action (or mixture over actions) is optimal given σ. This defines a best-response correspondence $\mathrm{BR} : \Delta(A(h)) \rightrightarrows \Delta(A(h))$.

Step 3: Existence of an interim equilibrium. The set $\Delta(A(h))$ of behavioural strategies is compact and convex. The best-response correspondence BR is upper hemicontinuous (since the continuation payoffs are continuous in σ) and convex-valued (since the agent's expected payoff is linear in their own mixing probabilities). By Kakutani's fixed-point theorem, there exists a behavioural strategy $\hat{\sigma}$ with $\hat{\sigma} \in \mathrm{BR}(\hat{\sigma})$: when all temporal parts play $\hat{\sigma}$, each finds $\hat{\sigma}$ optimal.

Step 4: Existence of a planning optimum. The planning objective,

$$V^{\text{plan}}(\sigma) \;=\; \sum_{\theta} \rho(\theta) \sum_{\mathbf{a}} \Pr_{\sigma}(\mathbf{a})\, U(\mathbf{a}, \theta),$$

is a continuous function on the compact set $\Delta(A(h))$, so a maximiser σ^* exists.

Step 5: The planning optimum is interim-optimal. Suppose for contradiction that σ^* is planning-optimal but not interim-optimal: there exists an action $a' \in A(h)$ such that the interim agent strictly prefers deviating to a'. That is,

$$\sum_{\theta,d} P(\theta, d)\, v_d(a', \sigma^*_{-d}, \theta) \;>\; \sum_{\theta,d} P(\theta, d)\, v_d(\sigma^*, \sigma^*_{-d}, \theta).$$

Consider the perturbed behavioural strategy $\sigma_\varepsilon = (1-\varepsilon)\sigma^* + \varepsilon\, \delta_{a'}$ for small $\varepsilon > 0$. The first-order change in the planning payoff is:

$$\frac{d}{d\varepsilon} V^{\text{plan}}(\sigma_\varepsilon)\Big|_{\varepsilon=0} \;=\; \sum_{\theta} \sum_{d} \rho(\theta)\, \sigma_d(\theta, \sigma^*) \big[v_d(a', \sigma^*_{-d}, \theta) - v_d(\sigma^*, \sigma^*_{-d}, \theta)\big].$$

Since $P(\theta, d) = \rho(\theta)\, \sigma_d(\theta, \sigma^*)/C(\sigma^*)$ and the interim agent strictly prefers a', this equals

$$C(\sigma^*) \sum_{\theta,d} P(\theta, d) \big[v_d(a', \sigma^*_{-d}, \theta) - v_d(\sigma^*, \sigma^*_{-d}, \theta)\big] \;>\; 0.$$

But this contradicts σ^* being a planning optimum, since V^{plan} could be improved by shifting toward a'. □

B.3 Proof of Proposition 3.2

Proof. Suppose a is an interim fixed point: when all temporal parts play a, each finds a optimal. Under the pure strategy a, the dots visited in state θ are $D(\theta, a)$, the normalising constant is $C(a) = \sum_\theta \rho(\theta)\, |D(\theta, a)| > 0$, and the thirder beliefs are $P(\theta, d) = \rho(\theta)/C(a)$ for each $d \in D(\theta, a)$.

At the pure strategy a, every dot is reached with certainty, so the reach probability $\sigma_d(\theta, a) = 1$ for all $d \in D(\theta, a)$. For any alternative action a', consider the perturbed behavioural strategy $\sigma_\varepsilon = (1-\varepsilon)\,\delta_a + \varepsilon\,\delta_{a'}$. Following the same calculation as Step 5 of the proof of Theorem 3.1, the first-order change in the planning payoff is

$$\frac{d}{d\varepsilon} V^{\text{plan}}(\sigma_\varepsilon)\bigg|_{\varepsilon=0} = \sum_\theta \sum_{d \in D(\theta,a)} \rho(\theta)\big[v_d(a', a_{-d}, \theta) - v_d(a, a_{-d}, \theta)\big].$$

Since $P(\theta, d) = \rho(\theta)/C(a)$, this equals

$$C(a) \sum_\theta \sum_{d \in D(\theta,a)} P(\theta, d)\big[v_d(a', a_{-d}, \theta) - v_d(a, a_{-d}, \theta)\big] \;\leq\; 0,$$

where the inequality follows from a being interim-optimal: the interim agent weakly prefers a to a' at every dot. Since this holds for every $a' \in A(h)$, the planning payoff cannot be increased by any local perturbation away from a, so a is planning-optimal among pure strategies. □

B.4 Proof of Theorem 4.1

Proof. For part (i), suppose $|D(\theta, n, \mathbf{a})| = |D(\theta, n', \mathbf{a})|$ for all n, n', θ, and $\mathbf{a}$. Since every agent visits the same number of dots, each agent's share is $|D(\theta, n, \mathbf{a})|/|D(\theta, \mathbf{a})| = 1/|N|$ in every $(\theta, \mathbf{a})$. The social weights are therefore $\lambda_n = 1/|N|$ for all n (independent of σ), and every agent exists in every state, so $E_n = \sum_\theta \rho(\theta)\, U_n(\sigma, \theta)$. The planning payoff reduces to $V^{\text{plan}} = (1/|N|) \sum_\theta \rho(\theta) \sum_n U_n(\sigma, \theta)$.

Fix any action $a' \in A(h)$ and consider the perturbation $\sigma_\varepsilon = (1-\varepsilon)\sigma + \varepsilon\,\delta_{a'}$. Write $h(\theta) = \bigcup_{\mathbf{a}} D(\theta, \mathbf{a})$ for the set of all dots reachable in state θ. Following the same calculation as Step 5 of the proof of Theorem 3.1, the first-order change in the planning payoff is

$$\frac{d}{d\varepsilon} V^{\text{plan}}(\sigma_\varepsilon)\bigg|_{\varepsilon=0} = \frac{1}{|N|} \sum_\theta \rho(\theta) \sum_n \sum_{d \in h(\theta)} \sigma_d(\theta, \sigma)\;\big[v_d^{(\theta,n)}(a', \sigma_{-d}, \theta) - v_d^{(\theta,n)}(\sigma, \sigma_{-d}, \theta)\big].$$

The thirder belief assigns $P(\theta, n, d) = \rho(\theta)\,\sigma_d(\theta, \sigma)/C(\sigma)$ to each dot, where

$$C(\sigma) \;=\; \sum_\theta \rho(\theta) \sum_n \sum_{d \in h(\theta)} \sigma_d(\theta, \sigma)$$

is the normalising constant. The interim first-order condition in the direction of a' is

$$\text{FOC}^{\text{int}}_{a'}(\sigma) \;=\; \sum_{\theta,n,d} P(\theta, n, d)\;\big[v_d^{(\theta,n)}(a', \sigma_{-d}, \theta) - v_d^{(\theta,n)}(\sigma, \sigma_{-d}, \theta)\big].$$

Substituting $P(\theta, n, d) = \rho(\theta)\,\sigma_d(\theta, \sigma)/C(\sigma)$ shows that the planning derivative equals $(C(\sigma)/|N|)\cdot \text{FOC}^{\text{int}}_{a'}(\sigma)$. Since $C(\sigma)/|N| > 0$, the planning derivative is non-positive in every direction a' if and only if the interim FOC is, so the planning optimum and interim optimum coincide. The argument of Theorem 3.1 applies.

For part (ii), suppose $|D(\theta_0, n_0, \mathbf{a}_0)| \neq |D(\theta_0, n_1, \mathbf{a}_0)|$ for some $(\theta_0, \mathbf{a}_0)$ and agents n_0, n_1. Then the social weights $|D(\theta_0, n, \mathbf{a}_0)|/|D(\theta_0, \mathbf{a}_0)|$ differ across agents in the compound state $(\theta_0, \mathbf{a}_0)$. In the planning derivative $dV^{\text{plan}}/d\sigma$, the compound-state formula weights each agent's payoff by their dot share *within each action profile*; agents with more dots in a given compound state receive proportionally larger weight. In the interim FOC, the thirder adjusts the probability of each $(\theta, \mathbf{a})$ proportionally to $|D(\theta, \mathbf{a})|$, which systematically overweights compound states in which more dots are reached. Because the dot shares differ across agents under $(\theta_0, \mathbf{a}_0)$, the planning and thirder weightings of the agents' payoffs diverge. This gap is generically non-zero (i.e. for all but a measure-zero set of payoff parameters); the duplicating AMD (Proposition 4.3) provides a concrete example in which $\sigma^* > \hat{\sigma}$ for *all* positive payoff parameters. □

B.5 Proof of Proposition 4.2

Proof. Parts (i) and (ii) follow from the social welfare construction, and the values are largely explained in the proposition itself. I add full detail here, at the cost of being a little repetitive. In state H (probability 1/2), the original agent has one awakening and the clone has none; in state T (probability 1/2), each agent has one awakening. The planner's social weights are therefore $\lambda_{\text{orig}} = 1/2 \cdot 1 + 1/2 \cdot 1/2 = 3/4$ and $\lambda_{\text{clone}} = 1/2 \cdot 1/2 = 1/4$. The original exists in both states with equal prior probability, so $E_{\text{orig}}(x) = \frac{1}{2}\, S(x, H) + \frac{1}{2}\, S(x, T)$; the clone exists only in Tails, so $E_{\text{clone}}(x) = S(x, T)$. Hence

$$V^{\text{plan}}(x) = \tfrac{3}{4}\big[\tfrac{1}{2}\, S(x, H) + \tfrac{1}{2}\, S(x, T)\big] + \tfrac{1}{4}\, S(x, T) = \tfrac{3}{8}\, S(x, H) + \tfrac{5}{8}\, S(x, T).$$

The interim agent, using thirder beliefs, assigns probability 1/3 to each of the three awakening-events (Heads-Monday, Tails-Monday, Tails-Tuesday). The interim social weights are $\lambda^{\text{int}}_{\text{orig}} = 2/3$ and $\lambda^{\text{int}}_{\text{clone}} = 1/3$. Each agent's conditional expected score is unchanged, so

$$V^{\text{int}}(x) = \tfrac{2}{3}\big[\tfrac{1}{2}\, S(x, H) + \tfrac{1}{2}\, S(x, T)\big] + \tfrac{1}{3}\, S(x, T) = \tfrac{1}{3}\, S(x, H) + \tfrac{2}{3}\, S(x, T).$$

Since $3/8 \neq 1/3$, the planning and interim coefficients on $S(x, H)$ differ, and strict properness of S ensures the maximisers are distinct.

For part (iii), observe that each agent has at most one payoff-relevant awakening per state: there is no state in which a single agent visits the information set twice with both visits affecting payoffs. Randomisation is beneficial only when different visits to the same information set would ideally take different actions, or the action set is discrete and 'convexification'/hedging helps. Here, each agent's single payoff-relevant visit should take the same deterministic action, so mixing over reports is weakly dominated by the optimal pure report. The Newcomb tension therefore persists for any behavioural strategy. □

References

Aumann, R. J., Hart, S., and Perry, M. (1997). The absent-minded driver. *Games and Economic Behavior*, 20(1):102–116.

Harsanyi, J. C. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of political economy*, 63(4):309–321.

Janda, P. (2024). Doppelgänger changes the game. *Episteme*, 21(4):1182–1207.

Kierland, B. and Monton, B. (2005). Minimizing inaccuracy for self-locating beliefs. *Philosophy and Phenomenological Research*, 70(2):384–395.

Kreps, D. M. and Wilson, R. (1982). Sequential equilibria. *Econometrica: Journal of the Econometric Society*, pages 863–894.

Kuhn, H. W. (1953). Extensive games and the problem of information. In Kuhn, H. W. and Tucker, A. W., editors, *Contributions to the Theory of Games*, volume II of *Annals of Mathematics Studies*, pages 193–216. Princeton University Press, Princeton.

Lewis, D. (2001). Sleeping beauty: reply to elga. *Analysis*, 61(3):171–176.

Nozick, R. (1969). Newcomb's problem and two principles of choice. In Rescher, N., editor, *Essays in Honor of Carl G. Hempel*, pages 114–146. D. Reidel, Dordrecht.

Osborne, M. J. and Rubinstein, A. (1994). *A Course in Game Theory*. MIT Press, Cambridge, MA.

Piccione, M. and Rubinstein, A. (1997). On the interpretation of decision problems with imperfect recall. *Games and Economic Behavior*, 20(1):3–24.

Ross, J. (2010). Sleeping beauty, countable additivity, and rational dilemmas. *Philosophical Review*, 119(4):411–447.

Schwarz, W. (2015). Lost memories and useless coins: revisiting the absentminded driver. *Synthese*, 192(9):3011–3036.

Spohn, W. (2025). The puzzle of the absent-minded driver is not about absent-mindedness, but about indexical belief. *Homo Oeconomicus*, pages 1–13.

Winkler, P. (2017). The sleeping beauty controversy. *The American Mathematical Monthly*, 124(7):579–587.