Episteme (2023), 1–26
doi:10.1017/epi.2023.24

ARTICLE

Doppelgänger Changes the Game

Pavel Janda

Institute of Philosophy, Czech Academy of Sciences, Jilská 1, 110 00 Prague, Czech Republic
Email: janda@flu.cas.cz

(Received 21 September 2021; revised 30 January 2023; accepted 14 February 2023)

Abstract
Thirders sometimes feel compelled to give the same answer – a credence of 1/3 – to the original and the duplicating Sleeping Beauty problem, which leads to some unwanted consequences. I will argue that they do not have to feel compelled to give the same answer, because the original and the duplicating version of the Sleeping Beauty problem are different types of decision problems. If one accepts that it is rationally permissible to give different answers to different types of decision problems, both versions do not require the same solution.

Keywords: Sleeping Beauty; absentmindedness; duplication; imperfect recall

1. Introduction

Philosophers have formulated the duplicating (doppelgänger) Sleeping Beauty problem to challenge the thirder position, i.e., the position that defends a credence of 1/3 in Heads for the original Sleeping Beauty case; see Elga (2000) for the original problem and, for example, Arntzenius (2003), Kierland and Monton (2005), Bostrom (2007), and Titelbaum (2014) for a discussion about the doppelgänger problem and duplication. The challenge, roughly speaking, is that: "many of the people who give the 1/3 answer for the case of Sleeping Beauty feel compelled to give that answer for the case of Duplicating Beauty" (Kierland and Monton 2005: 392), but the 1/3 answer creates unwanted consequences in the duplicating scenario (e.g., see Kierland and Monton 2005: 392; Bostrom 2007: 63; Leitgeb 2010).

I want to suggest an answer to that challenge. I believe that this wrong compulsion originates in neglecting the underlying decision-theoretic structures of both problems. I will employ game-theoretic tools to investigate the underlying decision-theoretic structures of both decision problems; my reasons for choosing game theory are that the Sleeping Beauty problem has its predecessors in game theory, game theory has been used to analyse the Sleeping Beauty problem (and similar games), and the formulation of the original Sleeping Beauty problem was inspired by game theory.[1] I will

[1]One might argue that game theory does not apply well to the Sleeping Beauty problem, since the Sleeping Beauty problem concerns a single agent playing against Nature. So, one should instead talk

argue that there are important structural differences between the classic and duplicating Sleeping Beauty problem on the basis of which thirders do not have to feel compelled to give the same answer to the duplicating and the original Sleeping Beauty problem. Specifically, the original and the duplicating Sleeping Beauty scenarios are different types of decision problems. The classic Sleeping Beauty problem is a game with absentmindedness, but the duplicating case is a game with imperfect recall (as understood in standard game theory, see Osborne and Rubinstein 1994; Piccione and Rubinstein 1997a; Lambert *et al.* 2019).[2] One could argue that, for the duplicating Sleeping Beauty case, one needs to redefine the standard game-theoretic concepts (I will discuss this point extensively in section 4). But the need to redefine the fundamental concepts of game theory to model the duplicating case would only underline the difference between the original and duplicating scenario since standard game-theoretic tools have been used to model the classic Sleeping Beauty problem (e.g., see Halpern 2006). In either case, I will conclude, both Sleeping Beauty scenarios are different types of decision problems, and if one accepts that it is rationally permissible to give different answers to different types of decision problems, thirders do not need to feel compelled to give the same answer to both scenarios. The goal of this paper is to prove Proposition 1, and discuss its philosophical consequences.

about decision theory rather than game theory. I am happy to call the Sleeping Beauty problem a "decision problem", but I do not think that it should prevent one from analysing such decision situations with game-theoretic tools. In fact, one can find multiple examples of decision situations that concern a single player or a single agent playing against Nature analysed and solved using game-theoretic tools. For example, decision situations in Osborne and Rubinstein (1994: 203) have only one player and one of them has one player and a chance node (which is similar to the Sleeping Beauty problem), and Osborne and Rubinstein (1994: 203) call them "one-player extensive games". Similarly, look at game (i) in Bonanno (2018: 120), Example 5 in Piccione and Rubinstein (1997b: 13), or the forgetful passenger in Aumann *et al.* (1997b). The most famous one-player game that is relevant to the Sleeping Beauty case is, of course, the absentminded driver paradox and its modifications (see Piccione and Rubinstein 1997a; Rubinstein 1998: Ch. 4, for overviews). The formulation of the original Sleeping Beauty problem was even inspired by a game-theoretic discussion about the absentminded driver case (see footnote 1 in Elga 2000: 143). Moreover, game-theoretic tools have already been applied to the Sleeping Beauty problem in Halpern (2006) and inspired further philosophical discussion (see Meacham 2008: 247, fn. 4). So, I am also happy to call the Sleeping Beauty problem a "game" and will do it from now on.

[2]First, the terms "game with absentmindedness" and "game with imperfect recall" are abbreviations for a game with an absentminded agent and a game with an agent with imperfect recall, respectively. Also, the game-theoretic terminology such as "imperfect recall" or "absentmindedness" might sound misleading since it automatically evokes forgetfulness or memory loss. Those terms can indeed amount to the fact that the agent forgets about her actions performed in the past. For example, in game theory, the usual interpretation of absentmindedness is that the agent must have forgotten that she has already been to and played in some part of the game (e.g., see Piccione and Rubinstein 1997a; Rubinstein 1998), but nothing hinges on that specific interpretation. Imperfect recall or absentmindedness does not have to be caused only by forgetfulness or memory loss. One can interpret imperfect recall or absentmindedness more generally as the uncertainty of whether the agent has moved in the past without specifying a reason for that uncertainty (Bonanno 2018: 120). Sleeping Beauty has been interpreted as a forgetful agent (e.g., see Monton 2002; Dorr 2003; Schervish *et al.* 2004), but, in the duplicating story, Beauty does not suffer from memory loss at any point, see Titelbaum (2014: 219). That is, Beauty might correctly remember that she reported her credence on Sunday but still be uncertain whether she has moved in the past. So, in the duplicating case, imperfect recall does not amount to forgetfulness but has a different cause, see Titelbaum (2014: 219–20). What makes Beauty uncertain is not forgetfulness but knowing that her perfect duplicate would also recollect the play on Sunday. This, I think, makes the duplicating case closely related to two roads to Shangri La discussed in Arntzenius (2003) or Dr. Evil discussed in Elga (2004).

Proposition 1. *The original Sleeping Beauty problem is a game with absentmindedness, but the duplicating Sleeping Beauty problem is a game with imperfect recall.*

To clarify, it might happen that, in the end, the same answer will be optimal for both scenarios, but it should not be an automatic requirement if my result is correct. One might also worry that the distinction between a game with absentmindedness and a game with imperfect recall is immaterial and has no impact on the properties and solutions of decision situations or philosophical discussions. I will defend the meaningfulness of the drawn distinction and its philosophical relevancy at the end of the paper (in section 5) by discussing its effect on some of the approaches to the Sleeping Beauty problem and the duplicating scenario. Probably the most important part of that discussion is showing that – if my arguments hold and the duplicating Sleeping Beauty problem is a game with imperfect recall – then thirders avoid the unpleasant consequences of creating a large number of Beauty's duplicates.

The paper is structured as follows. In section 2, I will discuss the game-theoretic background. In section 3, I will show that the original Sleeping Beauty case is a game with absentmindedness. In section 4, I will show that the duplicating scenario is a game with imperfect recall and not with absentmindedness. In section 5, I will discuss the philosophical relevancy of the drawn difference between the two games.

2. Game-Theoretic Background

Let an extensive-form game Γ consist of the following components (see, for example, Bonanno 2018: 75–6, or Osborne and Rubinstein 1994: 89–90, for further details). First, a finite set of players $P = \{1, \ldots, k\}$. Secondly, a finite rooted directed tree whose:

1. root has no directed edges leading to it, and
2. every other node has exactly one directed edge leading to it – there is a unique path (a unique sequence of directed edges) leading from the root to any other node.

Further, there is a set $\mathcal{A}$ of actions and a function $f_{\mathcal{A}}$ that assigns one action to every directed edge. No two edges leading out of the same node are assigned the same action. Usually, agents take pragmatic actions such as turn left, bet on horse number 23, etc. In this paper, the agent's actions of interest will be reports of her epistemic states, specifically, credences (mainly one's credences in Heads when a fair coin is tossed).

A node with no directed edges out of it – no further action is taken by any agent $i \in P$ – is called a *terminal node*. A game ends at terminal nodes. Every other node is called a *decision node*. Let D be a set of all the decision nodes in Γ. The concept of one decision node being a *predecessor* of another node will play an important part, so let me define it.

Definition 1 (Predecessor). *A node x is a predecessor of a node y, if there is a sequence of directed edges from x to y.*

The path from x to y must be continuous (there are no jumps between nodes) for x to be a predecessor of y.

Importantly, e.g., see Bonanno (2018: 119), I will assume that each decision node from D belongs only to a single agent $i \in P$, i.e., only that agent i makes a decision/plays at that

https://doi.org/10.1017/epi.2023.24 Published online by Cambridge University Press

node. Each agent i is thus assigned her own exclusive set of decision nodes $\mathcal{D}_i \subseteq D$, where the equality holds if we have a one-player game even without any chance nodes. Such a situation is rare but possible, e.g., see Osborne and Rubinstein (1994: 203–4).

An *information set* $I_i \subseteq \mathcal{D}_i$ of player i is created by partitioning $\mathcal{D}_i$. So, I_i is a non-empty collection of the decision nodes of the player i. So, the player concerned (and no other) is making a decision at those nodes (Webb 2007: 93). A player can have one or multiple information sets in Γ. Note that a single node forms an information set. If an information set I_i has more than one node, the standard interpretation is as follows, e.g., see Bonanno (2018: 117–18). Upon reaching I_i, the player i knows that she has reached I_i, but she does not know which sequence of directed edges leading to I_i is the actual one. In other words, she does not know at which of those nodes in I_i she is currently making her decision, since, to her, nodes in I_i are qualitatively identical and indistinguishable.

Games with imperfect recall and absentmindedness are defined negatively (they violate requirements for games with perfect recall). So, let me first define the properties of games with *perfect recall*, which are as follows (Bonanno 2018: 119):[3]

Condition 1 The actions available at any two nodes in the same information set I_i must be the same.[4]

Condition 2 If x and y are two nodes in the same information set I_i, then it is not the case that one node is a predecessor (see Definition 1) of the other.

Condition 3 If node $x \in I_i$ is a predecessor of node $y \in I_i^*$ (thus, by condition (2), $I_i \neq I_i^*$), and a is the action assigned to the directed edge out of x in the sequence of edges leading from x to y, then for every node $t \in I_i^*$ there is a predecessor $w \in I_i$ such that the action assigned to the directed edge out of w in the sequence of edges leading from w to t is that same action a.[5]

Condition 1 is a technical requirement. Any game that violates Condition 3 is a game with *imperfect recall*, e.g., see Bonanno (2004: 240; 2018: 119). A special case of imperfect recall is *absentmindedness*. Those games violate Condition 2 and are called games

[3]Traditionally (e.g., see Kuhn 1953), in game theory, one can interpret perfect recall as the following condition: if player i takes action a in an information set I_i and later on has to move again, then at the later time, she remembers that she took action a in that earlier information set (Bonanno 2018: 119). In other words, perfect recall requires – as a necessary condition – that a player always remembers what she knew in the past and what actions she herself took in the past. But, "to remember" is only a necessary and not a sufficient condition for a game to be a game with perfect recall. Given the comments in footnote 2, an agent can have a perfect memory but still be uncertain whether she has moved in the past, i.e., play a game without perfect recall. A possible alternative to my definition of perfect recall is in Osborne and Rubinstein (1994: 203–4), but Osborne and Rubinstein use a concept of "experience" and give it a specific technical meaning. Using a concept of experience, especially in the philosophical context, might be more confusing than helpful. So, I avoid such terminology.

[4]Remember that nodes in a single information set I_i are indistinguishable for agent i. The first condition makes sure that an agent cannot distinguish between different decision nodes in a single information set. If decision nodes in a single information set had different possible actions, they would not be qualitatively identical to that agent since she could use different sets of available actions to differentiate between those nodes.

[5]Condition 3 secures that, after moving from I_i to I_i^*, no matter at which node in I_i^* agent i finds herself, she knows that she must have played the action a. In other words, Condition 3 requires that for any node y in I_i^*, there is a predecessor node x in I_i such that they are connected – directly or indirectly – via a directed edge leading from x that is assigned the action a.

https://doi.org/10.1017/epi.2023.24 Published online by Cambridge University Press

with absentmindedness, e.g., see Rubinstein (1998: 70) or Halpern and Pass (2016: 280). Roughly speaking, in games with absentmindedness, agent i can take action in one of her information sets I_i that will lead her back to that same information set I_i (e.g., see Halpern and Pass 2016: 280, or Aumann *et al.* 1997a, for an example with a diagram).

3. The Original Beauty Is Absentminded

I will now show that the original Sleeping Beauty problem is a game with absentmindedness. Let me start by citing Elga's original Sleeping Beauty problem to give a point of reference:[6] "Some researchers are going to put you to sleep. During the two days that your sleep will last, they will briefly wake you up either once or twice, depending on the toss of a fair coin (Heads: once; Tails: twice). After each waking, they will put you to back to sleep with a drug that makes you forget that waking. When you are first awakened, to what degree ought you believe that the outcome of the coin toss is Heads?" (Elga 2000: 143).

Let me identify players, decision nodes, assign those nodes to the players, and determine information sets. In the original scenario, there is one player, Beauty, and Nature. By assumption, Beauty reports her credence on Sunday before the experiment starts. Let that node be called S, which will be the root of our decision tree (see Figure 1). Then, Nature makes a chance move with a coin flip, with the probability p of Heads and the probability $1 - p$ of Tails. Let that node be called S_{toss} since I assume that the toss happens on Sunday. When Beauty wakes up, she knows she is going through one of the two scenarios. In the first scenario, the coin lands Heads, she wakes up on Monday and reports her credence in Heads. Let that node be called MH, i.e., Monday and Heads. Then, the game ends. In the second scenario, the coin lands Tails, and she wakes up and reports her credence twice. First, she wakes up and reports her credence on Monday. Let that node be called MT, i.e., Monday and Tails. She is then put back to sleep and wakes up on Tuesday to report her credence. Let that node be called TT, i.e., Tuesday and Tails. Then, the game ends. So, $D = \{S, S_{toss}, MH, MT, TT\}$ is the set of all the decision nodes (including the chance node) in the original Sleeping Beauty game. Clearly, Beauty's set of decision nodes is $\mathcal{D}_B = \{S, MH, MT, TT\}$ and Nature gets $\mathcal{D}_N = \{S_{toss}\}$.

Now I need to determine information sets; I will focus on Beauty. By assumption, Beauty can always tell when she is on Sunday before the experiment starts. That is, she can distinguish S from any other node in $\mathcal{D}_B$, so $\{S\}$ is one of her information sets. Upon awakening, Beauty does not know the outcome of the coin toss, and the awakenings at MH, MT, and TT are subjectively indistinguishable to her. So, $\{MH, MT, TT\}$ is Beauty's other information set. Notice that, upon awakening, Beauty is uncertain whether she has already moved (reported her credence) on Monday or not. Moreover, reporting her credence at MT leads her to TT which is a node in the same information set as MT. It should not be surprising that the original Sleeping Beauty problem is a game with absentmindedness.

[6]There is a Sunday-toss and a Monday-toss version of the Sleeping Beauty problem. In this paper, I will assume that the coin flip happens on Sunday; a toss on Monday should make (Elga 2000: 144–5) and does make no difference to the final result (while some details, e.g., the structure of the decision tree, about how one gets to the final result might differ in the Monday-toss scenario).

https://doi.org/10.1017/epi.2023.24 Published online by Cambridge University Press

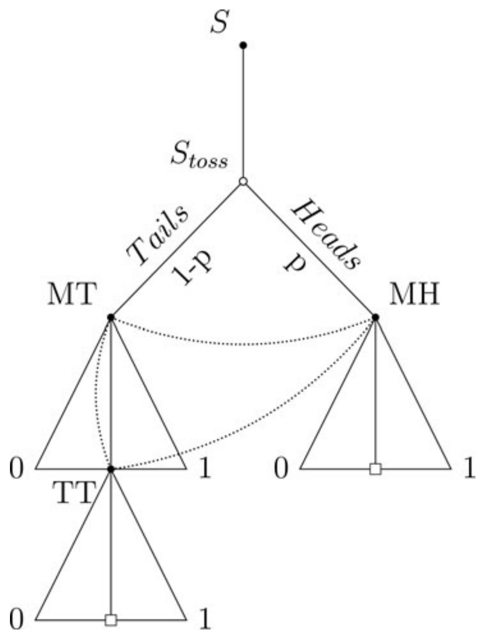

Fig. 1. I start with S, i.e., Sunday, when Beauty supposedly reports her credence in Heads. It is generally agreed that, on Sunday, she should report a credence of 1/2, so I use only one directed edge. Then comes Nature's chance node S_{toss} with the probability p of the coin landing Heads and the probability $1-p$ of the coin landing Tails. Beauty's decision nodes MH, MT, and TT form an information set $\{MH, MT, TT\}$, which is depicted by the dotted curves connecting the three nodes. The triangles at Beauty's decision nodes represent the idea that she can choose to report any probabilistic credence from the interval [0, 1] (see Myerson 1997: 140–8, section 3.13) for further details about similar decision problems). Reporting the extreme credence 0 or 1 is depicted by the edges of those triangles. The edges inside the triangles represent examples of possible non-extreme actions that Beauty may take; their current central position is only one of the possibilities and has no deeper meaning. Finally, the square nodes represent the terminal nodes – the game ends there.

Lemma 1. *The original Sleeping Beauty problem is a game with absentmindedness.*

One can visualise the original Sleeping Beauty problem with a decision tree in Figure 1.[7]

4. A Doppelgänger Changes the Game

Modifications of the original Sleeping Beauty problem have been used to argue against the thirder position; for example, see White (2006), Bostrom (2007), or Leitgeb (2010). I am interested in the cases in which a perfect duplicate of the original Beauty is created and woken up on Tuesday – instead of the original Beauty – if the coin lands Tails. A formulation and discussion about the duplicating case can be found, for example, in Kierland and Monton (2005: 391–3), Bostrom (2007: 62–5), Titelbaum (2014: 219), or Arntzenius (2003: 363–70). The following formulation comes from Builes (2020: 3038): "On Sunday, Beauty will be put to sleep. She will be woken up on Monday, and then let go. A coin will then be tossed on Monday night. If it lands Heads, nothing happens. If it lands Tails, a perfect subjective duplicate of Beauty, call her Tuesday Beauty, will be created, and this duplicate will be woken up on Tuesday morning. Tuesday Beauty will then be let go. Beauty is told that her Monday waking will be subjectively indistinguishable from the Tuesday waking of Tuesday Beauty."

[7]There is a precedence (see Halpern 2006) for modelling the Sleeping Beauty problem with a decision tree as an extensive-form game. My interpretation in Figure 1 is very similar to Example 5 in Piccione and Rubinstein (1997b: 13), which is the seminal game-theoretic paper that inspired the formulation of the original Sleeping Beauty problem.

https://doi.org/10.1017/epi.2023.24 Published online by Cambridge University Press

The main difference between the original and duplicating scenario is an addition of a new player – a perfect subjective duplicate of the original Beauty – who wakes up on Tuesday if the coin lands Tails. The question I want to explore is how to implement this change into the original game. My point will be that one should be careful about it because the chosen implementation has philosophical and game-theoretic implications.

First, at face value, the duplicating and the original scenario share many traits, for example, the following ones. A fair coin is tossed. There are two uncentred possibilities, Heads and Tails. The Heads-world has only one centre, *MH*. The Tails-world has two centres, *MT* and *TT*. Moreover, the awakening on Monday is still indistinguishable from the awakening on Tuesday, so *MH*, *MT*, and *TT* are subjectively indistinguishable upon awakening and thus belong to the same information set. So, one could be tempted to introduce the duplicate to the game by re-interpreting one of the centred-possibilities in Figure 1. Specifically, one assigns *TT* to the duplicate and keeps *MH* and *MT* for the original Beauty as in Figure 2. So, the information set in Figure 2 is $\{MH_{org}, MT_{org}, TT_{dup}\}$, where *org* stands for "original" and "*dup*" for "duplicate" indicating to whom those decision nodes belong. From the game-theoretic perspective, however, this is a problematic move because the game in Figure 2 is ill-formulated. By definition (see section 2), an information set is a collection of the decision nodes of a single player. One creates an information set by partitioning the set $\mathcal{D}_i$ of decision nodes of agent *i*, so an information set cannot contain decision nodes of more than one player. That is, set $\{MH_{org}, MT_{org}, TT_{dup}\}$ is not a well-formed information set, which leads to an ill-formulated game.

One could argue that the ill-formulated game in Figure 2 results from the inadequacy of the "classic" game-theoretic tools. So, for cases in which agents are unsure about who they are, as one of the reviewers suggested, one could aim at redefining

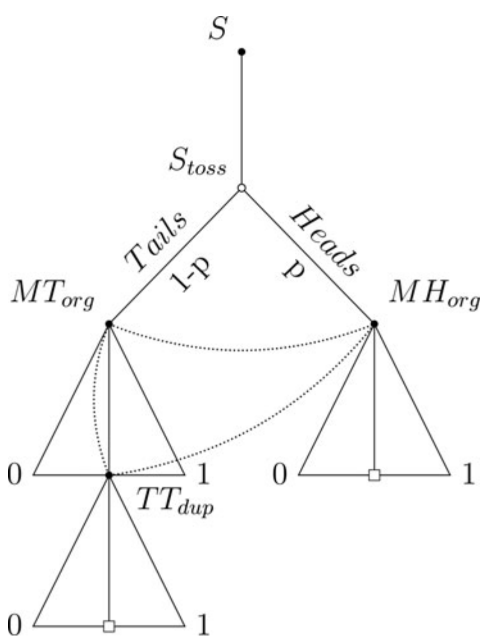

Fig. 2. A model for the duplicating Sleeping Beauty problem with decision nodes of multiple players belonging to a single information set.

https://doi.org/10.1017/epi.2023.24 Published online by Cambridge University Press

or extending the concept of an information set so that it can contain nodes of different players. I think that this is a possible but problematic position. Let me clarify why I think so. First, it slightly misrepresents my claim. I claim that if one uses a single information set containing decision nodes of multiple players (e.g., as in Figure 2), one cannot model the duplicating scenario using the classic game-theoretic tools such as standardly defined information sets. But I am *not* claiming that standard game theory with its "classic" tools cannot model the duplicating case at all. I argue that the standard game theory already has all the tools needed to model the original Sleeping Beauty problem (we have already seen that one in Figure 1) and the duplicating Sleeping Beauty problem (I will suggest my model shortly). But, for the sake of argument, suppose one wants to redefine the concept of an information set, so it can contain decision nodes of multiple players. In that case, one has to show that such a modification is not ad hoc to fit a specific case and why it is better than using the standardly defined concepts. Moreover, one has to show how such a modification connects with the rest of the game-theoretic tools and might need to modify other standard game-theoretic concepts (see footnote 9). For example, the definitions of a game with absentmindedness and a game with imperfect recall are formulated in terms of standard game-theoretic tools, including information sets. So, one would have to redefine those games and what their equilibria are. Even if one succeeds in all these modifications, the need for such modifications and the inability to model the duplicating Sleeping Beauty problem with the classic tools (while one can easily model the classic Sleeping Beauty problem with the standard tools as in Figure 1) still underlines my point about the structural difference between the classic and the duplicating scenario.

For the sake of argument, assume one overlooks all the required modifications (e.g., of the standard concept of an information set and possibly of other game-theoretic concepts) and decides to use the model in Figure 2. I still think that this decision tree is ill-fitted to model the duplicating scenario. By assumption, upon awakening, the original Beauty and her perfect duplicate cannot distinguish between possible awakenings at *MH*, *MT*, and *TT*. This means that, upon awakening, the original Beauty considers *MH*, *MT*, and *TT* to be possible but indistinguishable moments at which she makes a report (i.e., decision nodes among which she cannot discriminate but considers possible moments at which she makes a decision). Similarly, the duplicate considers *MH*, *MT*, and *TT* to be possible but indistinguishable moments at which she makes a report.[8] If one uses Figure 2, which has only one set of nodes *MH*, *MT*, and *TT*, then every node must have two players (the original Beauty and the duplicate) assigned to it (see Figure 3).[9] At the same time, one (including the original Beauty and the duplicate) also knows that who is awake on Monday can never be woken up and report her credence on Tuesday. That is, the same agent cannot make a decision at *MT* and *TT* (that the same agent reports on Monday and Tuesday cannot happen by the formulation of the duplicating case). But if one assigns *MT* to both agents (as in Figure 3), then *TT*

[8]Of course, the duplicate cannot be awake on Monday and the original Beauty on Tuesday. But one now models an agent's uncertainty about where she finds herself upon awakening. From that perspective, for example, Monday and Heads is a live possibility for the duplicate and Tuesday and Tails is a live possibility for the original Beauty.

[9]Moreover, one either needs to assign the directed edge from *MT* (which represents one's report at *MT*) to both agents or draw two lines from *MT* to *TT*, each belonging to one of the agents. Assigning a single node or directed edge to multiple players or drawing two edges from one node to another violates other classic game-theoretic concepts, e.g., see Bonanno (2018: 119). So, as hinted earlier, further modifications to the classic game-theoretic tools are needed.

https://doi.org/10.1017/epi.2023.24 Published online by Cambridge University Press

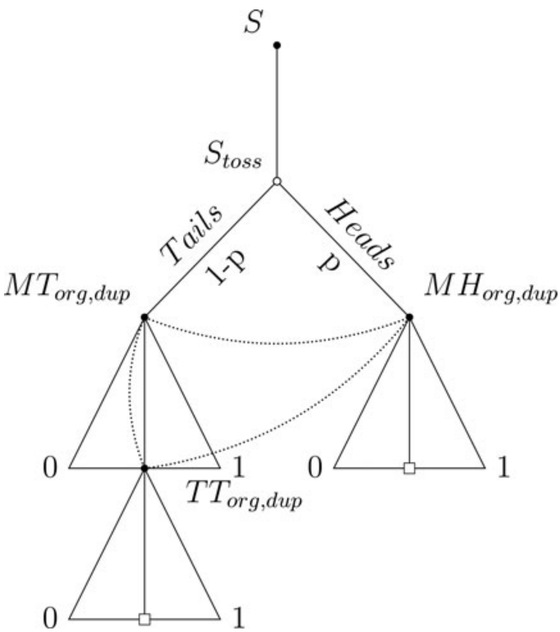

Fig. 3. Multiple agents assigned to the same nodes, where, e.g., $MT_{org,dup}$, means that *MT* belongs to the original Beauty and the duplicate at the same time.

cannot be assigned to either of those agents because at least one same agent would be assigned to *MT* and *TT*. And if, say, *MT* is assigned only to the original Beauty and *TT* only to the duplicate, then the original Beauty does not consider *TT* a possible moment for reporting her credence upon awakening. But the original Beauty should consider *TT* a possible moment where she makes a report since, by the formulation of the duplicating case, *MT*, *MH*, and *TT* are (from her perspective upon awakening) indistinguishable possible decision-making moments.

A conclusion from the previous discussion is that – if one wants to model the duplicating case with any type of game theory – one needs to do the following. Introduce a model (a decision tree) that allows both agents to be uncertain about who they are such that it allows them (upon awakening) to consider *MH*, *MT,* and *TT* to be a live possibility to each. Simultaneously, in that model, if a directed edge connects *MT* and *TT*, those nodes must belong to different players. Despite the agent's uncertainty about who she is, she knows that if she reports her credence on Monday, and if Tails, it will not be her who reports on Tuesday.

My suggested model (decision tree) for the duplicating Sleeping Beauty problem in Figure 4 uses only the classic game-theoretic tools; it is inspired by Halpern (2006) and Gilboa (1997). The model in Figure 4 uses two information sets: $\{MH_1, MT_1, TT_1\}$ that belongs to player 1 and $\{MH_2, MT_2, TT_2\}$ that belongs to player 2. Whether one interprets player 1 as the original Beauty and player 2 as the duplicate (or the other way around) is immaterial. But, for the sake of argument, assume that player 1 is the original Beauty, and player 2 is the duplicate. Consider $\{MH_1, MT_1, TT_1\}$ represented by the doted lines in Figure 4. Monday and Heads, Monday and Tails, and Tuesday and Tails are in the same information set belonging only to one player (original Beauty in this case), so they are subjectively indistinguishable for the original Beauty upon awakening. A similar case holds for the duplicate and $\{MH_2, MT_2, TT_2\}$ represented by the dashed lines in Figure 4. Importantly, in Figure 4, there is no path of directed edges

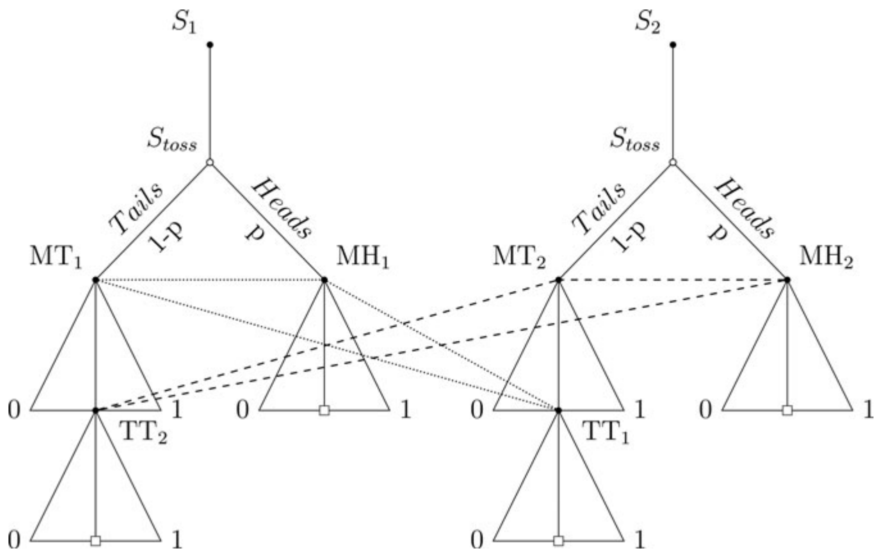

Fig. 4. Suggested model for the duplicating Sleeping Beauty problem, with the standard game-theoretic tools.

between the Monday and Tuesday awakenings belonging to the same information set. For example, the original Beauty (i.e., player 1) is, upon awakening, unsure whether she is at MT_1 or TT_1 (they belong to the same information set), but MT_1 and TT_1 are mutually exclusive possibilities since the directed edge leading from MT_1 is not a part of any sequence of directed edges leading to TT_1. In other words, waking up on Monday and Tuesday is mutually exclusive since the agent either wakes up in the right-hand tree or the left-hand tree, and there is no way of moving from one tree to the other. MT_1 and TT_2 are connected with a directed edge because one leads to the other, but those nodes belong to different agents (and different information sets) since different agents make a report on those occasions. Finally, there is no path of directed edges between S_1 (player 1 makes a report on Sunday) and TT_1 or S_2 (player 2 makes a report on Sunday) and TT_2. This means that if player 1 is the original Beauty (i.e., makes a report on Sunday), then there is no way that agent can wake up on Tuesday (i.e., make a report at TT_1), i.e., there is no path from S_1 to TT_1.

The game in Figure 4 is a game with imperfect recall but not a game with absentmindedness (in contrast to the original Sleeping Beauty case). Note that since there is no path of directed edges between MT_1 and TT_1, by Definition 1, MT_1 is not a predecessor of TT_1; the same applies to MT_2 and TT_2. So, if any MT and TT are in the same information set, there is no connected path of directed edges between that MT and TT. A consequence of this observation is that the duplicating scenario does not violate Condition 2, so it is not a game with absentmindedness.

Lemma 2. *The duplicating Sleeping Beauty scenario (as represented in Figure 4) is not a game with absentmindedness.*

Yet, the duplicating scenario (as represented in Figure 4) is a game with imperfect recall because it violates Condition 3.

https://doi.org/10.1017/epi.2023.24 Published online by Cambridge University Press

Lemma 3. *The duplicating Sleeping Beauty problem (as represented in Figure 4) is a game with imperfect recall.*

Taking together Lemma 1, Lemma 2, and Lemma 3 is enough to prove Proposition 1, as required.

5. Philosophical Relevancy

I have argued that if one wants to be able to model the duplicating scenario with the standard game-theoretic tools, then one must be careful about how Beauty's duplicate is introduced into the original Sleeping Beauty problem. I have argued for the model in Figure 4, and now I want to investigate some philosophical consequences of choosing that model. I will explore its relation to other existing philosophical approaches to the original and duplicating Sleeping Beauty problem.

5.1. The HT approach and the Elga approach

The first consequence is that some approaches that result in different recommendations for the classic case will make the same recommendation in the duplicating case. Specifically, consider the HT (Halpern and Tuttle) approach and the Elga approach from Halpern (2006). I chose to discuss the HT approach and the Elga approach because they are generalisations (see Halpern 2006: 125) of Elga's suggested solutions (see Elga 2000) to the original Sleeping Beauty problem but formulated in game-theoretic terms,[10] so easily applicable to my model. The idea is as follows (the comments in brackets are mine): "To summarize, the HT approach assigns probability among points [*nodes*] in an information set I by dividing the probability of a run r [*a run is what I have called a path of directed edges*] among the points in I that lie on r (and then normalizing so that the sum [*of the probabilities of nodes in I*] is one), while the Elga approach proceeds by giving each and every point in I that is on run r the same probability as that of r, and then normalizing" (Halpern 2006: 127).

The HT approach and the Elga approach differ in their recommendation for the original Sleeping Beauty problem. Consider Figure 1 with its information set {*MH*, *MT*, *TT*}. Two paths of directed edges, i.e., runs r, go through that information set. Let run r_2 start at S and go through MH and run r_1 start at S and go through MT and TT. So, only one node from {*MH*, *MT*, *TT*} lies on r_2 and two such nodes lie on r_1.

For example, assume that the probability of r_2 is the probability of a fair coin landing Heads, i.e., $p = 1/2$, and the probability of r_1 is the probability of a fair coin landing Tails, i.e., $1 - p = 1/2$. Given an information set I, the Elga approach gives every node from I that lies on a run r the same probability as that of r, and then renormalises probabilities over the nodes in I. So, MH gets the probability of r_2, i.e., 1/2, and both MT and TT get the probability of r_1, i.e., 1/2. So, every node in {*MH*, *MT*, *TT*} has the probability of $^1/_2$. After renormalisation (where $1/2 + 1/2 + 1/2 = 3/2$ is the normalising constant), every node in {*MH*, *MT*, *TT*} will have the probability of 1/3. Thus, the probability of Heads is 1/3 since the probability of MH is 1/3 (compare with Halpern 2006: 126). The Elga approach is a game-theoretic instance of Briggs' thirder rule, see (Briggs 2010: 10). One can translate Briggs' notation to game-theoretic concepts.

[10]From the game-theoretic perspective, Halpern's approaches bear similarity to the requirement of consistency of a belief system discussed in Piccione and Rubinstein (1997b: 12).

https://doi.org/10.1017/epi.2023.24 Published online by Cambridge University Press

Let A be an uncentred proposition (this remains unchanged), W uncentred worlds (runs r), N_W the number of centres in W (the number of decision nodes from a single information set lying on the given r), and $Cr_u(W)$ the probability of a run. For example, assume that A stands for Heads. Then, $\sum_{W \in A} Cr_u(W)N_W = 1/2$ since $Cr_u(W) = 1/2$ and there is one centre (i.e., node MH) in Heads (i.e., on r_2), so $N_W = 1$. Lastly, $\sum_W Cr_u(W)N_W$=3/2 is the normalising constant that sums over all the worlds (Briggs' $W^\star$ is notationally unnecessary, so I omit it). So, one's credence in Heads according to the thirder rule $\frac{\sum_{W \in A} Cr_u(W)N_W}{\sum_W Cr_u(W)N_W}$ is 1/3. For Tails, $\sum_{W \in A} Cr_u(W)N_W = 1$ since $Cr_u(W) = 1/2$ and there are two centres (i.e., MT and TT) in Tails (i.e., on r_1), so $N_W = 2$. Then, $\frac{\sum_{W \in A} Cr_u(W)N_W}{\sum_W Cr_u(W)N_W}$ is 2/3.

Keep assuming that the probability of r_1 and r_2 is 1/2. The HT approach assigns probability among nodes in an information set I by dividing the probability of a run r among the nodes in I that lie on that r. Since there is only one node, MH, on r_2, all the probability of 1/2 goes to that node. So, the probability of Heads is 1/2 (compare with Halpern 2006: 125–6). Since two nodes from {MH, MT, TT} lie on r_1, the probability of r_1 will be divided between those two nodes. In general, the HT approach does not give a rule on how to divide the probability among multiple nodes, but let me assume that 1/2 is divided equally between MT and TT (see (Halpern 2006: 126) for the same assumption). So, each gets 1/4. The HT approach is a game-theoretic instance of what Briggs calls the halfer rule, see Briggs (2010: 9). That is, one's credence in an uncentred proposition A equals to $Cr_u(A)$, which, in the game-theoretic jargon, is the probability of run(s) (i.e., uncentred worlds) where A is true. If A is Heads, then one's credence in A is the probability of r_2, i.e., 1/2. If A is Tails, then one's credence in A is the probability of r_1, i.e., 1/2.

Overall, the HT approach and the Elga approach differ in their recommendations for distributing probabilities among multiple nodes from a single information set that lie on the same run. The HT approach divides that probability between nodes, and the Elga approach gives each such node the full probability of the run. In the classic Sleeping Beauty scenario, both approaches differ in assigning probabilities between nodes MT and TT from the information set {MH, MT, TT} that lie on the run r_1. But for the duplicating scenario in Figure 4, there is no run with more than one node from a single information set. In Figure 4, there are four runs: r_1 starting at S_1 and going through MT_1 and TT_2, r_2 starting at S_1 and going through MH_1, r_3 starting at S_2 and going through MT_2 and TT_1, and r_4 starting at S_2 and going through MH_2. Each of these four runs has only a single decision node from any of the two information sets in Figure 4, i.e., {MH_1, MT_1, TT_1} and {MH_2, MT_2, TT_2}. Remember that the HT approach and the Elga approach differ in how they distribute probabilities over multiple nodes that lie on the same run and are from the same information set. But for any run with only one node from a single information set, both the HT approach and the Elga approach assign the whole probability of the run to that single node (e.g., see their assignment of probabilities to MH in the classic scenario). The Elga approach, by definition, assigns the whole probability of the run to any node on that run. The HT approach cannot divide the probability among multiple nodes because there is only one node, so that one node gets all the probability of the run. Since Figure 4 has no runs with multiple nodes from a single information set, the

HT approach and the Elga approach will give the same recommendation for the duplicating scenario in Figure 4; this observation is just a subcase of Lemma 3.1 in Halpern (2006: 128).

The concrete results will depend on the probabilities one assigns to the runs. For example, assume that the probability of each run is 1/2 since the probability of the coin landing Heads/Tails is 1/2. On Sunday, the agent can well locate herself. So, if, for example, she finds herself at S_1, she knows that the other tree cannot happen (i.e., r_3 cannot happen for player 1 since she cannot wake up as a duplicate if she is awake on Sunday). Player 1 then focuses only on the runs leading from S_1, where the probability of Heads (i.e., run r_2 when the coin lands Heads) and Tails (i.e., run r_1 when the coin lands Tails) is 1/2 by our assumption. So, e.g., by the Principal Principle, one can set her Sunday credences in Heads/Tails to 1/2. Upon awakening, one cannot locate herself in one of the trees and so must consider both of them, e.g., player 1 must take into account all the runs r_1, r_2, and r_3 with nodes from $\{MH_1, MT_1, TT_1\}$. Consider, for example, r_1 with two nodes MT_1 and TT_2; note that each node comes from a different information set. The Elga approach and also the HT approach gives the whole probability of r_1, i.e., 1/2, to MT_1 and TT_2 despite lying on the same run because MT_1 and TT_2 come from different information sets. So, every member of $\{MH_1, MT_1, TT_1\}$ and of $\{MH_2, MT_2, TT_2\}$ gets the probability of 1/2, which, after normalisation (which happens for each information set separately), gives the probability of 1/3 to every node. So, in our concrete case, one's credence in Heads will be 1/3 (since MH_1 and MH_2 get assigned the probability of 1/3); as we have seen, this is true for the Elga approach and also the HT approach.

The question is whether any instance of the thirder or halfer rule (for the original case) will give the same recommendation for Figure 4. It will not, see the next section. It all depends on one's starting assumptions. In this section, my only restriction was the game-theoretic requirement that the probabilities of directed edges leading from a single node sum to one (see footnote 14). One satisfies this requirement by assuming that the probability of Heads (i.e., the probability of r_2 and r_4) is 1/2 and the probability of Tails (i.e., the probability of r_1 and r_3) is 1/2. So, the sum of probabilities of r_1 and r_2 going from S_1 is one, and the sum of probabilities of r_3 and r_4 going from S_2 is one.

5.2. Compartmentalised conditionalisation

Meacham formulates a hypothetical-priors version of the HT approach and the Elga approach, compartmentalised and centred conditionalisation, where one uses hypothetical priors instead of probabilities of runs. Hypothetical priors are the credences one ought to have if one has no evidence whatsoever, see Meacham (2008: 248). Meacham assumes that one has hypothetical priors in centres (i.e., time-slices or centred worlds) and uncentred possible worlds. I will now discuss compartmentalised conditionalisation and leave centred conditionalisation for the next section.

In short, compartmentalised conditionalisation first divides one's credences among uncentred worlds and then divides the credence of each uncentred world among the alternatives (centres) at that uncentred world, see Meacham (2008: 257). One can operationalise it in the following three steps, see Meacham (2008: 249):

1. One takes her hypothetical priors and sets the credence in every centre (centred world) incompatible with her current evidence to 0.

https://doi.org/10.1017/epi.2023.24 Published online by Cambridge University Press

2. One normalises the credences in the remaining uncentred worlds such that the ratios between them are the same as the ratios between their hypothetical priors.
3. One normalises credences in the remaining centres at each world so that they sum to the credence assigned to that world, and the ratios between them are the same as the ratios between their hypothetical priors.

Meacham shows that compartmentalised conditionalisation assigns 1/2 in Heads for the classic and duplicating case. Assume that, on Sunday, by the Principal Principle (see Meacham 2008: 257) or other reason, one has a 1/2 credence in each, that the coin toss lands Heads/Tails. This means that one's hypothetical priors in Heads and Tails have a ratio of 1:1 (this follows from the properties of Bayesian conditionalisation (see Rédei and Gyenis 2021), which Meacham uses for his updating rules). On Sunday, Meacham says (2008: 257), the uncentred world Heads has a single centre (Sunday and Heads, i.e., *SH*), and the uncentred world Tails has a single centre (Sunday and Tails, i.e., *ST*). For the sake of convenience, also assume that the hypothetical priors in *MT* and *TT* are equal, i.e., their ratio is 1:1 (see Meacham 2008: 257, for the same assumption).

For the classic scenario in Figure 1, one's doxastic alternatives (centres) have changed upon awakening. One still has one alternative (Monday) at the Heads-world but two alternatives (Monday and Tuesday) at the Tails-world. None of those centred worlds is eliminated by one's evidence. So, the first step of compartmentalised conditionalisation has no effect. Similarly, upon awakening, despite the change in one's doxastic alternatives, uncentred worlds Heads/Tails from Sunday also remain (i.e., neither of them is eliminated by evidence). Since there is no change in uncentred worlds, and one knows, by assumption, that hypothetical priors in Heads and Tails have a 1:1 ratio, one will again have the credence of 1/2 in Heads/Tails, see Meacham (2008: 257). So, the second step is satisfied. The third step distributes credences in uncentred worlds to their centres according to one's hypothetical priors. By assumption, hypothetical priors in *MT* and *TT* are equal, i.e., their ratio is 1:1, so *MT* and *TT* each get 1/4 to preserve that ratio. The conclusion is that credence in Heads is 1/2. Given Meacham's treatment of duplication (2008: 251, 254–5),[11] in the duplicating scenario, the centre *TT* that originally belonged to Beauty will be replaced by the duplicate's centre (my interpretation is that something similar to the situation depicted in Figure 2 happens). Such change does not eliminate any of the uncentred worlds (i.e., Heads or Tails) at any stage of the game. The reasoning is then very similar to the classic case. Using the Principal Principle on Sunday again results in hypothetical priors for uncentred worlds, Heads and Tails, with a 1:1 ratio. Since, upon awakening, no centred or uncentred world gets eliminated, one's credence in Heads will again be 1/2 (and 1/4 goes again to *HT* and *MT*).

One can reproduce a similar reasoning about the duplicating scenario for Figure 4. Assume that on Sunday (S_1 or S_2), one can locate herself well and knows that the other tree cannot happen since there is no connection between both trees (there is no connected path to get from one to the other, see Figure 5), so she focuses only on runs going from the appropriate Sunday node. For example, focusing on S_1 and player 1, she knows that r_3 cannot happen (i.e., Tails can be true only if r_1 actualises), so she

[11]Meacham considers a case similar to the duplicating scenario with two worlds, *A* and *B*, with one centre in *A* and two centres in *B*, one of which belongs to the duplicate.

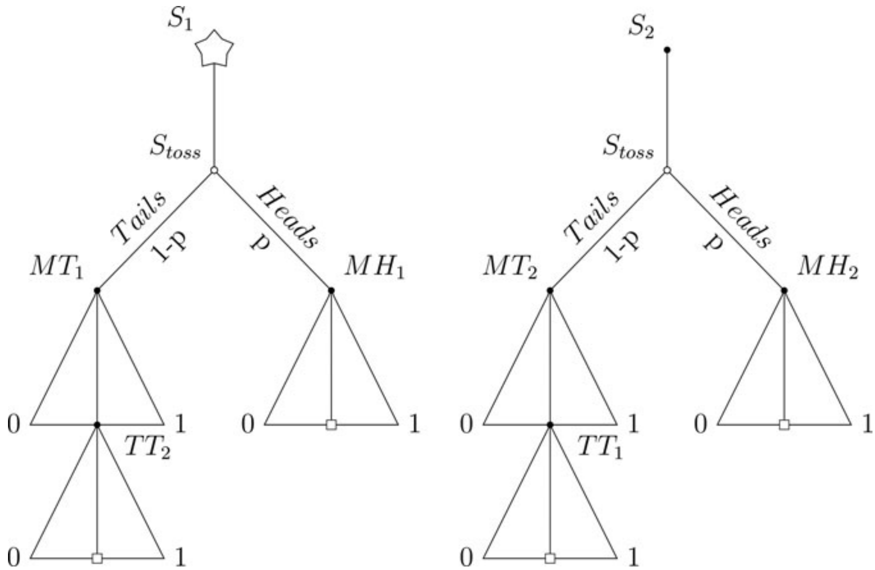

Fig. 5. Sunday perspective when a player can locate herself on Sunday, e.g., player 1 at S_1, indicated by a star. From that perspective, both trees are completely separated.

focuses only on r_1 and r_2. By assumption, the probability of Heads/Tails is 1/2, so one can use the Principal Principle to set her Sunday credences in Heads/Tails to 1/2. As before, it follows that hypothetical priors in Heads and Tails have a ratio of 1:1. Upon awakening, neither uncentred world Heads nor uncentred world Tails got eliminated. Since one's hypothetical priors in Heads/Tails are in a 1:1 relation (and none of these worlds gets eliminated), one's credence in Heads and Tails, upon awakening, is 1/2. If one again assumes that hypothetical priors between MT_1 and TT_1 (two of the new doxastic alternatives upon awakening) are in a 1:1 relation, then MT_1 and TT_1 are both assigned the credence of 1/4. So, in this case, the hypothetical priors between r_1, r_2, and r_3 would be 2:1:1, but this breaks none of our assumptions since the only assumptions about hypothetical priors were about the ratio 1:1 of Heads and Tails and the ratio 1:1 of MT_1 and TT_1, which all still hold.

The difference between my model in Figure 4 and Meacham is that, in my model, doxastic centres have not only changed (from Sunday to Monday nodes), but an agent's inability to locate herself in a world creates a new relation that connects two formerly separated trees (as in Figure 5) via an information set (as in Figure 4). This new relation represents an agent's inability to discriminate (upon awakening) between the elements of the information set that belongs to her. But since, for example, TT_1 is a part of this new relation for player 1 (i.e., is an element of the information set that belongs to player 1), r_3 that was eliminated on Sunday must be reintroduced on Monday. In my model, Tails is true if r_1 or r_3 (which are mutually exclusive) actualises, so credences (probabilities) in those runs (which one can see as uncentred worlds) must sum to 1/2 to meet the assumption that one's credence in Tails is 1/2. So, uncentred world Tails is a complex entity since it is a union or disjunction of two mutually exclusive uncentred worlds, represented by runs r_1 and r_3. In other words, if one thinks of an uncentred proposition as a set of uncentred worlds where that proposition is true,

https://doi.org/10.1017/epi.2023.24 Published online by Cambridge University Press

then the uncentred proposition Tails is a set of two uncentred worlds represented by runs r_1 and r_3. And these uncentred worlds r_1 and r_3 contain centres (i.e., decision nodes MT_1 and TT_1). So the difference (e.g., for player 1) between Sunday and Monday is not only about new doxastic alternatives but also about a new relation (represented by an information set) and consequently that r_3 (a run or an uncentred world), which was eliminated on Sunday (at S_1 for player 1), is a viable option on Monday. Despite all those additional changes, uncentred worlds Heads and Tails (the complex entity) are not eliminated, as a player moves from Sunday to Monday. I also standardly assume that Beauty knows how the game works from the beginning, so she is aware of these changes from the start.

5.3. Centred conditionalisation and multiple duplicates

To discuss centred conditionalisation, I need to introduce another mechanic of Meacham's approach, continuity. Roughly speaking, continuity tells one how to distribute hypothetical priors over centres in specific cases (i.e., when the centres are continuous). Thus, one can think of continuity as a replacement for the game-theoretic rule of how (in some cases) probabilities of runs distribute over nodes in a decision tree. Centred conditionalisation can be operationalised in the following three steps (see Meacham 2008: 249) with the continuity condition (see Meacham 2008: 252):

1. One takes her hypothetical priors in centres (centred worlds).
2. One sets the credence in every centre incompatible with her current evidence to 0.
3. One normalises the credences in the remaining centres such that the ratios between them are the same as the ratios between one's hypothetical priors. Credences in uncentred worlds are determined by sums of the credences in respective centres.

> **Continuity**: The ratio of priors between new alternatives is the same as the ratio of priors between any old alternatives they are continuous with.

Let me use Elga's conditions for continuity (as formulated by Meacham 2008: 258) for two centres. First, both centres are centred at the same world and individual. In other words, centres (nodes) must lie on the same run, i.e., a continuous path of directed edges, and be in information sets that belong to the same agent. I will also assume that if two nodes are continuous, one must be a direct predecessor of the other (chance nodes can be disregarded). This condition copies conditions (ii) and (iii) of Elga's continuity principle that the new alternative/centre (e.g., think of centres on Monday) is not centred at an earlier time than the old alternative (e.g., think of Sunday), and there is no other new alternative satisfying all the other required conditions for continuity that is centred at an earlier time than this new alternative.

First, consider the classic case (see Figure 1) and assume that, by the Principal Principle, one's credence in centres on Sunday, i.e., *SH* (Sunday and Heads) and *ST* (Sunday and Tails), is 1/2 in each. It means that one's hypothetical prior in *SH* and *ST* is also 1/2 in each (see Meacham 2008: 258). Upon awakening, centred conditionalisation itself does not tell one how to distribute credences among *MH*, *MT*, and *TT* because it will depend on one's hypothetical priors in *MH*, *MT*, and *TT*, which can be anything. Continuity becomes helpful here. Using the idea that *MH* and *MT* are continuous with *SH* and *ST*, respectively, one can deduce that hypothetical priors in *MH*

https://doi.org/10.1017/epi.2023.24 Published online by Cambridge University Press

and MT are equal (since, by continuity, they must preserve the ratio of hypothetical priors of SH and ST). But continuity does not tell one what actual value MH or MT has. To find that out, one can use Elga's principle of indifference (see Elga, 2000: 144; 2004), but I use a formulation from Meacham (2008: 258): one's credences in doxastic alternatives at the same world should be equal. In the game-theoretic terminology, one's credences in nodes on the same run and in the same information set (since one distributes credences over nodes in one's information set) should be equal. From Elga's principle of indifference, it follows that hypothetical priors in MT and TT are equal. Putting all together, hypothetical priors in MH, MT, and TT are all equal. For probabilistic functions, this holds only if one's hypothetical prior in each of MH, MT, and TT is 1/3. For the duplicating case, the reasoning is the same except that Elga's principle of indifference is now applied to doxastic possibilities MT centred at the original Beauty and TT centred at the duplicate (structurally, I imagine this scenario as the one in Figure 2).

Some philosophers (e.g., see Bostrom 2007: 63–5; Meacham 2008: 260; Builes 2020: 3040–2) claim that this approach becomes problematic if one starts multiplying Beauty's duplicates. One version among many could go as follows: "Let ϕ be a (non-indexical) proposition, to which Beauty assigns a prior credence of 1/2. Beauty is never woken up again after being put to sleep on Monday. If ϕ is true then there will be a total of $N > 0$ awakenings of doppelgangers in states that are subjectively indistinguishable from Beauty's Monday awakening" (Bostrom 2007: 64).

I believe that the currently discussed approach becomes problematic if one treats new duplicates as new centres in the same uncentred world (i.e., new nodes from a single information set on the same run). In other words, if one treats each new duplicate as a modification of Figure 2 and prolongs the Tail branch (run) by adding new nodes for each new duplicate (i.e., that duplicate's awakening), then the following problem occurs. For example, consider Figure 6 where there are two duplicates in total with nodes TT_1 and TT_2 belonging to duplicate 1 and duplicate 2, respectively (MH_{org} and MT_{org} belong to the original Beauty). Assume one uses Figure 6 and let Beauty's Sunday credences in Heads and Tails (one can think of them as SH and ST, respectively) be 1/2, so hypothetical priors in Heads and Tails are equal. Then, by continuity with Sunday, one's credences (upon awakening) in MH_{org} and MT_{org} are equal. Moreover, by Elga's principle of indifference, one's credences (upon awakening) in MT_{org}, TT_1, and TT_2 are also equal. So, upon awakening, one's credence in all of MH_{org}, MT_{org}, TT_1, and TT_2 are equal, which means that each node gets a credence of 1/4, if one has probabilistic credences. So, upon awakening, one's credence in Heads is 1/4 and in Tails 3/4. The more duplicates one adds – in the style of Figure 6 – the more the probability ratio will be in favour of Tails. For example, if one has 9 duplicates in total, one's credence (upon awakening) in Heads will be 1/10 and 9/10 in Tails.

First, let me consider how centred conditionalisation works with Figure 4 and then consider multiple duplicates. On Sunday, the agent knows that TT (which TT depends on the choice of S_1 or S_2) is impossible and concentrates only on the remaining tree. Given a fair coin, by the Principal Principle, let one's Sunday credence in Heads and Tails be 1/2. So, on Sunday, one's credence in Sunday and Heads is 1/2, and one's credence in Sunday and Tails is also 1/2. It follows that hypothetical priors in Sunday and Tails and Sunday and Heads are equal. In my model (Figure 4), there are no nodes Sunday and Heads (SH) or Sunday and Tails (ST) with which Monday nodes (MH_1 and MT_1 or MH_2 and MT_2) can be continuous. But one can use the Sunday node that (disregarding the chance node) is a direct predecessor of Monday nodes to

https://doi.org/10.1017/epi.2023.24 Published online by Cambridge University Press

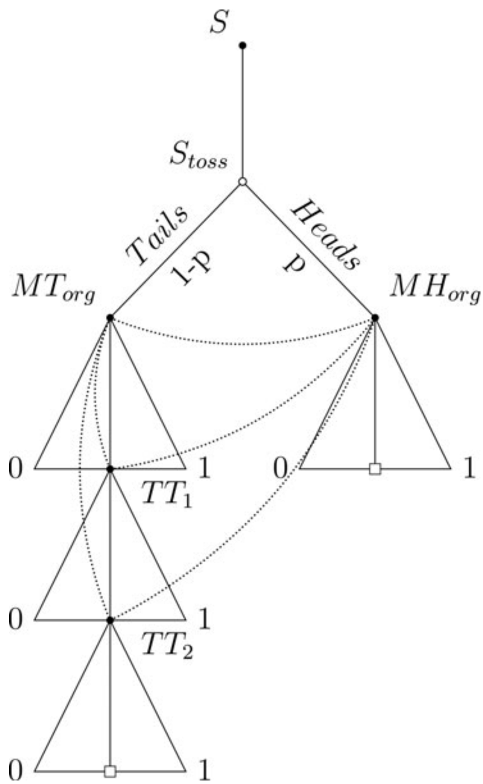

Fig. 6. Multiple duplicates on one run.

implement continuity (a direct predecessor means that there is a continuous path of directed edges, e.g., between S_1 and MH_1 and between S_1 and MT_1, and no other non-chancy nodes are between them). So, assuming that Monday nodes are continuous with respect to the appropriate Sunday node (i.e., MH_1 and MT_1 with S_1 and MH_2 and MT_2 with S_2), hypothetical priors in Monday nodes will be the same as hypothetical priors in Heads and Tails one had on Sunday. That is, hypothetical priors in Monday nodes are also equal, i.e., have a ratio of 1:1. The crucial difference is that, in Figure 4, one cannot apply Elga's principle of indifference since there is no uncentred world with more than one centre of a single player (i.e., a run with more than one node from a single information set). So, for example, the credences of player 1 will entirely depend on her hypothetical priors in MH_1, MT_1, and TT_1. But one only knows that those priors must be equal for MH_1 and MT_1, which leaves many options open. For example, one could go with 1/3 in each, so one's credence in Heads is 1/3, but this is only one of many options.

The non-applicability of Elga's principle of indifference is the reason why the model in Figure 4 can accommodate more than one duplicate without raising credence in Tails by introducing more and more duplicates. Let me again consider only two duplicates, but further extensions are based on the same idea.

In Figure 7, for example, the ratio of hypothetical priors in MH_1 and MT_1 that belong to player 1 (who acts at S_1 and whose information set is represented by the dotted curves) is equal by continuity with the ratio of hypothetical priors in Heads and

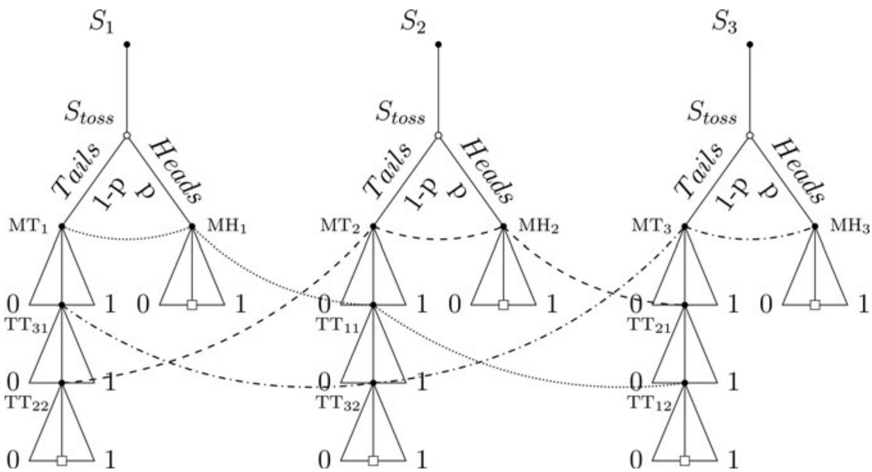

Fig. 7. A case with two duplicates, where player 1 starts at S_1, player 2 at S_2, and player 3 at S_3. For example, TT_{11} means that player 1 is the duplicate number 1 (where the number of the duplicate is given by the order in which they wake up) or TT_{32} means that player 3 is the duplicate number 2.

Tails at S_1 (i.e., S_1 is continuous with MH_1 and MT_1). But there is no restriction on hypothetical priors of player 1 assigned to TT_{11} in the second tree and TT_{12} in the third tree. First, because Elga's principle of indifference does not apply to them since they are the only nodes from the information set of player 1 on the given run. Secondly, those nodes are not continuous with any other node in their respective trees since all the other nodes are centred at different agents. So, for example, it is fine to hold a credence of 1/3 in MH_1 and MT_1 (the ratio of hypothetical priors in MH_1 and MT_1 is 1:1, like Heads and Tails at S_1) and have hypothetical priors of $1/3n$ in all the remaining nodes that belong to player 1, where n is the number of the remaining nodes (i.e., the number of duplicates). So, credences of player 1 in all nodes except MH_1 and MT_1 sum to 1/3. In our case, for two duplicates, it means that player 1 has a credence of 1/6 in both TT_{11} and TT_{12}. In other words, adding a second duplicate did not change the credence of player 1 in MH_1 (it is still 1/3), so her credence in Heads remains unchanged. And her credence in Tails is still 2/3. Similarly, adding more duplicates and accommodating them in the style of Figure 7 will not change credence in Heads of player 1 (or any other player in the game). So, sliding towards a higher probability of Tails as more and more duplicates are added does not happen.

5.4. Accuracy-based approaches

Finally, consider Kierland and Monton's suggestion that Beauty could find her optimal credence in Heads by minimising expected inaccuracy. Kierland and Monton consider two epistemic goals or methods she could follow: minimising expected average or total inaccuracy. The difference is that: "When one calculates expected total inaccuracy, one sums the inaccuracy for each temporal part, while when one calculates expected average inaccuracy, one averages the inaccuracy for each temporal part" (Kierland and Monton 2005: 389).

https://doi.org/10.1017/epi.2023.24 Published online by Cambridge University Press

Let me first consider expected total inaccuracy with the original case from Figure 1, where Beauty has three decision nodes (i.e., centres, temporal parts, or time-slices), *MH*, *MT*, and *TT*. The expected total inaccuracy calculates the inaccuracy of Beauty's credence in Heads, $c(H)$, at each node (I follow Kierland and Monton and use the Brier score[12] to evaluate inaccuracy), weights each score by the probability of reaching the respective node, and sums the weighted scores. For Figure 1, one has: $P_{MH}(1-c(H))^2 + P_{MT}(0-c(H))^2 + P_{TT}(0-c(H))^2$. Kierland and Monton assume that $P_{MH} = P_{MT} = P_{TT} = 1/2$ (e.g., see Kierland and Monton 2005: 389–90), so one has $\frac{1}{2}(1-c(H))^2 + \frac{1}{2}[2(0-c(H))^2]$, which is minimised at $c(H) = 1/3$. I used this form of the expectation formula to show that using expected total inaccuracy for the classic case results in weighted scoring, where the score $(1-c(H))^2$ is weighted by 1 since there is one node (*MH*) with this inaccuracy score and the score $(0-c(H))^2$ is weighted by 2 because there are two nodes (*MT* and *TT*) with this inaccuracy score. The problem is that strictly proper scoring rules (e.g., the Brier score) weighted with unequal positive weights stop being strictly proper (see footnote 12), but strict propriety is one of the fundamental assumptions in accuracy-based arguments, e.g., see Pettigrew (2016: 66). So, unless one can find a strictly proper weighted scoring rule (with unequal positive weights), I do not consider expected total inaccuracy a suitable method for the classic case; although, Kierland and Monton say that they do not have a conclusive reason to prefer expected total or average inaccuracy minimisation for the classic case, see Kierland and Monton (2005: 390).

Kierland and Monton, however, say that one should refrain from using expected total inaccuracy minimisation for the duplicating case (see Kierland and Monton 2005: 393). Remember that expected total inaccuracy sums the inaccuracy score (weighted by a probability) for each node, but it does not say to whom those nodes should belong. This is fine for the classic case where all nodes belong to a single agent (i.e., Beauty), but it matters in the duplicating case. When Kierland and Monton minimise expected total inaccuracy for the duplicating case, they minimise it for the original Beauty and a possible duplicate of her together (see Kierland and Monton, 2005: 393). So, e.g., considering Figure 2, the expectation formula is given by $P_{MH_{org}}(1-c(H))^2 + P_{MT_{org}}(0-c(H))^2 + P_{TT_{dup}}(0-c(H))^2$. For $P_{MH_{org}} = P_{MT_{org}} = P_{TT_{dup}} = 1/2$, one gets the optimal credence of 1/3 in Heads. This expectation formula is very similar to the one from the previous paragraph for the classic case, but the current formula mixes inaccuracy scores of the original Beauty (i.e., the inaccuracy scores for MH_{org} and MT_{org}) and her duplicate (i.e., the inaccuracy score for TT_{dup}). Kierland and Monton say that scores of different agents should not be mixed, i.e., one should focus only on her own expected inaccuracy and should not care about the inaccuracies of other people. Since their approach to computing expected total inaccuracy for the duplicating case mixes inaccuracy scores of different agents, Kierland and Monton dismiss

[12]The Brier score is the squared Euclidean distance $(v_n(H) - c(H))^2$ between one's credence $c(H)$ and the ideal (vindicated) credence $v_n(H)$ in H at the given node n (i.e., $v_n(H) = 1$ if H is true at n and 0 if it is false), see Kierland and Monton (2005: 385) or Pettigrew (2016: 4) for more details. The Brier score is a strictly proper scoring rule (i.e., an inaccuracy measure of one's credence in a single proposition, see Pettigrew 2016: 36), which means that the expected inaccuracy $p(1-c(H))^2 + (1-p)(0-c(H))^2$ is minimised at $c(H) = p$, for $0 \leq p \leq 1$ (see Pettigrew 2016: 86). It is important to note that if one uses positive weights $w_1, w_2 \in \mathbb{R}$ such that $w_1 \neq w_2$ to weight the inaccuracy scores given by the Brier score, then the weighted expectation $pw_1(1-c(H))^2 + (1-p)w_2(0-c(H))^2$ is not minimised at $c(H) = p$. In other words, the weighted Brier score (for unequal positive weights) is not strictly proper; see Levinstein (2019) for a general discussion about the strict propriety and weighted scoring rules.

https://doi.org/10.1017/epi.2023.24 Published online by Cambridge University Press

expected total inaccuracy minimisation as an appropriate epistemic goal for the duplicating case (see Kierland and Monton 2005: 393).

Still considering the duplicating case, Builes (2020: 3038–40) computes expected total inaccuracy focusing on minimising one's own expected inaccuracy and concludes that Beauty ought to have a credence of 1/2 in Heads when she wakes up. The idea is that, regardless of whether the coin lands Heads or Tails, there always is only one epistemically possible centre (node) that is her (see Builes 2020: 3038–9, fn 9). If the coin lands Heads, Beauty's total inaccuracy will be $(1 - c(H))^2$ since there is only one node where Heads is true. If the coin lands Tails, the agent knows that she is either Beauty (and not the duplicate) or that she is the duplicate (and not the original Beauty). Those are two mutually exclusive options, each with the inaccuracy score of $(0 - c(H))^2$. So, only one of them is epistemically possible (or relevant), and one's total inaccuracy is $(0 - c(H))^2$. Thus, for a fair coin, Beauty should minimise $1/2(1 - c(H))^2 + 1/2(0 - c(H))^2$, which gives $c(H) = 1/2$ as optimal. In contrast, for the classic case, Builes argues (2020: 3036–7) that minimising expected total inaccuracy recommends the credence of 1/3 in Heads. Given Builes' approach, minimising expected total inaccuracy recommends different optimal credences in Heads for the classic and the duplicating case. Builes further argues (2020: 3039) that the difference in recommendation results from the change in personal identity facts from one scenario to the other. So, the recommendations that minimising expected total inaccuracy gives explicitly depend on the personal identity facts, which is inconsistent with Time-Slice Rationality. Thus, time-slicers should minimise expected average inaccuracy and not expected total inaccuracy (see Builes 2020: 3040).

One can reconstruct Builes' reasoning for Figure 2 and Figure 4. I expressed my worries about using Figure 2 to model the duplicating case,[13] so I will consider only Figure 4. Consider player 1 with $\{MH_1, MT_1, TT_1\}$; from now on, I will use player 1 as my main example when discussing Figure 4, but all my reasoning can be easily reapplied to player 2 and $\{MH_2, MT_2, TT_2\}$. If Heads, she will be at MH_1 with a score of $(1 - c(H))^2$ or not woken up at all. So, there is only one relevant node of her, and her total inaccuracy is $(1 - c(H))^2$. If Tails, player 1 will be either at MT_1 or TT_1 with a score of $(0 - c(H))^2$ at each (but she cannot be at both nodes since the trees represent mutually exclusive situations). So, there is again only one relevant node for player 1 if Tails and her total inaccuracy is $(0 - c(H))^2$. Since, regardless of whether the coin lands Heads or Tails, there is always only one relevant node for player 1, she should minimise $1/2(1 - c(H))^2 + 1/2(0 - c(H))^2$ if the coin is fair. If, for example, player 1 is the original Beauty, then her optimal credence in Heads is 1/2. But one could argue that all nodes in $\{MH_1, MT_1, TT_1\}$ belong to player 1, so if one uses inaccuracy scores from all those three nodes to compute expected total inaccuracy for player 1, then player 1 still focuses on her own inaccuracy. That is, there is no mixing of inaccuracy scores from nodes belonging to different agents since player 1 uses inaccuracy scores from nodes that belong only to her and considers no scores from nodes that belong to player 2. Expected total inaccuracy of player 1 is then given by $P_{MH_1}(1 - c(H))^2 + P_{MT_1}(0 - c(H))^2 + P_{TT_1}(0 - c(H))^2$, which is almost the same expectation formula as Kierland and Monton's formula for the duplicating

[13]A relevant point here for Figure 2 is that if one applies Builes' reasoning to the duplicate, one will need to assign Beauty's node MH_{org} also to the duplicate to compute the duplicate's expected total inaccuracy. Even the duplicate, upon awakening, considers the possibility that she is the original Beauty and the coin landed Heads. So, a single node will belong to two different agents, which, I argue, is an issue (see section 4).

https://doi.org/10.1017/epi.2023.24 Published online by Cambridge University Press

case. For $P_{MH_1} = P_{MT_1} = P_{TT_1} = 1/2$,[14] it still holds that the optimal credence in Heads is 1/3, and the weighted expectation is not strictly proper, i.e., is unsuitable for accuracy-based arguments. But the difference is that one uses nodes MH_1, MT_1, and TT_1 instead of MH_{org}, MT_{org}, and TT_{dup}, so all the nodes now belong to the same agent, i.e., player 1.

I have so far assumed that focusing on one's own expected inaccuracy means not mixing the inaccuracy scores of different players. That is, in Figure 4, one does not mix scores of player 1 from nodes in the set $\{MH_1, MT_1, TT_1\}$ with scores of player 2 from nodes in $\{MH_2, MT_2, TT_2\}$. But there is a stricter interpretation of how to focus on one's own expected inaccuracy. That is, no player should mix the inaccuracy scores from the nodes where she is the original Beauty with the inaccuracy score(s) from the node(s) where she is the duplicate. For player 1, this excludes inaccuracy scores from any node in $\{MH_2, MT_2, TT_2\}$ but also separates inaccuracy scores she gets at MH_1 and MT_1 (since this is where player 1 is the original Beauty) from the score at TT_1 (it is where she is the duplicate). Then, player 1 always computes expected total inaccuracy with respect to only one tree and ignores the other (since she is the original Beauty in the tree with MH_1 and MT_1 and the duplicate in the tree with TT_1). Given a fair coin, $c(H) = 1/2$ is optimal for the tree with MH_1 and MT_1 and $c(H) = 0$ is optimal for the tree with TT_1. These, however, are not helpful recommendations since they only say what to do if one is in a specific tree. But that is all one can do under the stricter interpretation. For the sake of argument, assume one tries to find a followable advice by taking the expected total inaccuracy for each tree and weights it by the probability that one finds herself in the given tree to create an overall expectation formula. One now mixes inaccuracy scores from MH_1 and MT_1 with the score from TT_1, which is prohibited in this interpretation of "one's own inaccuracy".

In general, I agree that one should focus on one's own expected inaccuracy. But, as I have shown, one can still take multiple approaches to compute expected total inaccuracy, and I do not have a definitive reason to prefer one over the other. However, the consensus (albeit for different reasons) between Builes, Kierland and Monton, and me is that one should not minimise expected total inaccuracy in the duplicating case; I have also expressed my strict-propriety-related reason why I think that it is a wrong approach to the classic case.

Let me now turn to the minimisation of expected average inaccuracy. The idea is that the sum of every achievable score is averaged by the number of nodes (i.e., time-slices, temporal parts, or centres) where that score can be achieved. For the classic scenario in Figure 1, one has two nodes with the result Tails (i.e., score of $(0 - c(H))^2$) and one node with the result Heads (i.e., score of $(1 - c(H))^2$). So, one divides the sum $(0 - c(H))^2 + (0 - c(H))^2$ by 2 and $(1 - c(H))^2$ by 1, which gives: $P_{MH}(1 - c(H))^2 + \frac{1}{2}[P_{MT}(0 - c(H))^2 + P_{TT}(0 - c(H))^2]$. One then multiplies the score for each achievable node by the probability of it being achieved. For $P_{MH} = P_{MT} = P_{TT} = 1/2$, one gets $\frac{1}{2}(1 - c(H))^2 + \frac{1}{2}(0 - c(H))^2$, which is minimised at $c(H) = 1/2$, given the strict propriety of the Brier score. In contrast to the expected total inaccuracy, averaging the scores ensures that one does not end up with a weighted

[14]As usual in game theory (e.g., see the idea behind behavioural strategies in Bonanno 2018: 229), I assume that probabilities distributed over edges leading from a single information set sum to 1 (note that a single decision node forms an information set, e.g., see Osborne and Rubinstein 1994: 202, Example 202.1; or a discussion about a perfect-information frame in Bonanno 2018: 120). But probabilities of edges leading from different information sets are not restricted in this way. So, P_{MH_1} and P_{MT_1} must sum to 1 but P_{MH_1}, P_{MT_1}, and P_{TT_1} do not have to.

https://doi.org/10.1017/epi.2023.24 Published online by Cambridge University Press

scoring rule (e.g., the weighted Brier score discussed in footnote 12) that is not strictly proper due to unequal positive weights.

In the duplicating case (consider Figure 4), expected average inaccuracy works the same as in the classic case, but one uses the decision nodes from the information set $\{MH_1, MT_1, TT_1\}$ or $\{MH_2, MT_2, TT_2\}$. That is, an agent averages over the inaccuracy scores she receives at nodes in her information set. For example, consider player 1 and $\{MH_1, MT_1, TT_1\}$. If $P_{MH_1} = P_{MT_1} = P_{TT_1} = 1/2$, then one has $P_{MH_1}(1 - c(H))^2 + \frac{1}{2}[P_{MT_1}(0 - c(H))^2 + P_{TT_1}(0 - c(H))^2] = 1/2(1 - c(H)^2) + 1/2(0 - c(H))^2$ that is minimised at $c(H) = 1/2$. Consider now the duplicating case with $n \in \mathbb{N}$ duplicates (a generalisation of Figure 7 for $n + 1$ agents), so a single agent considers $n + 1$ nodes where Tails is true. Let, for example, TT_{1n} be a node at which player 1 considers to be an nth duplicate. If $P_{MH_1} = P_{MT_1} = P_{TT_{11}} = \cdots = P_{TT_{1n}} = 1/2$ (see footnote 14), the expected average inaccuracy will always recommend the credence of 1/2 in Heads because any number of $(0 - c(H))^2$ (i.e., scores for nodes where Tails is true) will be averaged out. That is, for n duplicates, the expected average inaccuracy formula is given by $\frac{1}{2}(1 - c(H))^2 + \frac{1}{2}\frac{1}{n+1}[(n + 1)(0 - c(H))^2] = \frac{1}{2}(1 - c(H))^2 + \frac{1}{2}(0 - c(H))^2$, which is minimised at $c(H) = 1/2$.

One could ask why one should not differentiate, for example, between MH_1 and MT_1 on one side and TT_1 on the other side (as I discussed earlier for the expected total inaccuracy). I think this would go against the idea of expected average inaccuracy. Kierland and Monton defended the expected average inaccuracy for the duplicating scenario by saying that: "If her goal is to minimise her own expected inaccuracy, then she should minimise the expected average inaccuracy for her and the possible duplicate of her, since she does not know which of those possible individuals she is" (Kierland and Monton 2005: 393). The general idea is that averaging minimises the expected inaccuracy of one's current node (time-slice), so one always focuses on her own inaccuracy (compare with Kierland and Monton 2005: 390, or Builes 2020: 3039). That is, the reason (exploring different ways in which one can focus on one's own expected inaccuracy) why I considered MH_1 and MT_1 separately from TT_1 for the expected total inaccuracy does not apply here since expected average inaccuracy, by definition, focuses on the expected inaccuracy of one's current node (i.e., one's own expected inaccuracy). Moreover – as mentioned in the quote – when one minimises expected average inaccuracy, one should also consider the possibility that she is the duplicate. So, separating MH_1 and MT_1 (where player 1 is the original Beauty) from TT_1 (where player 1 is the duplicate) does not make sense for the expected average inaccuracy. But also note that, when computing expected average inaccuracy (e.g., in Figure 4), one still keeps separate inaccuracy scores from nodes belonging to different players. That is, player 1 computes expected average inaccuracy using only nodes from $\{MH_1, MT_1, TT_1\}$ and player 2 computes it using inaccuracy scores only from nodes in $\{MH_2, MT_2, TT_2\}$. Since each information set contains a TT node, despite considering nodes only from one's own information set, each player also considers the possibility that she is the duplicate.

Overall, I agree with Builes, Kierland, and Monton – and my models in Figure 1 and Figure 4 (i.e., its generalisations for n duplicates) can accommodate it – that (for a fair coin) the expected average inaccuracy recommends a credence of 1/2 in Heads for both scenarios. I also agree that the expected total inaccuracy is not an optimal accuracy-based approach to the duplicating case. Moreover, I agree with Builes that minimising expected total inaccuracy is not a suitable method for the classic scenario, but I think this holds not only for time-slicers. For reasons discussed in section 4, I, however,

https://doi.org/10.1017/epi.2023.24 Published online by Cambridge University Press

disagree that the classic and the duplicating scenario are structurally identical (at least in game-theoretic terms) or that *the only* (epistemically) relevant difference between those two scenarios concerns identity facts, e.g., compare with Builes (2020: 3038–9) or Kierland and Monton (2005: 392). I consider the change/difference in the personal identity facts (introducing a new player or players, i.e., a duplicate or duplicates) from one scenario to the other to be fundamental or the most relevant in the sense that it is the change/difference that leads to other – at least game-theoretic but still epistemically relevant – changes/differences between the classic and duplicating Sleeping Beauty scenario.

6. Conclusion

Thirders sometimes feel compelled to give the same answer to the duplicating and the original Sleeping Beauty problem. A worry is that the thirder position creates unwanted consequences in the duplicating case. In this paper, I have used game-theoretic tools to argue that the original and the duplicating Sleeping Beauty scenarios are different types of decision problems. The original case is a game with absentmindedness, but the duplicating scenario is not. If one accepts that it is or should be rationally permissible to give different answers to different types of decision problems, then thirders do not need to feel compelled to give the same answer to both scenarios. I also discuss the relevancy of my argument for some philosophical approaches to the classic and duplicating Sleeping Beauty problem.[15]

References

Arntzenius F. (2003). 'Some Problems for Conditionalization and Reflection.' *Journal of Philosophy* **100**(7), 356–70.

Aumann R.J., Hart S. and Perry M. (1997a). 'The Absent-minded Driver.' *Games and Economic Behavior* **20**(1), 117–20.

Aumann R.J., Hart S. and Perry M. (1997b). 'The Forgetful Passenger.' *Games and Economic Behavior* **20**(1), 102–16.

Bonanno G. (2004). 'Memory and Perfect Recall in Extensive Games.' *Games and Economic Behavior* **47**(2), 237–56.

Bonanno G. (2018). *Game Theory*, 2nd edition. CreateSpace Independent Publishing Platform.

Bostrom N. (2007). 'Sleeping Beauty and Self-location: A Hybrid Model.' *Synthese* **157**(1), 59–78.

Briggs R. (2010). 'Putting a Value on Beauty.' In T.S. Gendler and J. Hawthorne (eds), *Oxford Studies in Epistemology*, vol. 3, pp. 3–34. Oxford: Oxford University Press.

Builes D. (2020). 'Time-Slice Rationality and Self-Locating Belief.' *Philosophical Studies* **177**(10), 3033–49.

Dorr C. (2003). 'Sleeping Beauty: In Defence of Elga.' *Analysis* **62**(276), 292–6.

Elga A. (2000). 'Self-Locating Belief and the Sleeping Beauty Problem.' *Analysis* **60**(2), 143–7.

Elga A. (2004). 'Defeating Dr. Evil with Self-Locating Belief.' *Philosophy and Phenomenological Research* **69**(2), 383–96.

[15]The author declares that they have no conflict of interest.

I confirm that the work on this paper was supported by the Formal Epistemology – the Future Synthesis grant, in the framework of the Praemium Academicum programme of the Czech Academy of Sciences. These funding sources have no involvement in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

I am grateful to my colleagues from the Czech Academy of Sciences, the LoPSE group at the University of Gdańsk, and the University of Bristol for their comments on various stages of my paper. Thanks to anonymous referees for their very useful comments, to the editors, and to Jonáš Gray for linguistic advice.

Gilboa I. (1997). 'A Comment on the Absent-Minded Driver Paradox.' *Games and Economic Behavior* **20**(1), 25–30.

Halpern J.Y. (2006). 'Sleeping Beauty Reconsidered: Conditioning and Reflection in Asynchronous Systems.' In T.S. Gendler and J. Hawthorne (eds), *Oxford Studies in Epistemology*, vol. 1, pp. 111–42. Oxford: Oxford University Press.

Halpern J.Y. and Pass R. (2016). 'Sequential Equilibrium in Games of Imperfect Recall.' In C. Baral, J. Delgrande and F. Wolter (eds), *Principles of Knowledge Representation and Reasoning: Proceedings of the Fifteenth International Conference (KR 2016)*, pp. 278–87. Cambridge, MA: AAAI Press.

Kierland B. and Monton B. (2005). 'Minimizing Inaccuracy for Self-Locating Beliefs.' *Philosophy and Phenomenological Research* **70**(2), 384–95.

Kuhn H. (1953). 'Extensive Games and the Problem of Information.' In H.W. Kuhn and A.W. Tucker (eds), *Contributions to the Theory of Games II*, pp. 193–216. Princeton, NJ: Princeton University Press.

Lambert N., Marple A. and Shoham Y. (2019). 'On Equilibria in Games with Imperfect Recall.' *Games and Economic Behavior* **113**, 164–85.

Leitgeb H. (2010). 'Sleeping Beauty and Eternal Recurrence.' *Analysis* **70**(2), 203–5.

Levinstein B.A. (2019). 'An Objection of Varying Importance to Epistemic Utility Theory.' *Philosophical Studies* **176**(11), 2919–31.

Meacham J. (2008). 'Sleeping Beauty and the Dynamics of De Se Beliefs.' *Philosophical Studies* **138**(2), 709–38.

Monton B. (2002). 'Sleeping Beauty and the Forgetful Bayesian.' *Analysis* **62**(1), 47–53.

Myerson R.B. (1997). *Game Theory: Analysis of Conflict*, 1st edition. Cambridge, MA: Harvard University Press.

Osborne M.J. and Rubinstein A. (1994). *A Course in Game Theory*, 1st edition. Cambridge, MA: MIT Press.

Pettigrew R. (2016). *Accuracy and the Laws of Credence*, 1st edition. Oxford: Oxford University Press.

Piccione M. and Rubinstein A. (1997a). 'The Absent-Minded Driver's Paradox: Synthesis and Responses.' *Games and Economic Behavior* **20**(1), 121–30.

Piccione M. and Rubinstein A. (1997b). 'On the Interpretation of Decision Problems with Imperfect Recall.' *Games and Economic Behavior* **20**(1), 3–24.

Rédei M. and Gyenis Z. (2021). Having a Look at the Bayes Blind Spot. *Synthese* **198**(4), 3801–32.

Rubinstein A. (1998). *Modeling Bounded Rationality*, 1st edition. Cambridge, MA: MIT Press.

Schervish M.J., Seidenfeld T. and Kadane J. (2004). 'Stopping to Reflect.' *Journal of Philosophy* **101**(6), 315–22.

Titelbaum M.G. (2014). *Quitting Certainties: A Bayesian Framework Modeling Degrees of Belief*, 1st edn. Oxford: Oxford University Press.

Webb J.N. (2007). *Game Theory: Decisions, Interaction and Evolution*, 1st edition. Cham: Springer.

White R. (2006). 'The Generalized Sleeping Beauty Problem: A Challenge for Thirders.' *Analysis* **66**(290), 114–19.

Appendix A. Proofs for Section 3 (The Original Beauty Is Absentminded)

Proof. To prove that the original Sleeping Beauty problem is a game with absentmindedness, I need to show that it violates Condition 2. Notice that both decision nodes *MT* and *TT* are in the same information set {*MH*, *MT*, *TT*} that belongs to Beauty. Moreover, by assumption, if the coin lands Tails, reporting credence in Heads on Monday leads to an awakening on Tuesday (and another report of Beauty's credence in Heads). So, *MT* is a predecessor of *TT*, i.e., there is a directed edge leading from *MT* to *TT* (see Figure 1). This violates Condition 2, as required. □

Appendix B. Proofs for Section 4 (A Doppelgänger Changes the Game)

Proof. One needs to show that neither the original Beauty nor the duplicate violates Condition 2. I will focus on player 1 (one can interpret player 1 as the original Beauty or the duplicate). One can also use player 2 in the same way, and one can also expand the following reasoning for cases with multiple duplicates (see discussion in section 5). I will use proof by contradiction. For the *reductio*, assume that the duplicating scenario is a game with absentmindedness, so it violates Condition 2. By Condition 2, I need to find at least two nodes such that: 1.) one node is a predecessor of the other node, and 2.) those two nodes come from the same information set. In the duplicating case, $\mathcal{D}_1 = \{S_1, MH_1, MT_1, TT_1\}$ is the set of decision nodes belonging to player 1. By assumption, player 1 can distinguish Sunday, but, upon awakening, she cannot distinguish among MH_1, MT_1, and TT_1. So, let me partition $\mathcal{D}_1$ into two sets: $\{MH_1, MT_1, TT_1\}$ and $\{S_1\}$. Condition 2 can be violated by player 1 only in those two information sets. Let me then consider those two sets one by one.

1. The violation of Condition 2 requires at least two nodes in the same information set, and so I can exclude $\{S_1\}$ straight away.
2. Let me now consider $\{MH_1, MT_1, TT_1\}$. Clearly, MH_1, MT_1, and TT_1 come from the same information set, but there is no node in $\{MH_1, MT_1, TT_1\}$ that is a predecessor of another node from $\{MH_1, MT_1, TT_1\}$. First, no directed edge from the Tuesday awakening leads to any of the Monday awakenings. So, TT_1 cannot be a predecessor of either MH_1 or MT_1. Secondly, suppose that the agent reports her credence in Heads on Monday, so there is a directed edge from a Monday node. First, it cannot lead from MH_1 by the definition of the Sleeping Beauty problem (i.e., there is no Tuesday awakening if Heads). Secondly, that edge cannot lead to TT_1. By definition, an information set is a set of nodes at which an agent acts. If there were an edge leading from MT_1 to TT_1, then there would be a scenario where player 1 acts (reports her credence in Heads) on Monday and Tuesday. But it is impossible since the same agent cannot report her credence in Heads on Monday and Tuesday; those are mutually exclusive events for the same agent in the duplicating case. We have seen that none of the Monday nodes (MH_1 or MT_1) is a predecessor of the Tuesday node TT_1, i.e., there is no directed edge from either MH_1 or MT_1 to TT_1.

This completes all the options how Condition 2 could be violated by player 1 in the duplicating scenario. Since the same reasoning holds for player 2 and $\mathcal{D}_2 = \{S_2, MH_2, MT_2, TT_2\}$, the duplicating Sleeping Beauty scenario is not a game with absentmindedness. □

Proof. To show that the duplicating Sleeping Beauty problem is a game with imperfect recall, I need to show that Condition 3 is violated at least for one agent. Let me focus on player 1 (player 2 would work equally well), which I now specifically interpret as the original Beauty.

We know that the original Beauty reports her credence in Heads on Sunday (S_1) and can fully distinguish S_1 from any other point in the game. So, S_1 forms a singleton information set $\{S_1\}$. The original Beauty also wakes up on Monday for sure, and so Sunday is a predecessor of her Monday awakening; either MH_1 or MT_1, depending on the coin toss result. But, in the duplicating scenario, three nodes (MH_1, MT_1, and TT_1) are in Beauty's information set. So, upon awakening, when the original Beauty moves from Sunday to Monday, she moves from the information set $\{S_1\}$ to a new information set $\{MH_1, MT_1, TT_1\}$.

If Condition 3 holds, then, for every node in $\{MH_1, MT_1, TT_1\}$, there is a predecessor in $\{S_1\}$ such that the action assigned to the directed edge out of that predecessor in the sequence of edges leading to any node in $\{MH_1, MT_1, TT_1\}$ is the same action a. Clearly, a necessary condition for Condition 3 to hold is that there is a predecessor in $\{S_1\}$ for every node in $\{MH_1, MT_1, TT_1\}$. Since $\{S_1\}$ is a singleton set, only S_1 can be a predecessor of any node in $\{MH_1, MT_1, TT_1\}$. This means that there must be a path of directed edges leading from S_1 to TT_1. That is, however, impossible. By assumption, the original Beauty reports her credence in Heads on Sunday and the duplicate wakes up on Tuesday. So, there cannot be a directed edge (or a path of directed edges) that leads from S_1 to TT_1. In other words, S_1 is not a predecessor of TT_1, and so not every node in $\{MH_1, MT_1, TT_1\}$ has a predecessor in $\{S_1\}$. But that there is a predecessor in $\{S_1\}$ for every node in $\{MH_1, MT_1, TT_1\}$ is a necessary condition for Condition 3 to hold. Thus, Condition 3 is violated, and the duplicating Sleeping Beauty problem is a game with imperfect recall. □

Pavel Janda is a postdoctoral researcher at the Department of Analytic Philosophy, Institute of Philosophy of the Czech Academy of Sciences. He received a PhD in philosophy from the University of Bristol in 2018. He specialises in formal epistemology.

Cite this article: Janda P (2023). Doppelgänger Changes the Game. *Episteme* 1–26. https://doi.org/10.1017/epi.2023.24