The Computational Complexity of Single-Player Imperfect-Recall Games

Emanuel Tewolde[1] , **Caspar Oesterheld**[1] , **Vincent Conitzer**[1,2] and **Paul W. Goldberg**[2]

[1] Foundations of Cooperative AI Lab (FOCAL), Carnegie Mellon University
[2] University of Oxford

emanueltewolde@cmu.edu, oesterheld@cmu.edu, conitzer@cs.cmu.edu, paul.goldberg@cs.ox.ac.uk

Abstract

We study single-player extensive-form games with imperfect recall, such as the Sleeping Beauty problem or the Absentminded Driver game. For such games, two natural equilibrium concepts have been proposed as alternative solution concepts to ex-ante optimality. One equilibrium concept uses *generalized double halving* (GDH) as a belief system and *evidential decision theory* (EDT), and another one uses *generalized thirding* (GT) as a belief system and *causal decision theory* (CDT). Our findings relate those three solution concepts of a game to solution concepts of a polynomial maximization problem: global optima, optimal points with respect to subsets of variables and Karush–Kuhn–Tucker (KKT) points. Based on these correspondences, we are able to settle various complexity-theoretic questions on the computation of such strategies. For ex-ante optimality and (EDT,GDH)-equilibria, we obtain NP-hardness and inapproximability, and for (CDT,GT)-equilibria we obtain CLS-completeness results.

1 Introduction

Most formalisms that are used for reasoning under uncertainty and for decision making in AI – for example, HMMs, (dynamic) Bayesian networks, influence diagrams, MDPs, POMDPs, multiagent versions of these – assume what is known as *perfect recall*: the agent does not forget anything it knew before. This may seem to be a very natural assumption: in the design of AI agents, generally we have plenty of reliable memory available. Moreover, the property of perfect recall ensures various desirable properties in the context of extensive-form games, including polynomial-time solvability of two-player zero-sum games [Koller and Megiddo, 1992] (and hence, *a fortiori*, single-player games). Finally, even when modeling humans – as in, for example, behavioral game theory [Camerer, 2003] – in spite of our clearly imperfect memory, usually perfect-recall models are used. So why use models with imperfect recall in AI?

It turns out there are a number of reasons why imperfect recall is relevant for AI agents; moreover, in cases where it is relevant, it is clear what the agent will and will not remember – unlike in the case of human memory, which is harder to predict and consequently to model in standard representations of imperfect recall. Imperfect-recall games already appear in the AI literature in the context of solving very large games such as poker: one technique for solving such games is *abstraction* – i.e., reducing the game to a smaller, simplified one to solve instead – and this process can give rise to imperfect recall in the abstracted game [Waugh *et al.*, 2009; Lanctot *et al.*, 2012; Kroer and Sandholm, 2016]. But imperfect recall is also of interest for other reasons. First, we may deliberately choose to have our agents forget: for example, the agent may temporarily need access to data that is sensitive from a privacy perspective, and therefore best forgotten afterwards. Conitzer and Oesterheld [2023] give the example of an AI driver assistant that can take over whenever the human car driver makes a major error. When that happens, the AI needs to reason about how good the human driver is in general, about whom it is not allowed to store information. An AI agent could also take the form of a highly distributed system operating across many nodes, where not all the nodes have access to the same information; hence, it may act at one node without having access to information that it did have available when acting at another node. Relatedly, the same agent (in the sense of being based on the same source code) may be instantiated multiple times, for example by human users deploying it in multiple contexts. In such cases it can still be useful to consider this family of instantiations as a single agent, but again these instantiations will not all have access to the same information. Finally, again building on the previous case, an agent may be acting not only in the real world, but also in simulations; for example, it may be simulated by another agent that wants to ensure that another instantiation of the same agent will act in a trustworthy fashion in the real world later [Kovarik *et al.*, 2023]. In this case, the real-world instantiation of the agent will generally not have access to the information that the simulation had access to earlier.

Notably, we need to model this phenomenon as imperfect recall rather than merely as imperfect information. In single-agent *perfect-recall* imperfect-information games (say, a POMDP), there is never a (strict) reason to randomize, whereas in imperfect-recall games the agent might have to randomize in order to perform well overall; cf. the Absentminded Driver example [Piccione and Rubinstein, 1997] (see appendix). For example, suppose we deploy a content rec-

ommendation system to many people's phones, in an edge-computing sort of setup: We are not in constant communication with the phones, so the nodes of our system have to act independently each day before getting back in touch with us. Over the next day, we would like to experiment (in an optimal way) what kind of content to recommend. With a pure strategy, we would show all users the same content and learn very little from it. Instead, we would prefer to randomize the content shown on each phone, that is, use a mixed strategy. Therefore, this situation cannot be a perfect-recall game (even of imperfect information).

Being able to make decisions with imperfect recall also represents a technical frontier. Many existing techniques inherently rely on perfect recall. Solving two-player zero-sum games becomes NP-hard as soon as one player has imperfect recall [Koller and Megiddo, 1992]. Moreover, in these contexts, there remains controversy at the very foundations of how to do probabilistic reasoning and decision making. For example, the Sleeping Beauty problem [Elga, 2000] (see appendix) asks one to give the probability of a state of the world in an imperfect-recall setting; some (Thirders) believe that the correct answer is $1/3$, and others (Halvers) believe it is $1/2$, see Section 3.2. Only recently has a clear picture started to emerge regarding how each of these positions can be combined with a corresponding form of decision theory to make good decisions [Hitchcock, 2004; Draper and Pust, 2008; Briggs, 2010; Conitzer, 2015; Oesterheld and Conitzer, 2022]; here we build on that recent conceptual work to define and study several foundational computational problems.

In this paper, we use extensive-form games to represent settings with imperfect recall. Even though we are considering a single-agent setting, the extensive form is still especially natural to use to model imperfect-recall settings, specifically with the use of information sets. Indeed, as we will discuss, game-theoretic phenomena such as notions of equilibrium naturally come up in the presence of imperfect recall even when there is just a single agent. Intuitively that is because it is more challenging for that agent to coordinate its actions with those it takes at other times. Moreover, randomization is in general necessary. We consider behavior strategies, which map each information set to a probability distribution over actions. Based on recent literature, we study three distinct solution concepts: (1) *ex ante* optimality, where the behavior strategy is one that maximizes expected utility at the outset; (2) equilibria based on causal decision theory and generalized thirding, in which an agent would not want to change its action at any information set, under the assumption that at all other game tree nodes (including ones in the same information set) the agent would follow the original strategy; and (3) equilibria based on evidential decision theory and generalized double halving, in which an agent would not want to switch to a different distribution over actions at a given information set, assuming that the agent would also use the *new* distribution at other nodes in that information set (but would use the original strategy at all other information sets).

Section 2 and 3 define those solution concepts and cover previously known characterizations and hardness results. Section 4 presents our novel results: First we show that the equilibria based on causal decision theory and generalized thirding are exactly the Karush–Kuhn–Tucker points of a corresponding utility maximization problem. This makes gradient descent methods applicable to the computation of such an equilibrium, and, relatedly, we derive that problems of finding such an equilibrium – up to an inverse exponential precision – are complete for the class CLS (Continuous Local Search). Finally, we derive various NP-hardness results for maximizing over the set of equilibria and for finding an equilibrium based on evidential decision theory and generalized double halving. Naturally, all these complexity results also have implications for learning or dynamics that converge to these solutions.

2 Background for Imperfect-Recall Games

2.1 Single-Player Extensive-Form Games with Imperfect Recall

We first define single-player extensive-form games, allowing for imperfect recall. The concepts we use in doing so are standard; for more detail and background, see, e.g., Fudenberg and Tirole [1991], Nisan *et al.* [2007] and Piccione and Rubinstein [1997].

Definition 1. *A single-player extensive-form game with imperfect recall (denoted Γ), sometimes also called an extensive decision problem with imperfect recall, consists of:*

1. *A rooted tree, with nodes $\mathcal{H}$ and where the edges are labeled with* actions. *The game starts at the root node h_0 and finishes at a leaf node, also called* terminal *node. The terminal nodes in $\mathcal{H}$ will be denoted as $\mathcal{Z}$. The set A_h refers to the set of actions available at a nonterminal node $h \in \mathcal{H} \setminus \mathcal{Z}$.*

2. *A utility function $u : \mathcal{Z} \to \mathbb{R}$, where $u(z)$ represents the payoff that the player receives from finishing the game at terminal node z.*

3. *A partition $\mathcal{H} \setminus \mathcal{Z} = \mathcal{H}_* \sqcup \mathcal{H}_c$ of nonterminal nodes into a set of the player's* decision nodes $\mathcal{H}_*$ *and a set of* chance nodes $\mathcal{H}_c$. *This partition indicates whether the single player or exogenous stochasticity determines the action at any given node.*

4. *For each chance node $h \in \mathcal{H}_c$, a fixed distribution $\mathbb{P}_c(\cdot \mid h)$ over A_h according to which chance determines an action at h.*

5. *A partition $\mathcal{H}_* = \sqcup_{I \in \mathcal{I}_*} I$ of the player's decision nodes into* information sets *("info sets" for short). We require $A_h = A_{h'}$ for any nodes h, h' in the same info set I.*

Throughout this paper, we let $\ell := |\mathcal{I}_*|$ denote the number of info sets. For computational purposes, we assume that a game Γ is represented by its game tree structure of size $\Theta(|\mathcal{H}|)$ (which includes the info set partition), and by a binary encoding of its chance node probabilities and its utility payoffs. The last two shall take on rational values only.

Any node $h \in \mathcal{H}$ uniquely corresponds to a (node, action)-history $\text{hist}(h)$ from root h_0 to h in the game tree. Define functions d, ν, a such that the node history and action history from h_0 to h consist of the sequences $\big(\nu(h,0), \nu(h,1), \dots, \nu(h,d(h))\big)$ and $\big(a(h,0), a(h,1), \dots, a(h,d(h)-1)\big)$ respectively. In other
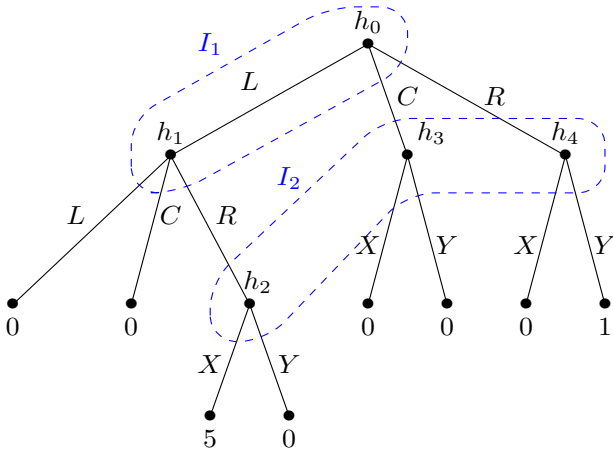
Figure 1: Running example of a single-player extensive-form game with imperfect recall. It has two info sets I_1 and I_2, and five nonterminal nodes.

words, function $d : \mathcal{H} \to \mathbb{N}_0$ identifies the tree depth of a node, $\nu : \mathcal{H} \times \mathbb{N}_0 \to \mathcal{H}$ the node ancestor at a specified depth, and $a : \mathcal{H} \times \mathbb{N}_0 \to \sqcup_{h \in \mathcal{H}} A_h$ the action ancestor at a specified depth. In particular, for all $h \in \mathcal{H}$, we have $\nu(h, 0) = h_0$ and $\nu(h, d(h)) = h$. We restrict the domain of functions ν and a to inputs (h, k) with $k \leq d(h)$ and $k \leq d(h) - 1$ respectively, and note that a maps (h, k) into $A_{\nu(h,k)}$.

The depth of Γ is defined to be the maximal depth of the leaf nodes. For notational convenience, we add a singleton info set to Γ for each chance node in Γ. The collection of these info sets, each consisting of a single element in $\mathcal{H}_c$, shall be denoted by $\mathcal{I}_c$. For each nonterminal node $h \in \mathcal{H} \setminus \mathcal{Z}$, let $I_h \in \mathcal{I}_* \sqcup \mathcal{I}_c$ denote its info set. For each info set $I \in \mathcal{I}_* \sqcup \mathcal{I}_c$, let A_I denote its action set.

Nodes of the same info set are assumed to be indistinguishable to the player during the game (even though the player is always aware of the full game structure). There may be information about the history of play that the player holds at some node, and that the player forgets somewhere further down its subtree. For instance, consider the game in Figure 1. Once the player arrives at node h_3, she cannot distinguish it from possibly being at node h_2. Thus she has already forgotten that she has only taken one action (action C) so far. In contrast to that, games with perfect recall have every info set reflect that the player remembers all her earlier actions. In particular, the player does not forget which info sets she entered in which order in the history of play.

With imperfect recall, it could furthermore be the case that multiple nodes of the same history (of some terminal node) belong to the same info set, as in info set I_1 in the game of Figure 1. The inability of a player to distinguish between two nodes on the same history is a property that we will refer to as *absentmindedness*; cf. the Absentminded Driver from Piccione and Rubinstein [1997] (see appendix).

Let $\Delta(A_I)$ denote the set of probability distributions over the actions in A_I. A (behavioural) *strategy* $\mu : \mathcal{I}_* \to \sqcup_{I \in \mathcal{I}_*} \Delta(A_I)$ of the player assigns to each info set I a probability distribution $\mu(\cdot \mid I) \in \Delta(A_I)$. At info set I, the player

will then randomly draw an action according to $\mu(\cdot \mid I)$. By abuse of notation, we extend any strategy μ of the player to info sets $I_h \in \mathcal{I}_c$ of chance nodes $h \in \mathcal{H}_c$ by setting $\mu(\cdot \mid I_h) := \mathbb{P}_c(\cdot \mid h)$ there.

Given that the player is currently at node $\bar{h} \in \mathcal{H} \setminus \mathcal{Z}$ and that she plays according to strategy μ, we can calculate the probability of reaching node $h \in \mathcal{H}$ by multiplying the probabilities of the actions on the path from $\bar{h}$ to h:

$$\mathbb{P}(h \mid \mu, \bar{h}) = \prod_{k=d(\bar{h})}^{d(h)-1} \mu\big(a(h, k) \mid I_{\nu(h,k)}\big) \quad \text{if } \bar{h} \in \text{hist}(h)$$

and $\mathbb{P}(h \mid \mu, \bar{h}) = 0$ otherwise. As a special case, we define the reach probability $\mathbb{P}(h \mid \mu) := \mathbb{P}(h \mid \mu, h_0)$ of a node $h \in \mathcal{H}$ to be its reach probability from the root h_0 of Γ. Naturally, the reach probability of the root is 1.

The expected utility payoff for being at node $h \in \mathcal{H} \setminus \mathcal{Z}$ and using strategy μ from then on can be determined by $U(\mu \mid h) := \sum_{z \in \mathcal{Z}} \mathbb{P}(z \mid \mu, h) \cdot u(z)$. Furthermore, let $U : \mu \mapsto U(\mu) := U(\mu \mid h_0)$ be the function that takes a strategy μ of Γ and returns the expected utility payoff of the player from following μ from the game start to termination. $U(\mu)$ is also called the *ex-ante* (expected) utility of μ.

2.2 Utility as a Polynomial Function

Fix an ordering $I_1, \ldots, I_\ell$ of the info sets in $\mathcal{I}_*$ and denote $m_i := |A_{I_i}|$ for all $i \in [\ell]$. Moreover, fix an ordering $a_1, \ldots, a_{m_i}$ of the actions in A_{I_i} for all $i \in [\ell]$.

We can uniquely describe a strategy μ of Γ by the probability values that it assigns to each action a_j at info set I_i, for $i \in [\ell]$ and $j \in [m_i]$. A strategy μ is a vector $\mu = (\mu_{ij})_{i,j} \in \times_{i=1}^\ell \mathbb{R}^{m_i}$ such that each subvector $\mu_{i\cdot} = (\mu_{ij})_j$ lies in the simplex $\Delta(A_{I_i}) \equiv \Delta^{m_i-1} := \{y \in \mathbb{R}^{m_i} : y_j \geq 0 \, \forall j, \sum_{j=1}^{m_i} y_j = 1\}$. Therefore, the strategy space of Γ is $\times_{i=1}^\ell \Delta(A_{I_i}) \equiv \times_{i=1}^\ell \Delta^{m_i-1}$.

The expected utility function U of a strategy μ can be fully written out as

$$U(\mu) = \sum_{z \in \mathcal{Z}} \left(u(z) \cdot \prod_{k=0}^{d(z)-1} \mu\big(a(z, k) \mid I_{\nu(z,k)}\big) \right).$$

As noted by Piccione and Rubinstein [1997], this is a polynomial function in the variables $(\mu_{ij})_{i,j}$. Recall that at chance nodes, the probabilities are exogenously fixed constants. Thus, the degree of the polynomial function U is upper-bounded by the maximum number of times the player of Γ might have to take a decision in order to reach a terminal node. Note that polynomial U can be constructed in polynomial time in the encoding size of Γ.

Example 2. *In the game of Figure 1, we get $\ell = 2$, $m_1 = 3$, $m_2 = 2$. Let the actions be ordered as (L, C, R) and (X, Y). Then, for any point $\mu \in \mathbb{R}^3 \times \mathbb{R}^2$, we have $U(\mu) = 5\mu_{11}\mu_{13}\mu_{21} + \mu_{13}\mu_{22}$.*

We show in the appendix that one can also reduce any multivariate polynomial $p : \times_{i=1}^\ell \mathbb{R}^{m_i} \to \mathbb{R}$ to a single-player extensive-form game Γ with imperfect recall such that its expected utility function satisfies $U(\mu) = p(\mu)$ on $\times_{i=1}^\ell \mathbb{R}^{m_i}$.

2.3 (Computing) Ex-ante Optimal Strategies

Suppose we want to *solve* a given game Γ. From a planning perspective, one would naturally search for a strategy that promises the highest payoff at a time before the player enters the game.

Definition 3. *We say a strategy μ^* is ex-ante optimal for Γ if it solves*

$$\max_{\mu} U(\mu) \quad \text{s.t.} \quad \mu \in \bigtimes_{i=1}^{\ell} \Delta(A_{I_i}) \,. \tag{1}$$

Due to Koller and Megiddo [1992], we can find an ex-ante optimal strategy for a single-player game with perfect recall in polynomial time. This will not be the case anymore in the presence of imperfect recall, as we will show next.

The class ZPP contains all those decision problems that can be solved in expected polynomial time by a randomized (Las Vegas) algorithm. Let OPT be an optimization problem with max and min values $\bar{q}$ and $\underline{q}$. Then, a fully polynomial-time approximation scheme (FPTAS) for OPT computes a solution to an instance of OPT with an objective value that is at most $\epsilon \cdot (\bar{q} - \underline{q})$ away from the optimal value. This computation must take polynomial time in $1/\epsilon$ and the encoding size of the instance. For a more precise definition, see de Klerk [2008].

Proposition 4. *Consider the problem that takes a game Γ and target value $t \in \mathbb{Q}$ (encoded in binary) as inputs and asks whether there is a strategy μ for Γ with ex-ante expected utility $U(\mu) \geq t$. This problem is NP-hard. Moreover:*
(1.) Unless NP = ZPP, there is no FPTAS for this problem. NP-hardness and conditional inapproximability hold even if the game instance Γ has a tree depth of 3 and only one info set.
(2.) NP-hardness holds even if Γ has no absentmindedness, a tree depth of 4 and the player has 2 actions per info set.
(3.) NP-hardness holds even if Γ has no absentmindedness, a tree depth of 3 and the player has 3 actions per info set.

Note that finding the ex-ante optimal strategy of Γ is at least as hard as this decision problem of whether a target value can be achieved in Γ. On the other hand, with an efficient solver of the decision version, one can recover the optimal ex-ante utility $U^* := \max_{\mu} U(\mu)$ through binary search.

An early such NP-hardness result was given by Koller and Megiddo [1992]. Gimbert *et al.* [2020] leverage the correspondence to polynomial optimization (Section 2.2) to further show that, in fact, this decision problem of whether a target value can be achieved in Γ is $\exists \mathbb{R}$-complete. The complexity class $\exists \mathbb{R}$, called *the existential theory of the reals* [Renegar, 1992; Schaefer and Stefankovic, 2017], consists of all those problems that reduce to deciding whether a sentence of the following form is true: $\exists x_1 \ldots \exists x_n F(x_1, \ldots x_n)$, where the x_i are real-valued variables and where F is a quantifier-free formula that may contain equalities and inequalities of real polynomials. $\exists \mathbb{R}$ lies in between NP and PSPACE [Shor, 1990; Canny, 1988]. The decision problem of Proposition 4 can also be shown to be a member of NP in an approximate sense; namely, when it is allowed to incorrectly return "yes"

to the problem instance (Γ, t, ϵ) if there exists a strategy profile μ with $U(\mu) \geq t - \epsilon$. Here, $\epsilon > 0$ represents an inverse-exponential precision parameter.

A proof of Proposition 4 can be found in the appendix. Result (1.) is based on the known hardness of maximizing a polynomial function over a simplex [de Klerk, 2008]. Result (3.) reduces from the hard problem of finding an optimal joint strategy in multiplayer common payoff games [Chu and Halpern, 2001]. The proofs reveal that NP-hardness remains even if the encoding size of the chance node probabilities and utility values are in $O(|\mathcal{H}|)$.

3 Equilibria in Imperfect-Recall Games

Proposition 4 shows a strong obstacle to finding or approximating ex-ante optimal strategies for single-player extensive-form games with imperfect recall. In light of these limitations, we will relax the space of solutions to *equilibrium* strategies. This solution concept argues that, whenever the player finds herself in an info set, she has no influence over which actions she chooses at other info sets. Therefore, at an equilibrium strategy μ, the player will play the best action at each info set, assuming that she has been playing according to μ up to the current decision point and that she will continue to do so at future decision points. Prior work has given a detailed description of viable equilibrium concepts in single-player games with imperfect recall [Piccione and Rubinstein, 1997; Briggs, 2010; Oesterheld and Conitzer, 2022]. We will consider two well-motivated equilibrium concepts that have been proposed and where an ex-ante optimal strategy also constitutes an equilibrium. In games without absentmindedness, these two equilibrium concepts coincide. In games with absentmindedness, the concepts differ in how expected utilities are evaluated for an action a at a current info set I, given that the player plays according to strategy μ anywhere "else". Computing such expected utilities requires
1. A Belief System: A method to form beliefs (i.e., a probability distribution) over being at a specific node/history of Γ given that the player is at info set I; and
2. A Decision Theory: An understanding of how an action choice at the current node affects the freedom to choose an action at other nodes of *the same* info set.

In the sequel, let I be the player's current info set at which she finds herself during play while playing μ in game Γ.

3.1 Decision Theories

Causal Decision Theory (CDT) postulates that the player can take an action $\alpha \in \Delta(A_I)$ at the current node without violating that the player has been playing according to μ at past arrivals at I, or that she will be playing according to μ at future arrivals at I. The intuition behind CDT is that the player's choice to deviate away from μ at the current node does not *cause* any change in behaviour at any other node of the same info set I.

In contrast to that, Evidential Decision Theory (EDT) postulates that if the player takes an action $\alpha \in \Delta(A_I)$ at the current node, then she will have also deviated to α whenever she arrived in I in past play, and she will be deviating to α whenever she arrives in I again in future play. Indeed, EDT

argues that the choice to deviate to α now is evidence for the player taking the same deviation choice in the past and future.

Denote with $\mu_{I \mapsto \alpha}$ an EDT deviation, i.e., the strategy of Γ that plays according to μ at every info set except at the info set $I \in \mathcal{I}_*$ where it plays according to $\alpha \in A_I$. By contrast, a CDT deviation may result in different actions taken at the same info set. This might not constitute a valid strategy that the player could have picked before the game started.

Example 5. *Consider the game in Figure 1 and suppose the player enters the game with the strategy $\mu = (R, X)$. Say, upon visiting info set I_1, the player plans to deviate from the μ-prescribed action R to the action L this one time only. Then, CDT argues that the player will stick to her μ-prescribed action R at the other node of I_1, leading to one of the two action histories (L, R, X) or (R, X). EDT, on the other hand, argues that such a deviation will then happen at both nodes of I, leading to the action histories (L,L).*

3.2 Self-Locating Belief Systems

Let $I^{1\text{st}} \subseteq I$ refer to those nodes $h \in I$ that are the first node of their history to enter info set I. Define the reach probability and (expected) visit frequency of I under μ as $\mathbb{P}(I \mid \mu) := \sum_{h \in I^{1\text{st}}} \mathbb{P}(h \mid \mu)$ and $\text{Fr}(I \mid \mu) := \sum_{h \in I} \mathbb{P}(h \mid \mu)$. Note that the reach probability and the visit frequency can only differ in games with absentmindedness, and that the visit frequency can be greater than 1. However, we have in general that $\mathbb{P}(I \mid \mu) > 0$ if and only if $\text{Fr}(I \mid \mu) > 0$. Finally, denote with $\chi : P \to \{0, 1\}$ the function that takes a Boolean property P as input and evaluate 1 if and only if P is true.

The first belief system argues that one should focus on the visit frequencies:

Definition 6. *Let I be an info set with $\text{Fr}(I \mid \mu) > 0$ under μ, and let $h \in \mathcal{H}_*$ be a player node. Then, Generalized Thirding (GT) determines the probability of the player to be at h, given that she uses μ and is currently in I, through*

$$\mathbb{P}_{\text{GT}}(h \mid \mu, I) := \chi(h \in I) \cdot \frac{\mathbb{P}(h \mid \mu)}{\text{Fr}(I \mid \mu)}.$$

The second belief system argues that one should rather focus on the reach probabilities. Note that the statement $I \cap \text{hist}(z) \neq \emptyset$ evaluates as true if and only if I occurs in history $\text{hist}(z)$ of terminal node $z \in \mathcal{Z}$ at least once.

Definition 7. *Let I be an info set with $\mathbb{P}(I \mid \mu) > 0$ under μ, and let $z \in \mathcal{Z}$ be a terminal node. Then, Generalized Double Halving (GDH) determines the probability of the player being on the path $\text{hist}(z)$ to terminal node z, given that she uses μ and is currently in I, through*

$$\mathbb{P}_{\text{GDH}}(\text{hist}(z) \mid \mu, I) := \chi(I \cap \text{hist}(z) \neq \emptyset) \cdot \frac{\mathbb{P}(z \mid \mu)}{\mathbb{P}(I \mid \mu)}.$$

GT and GDH were introduced as "consistency" and "z-consistency" by [Piccione and Rubinstein, 1997].

With the current definitions, GT and GDH assign probabilities to different type of events (to be at player node versus to be in the history of a terminal node). In the appendix, we phrase GT and GDH in each other's language. In the language of GDH, GT assigns the event of being in history $\text{hist}(z)$ of a terminal node z a higher probability if the reach probability of z under μ is higher (same as GDH) and if $\text{hist}(z)$ visits info set I very often (whereas GDH only cares about I being visited at least once by $\text{hist}(z)$).

Example 8. *Consider the game in Figure 1 again and suppose the player enters the game with the strategy $\mu = (\frac{1}{2}L + \frac{1}{2}R, X)$. Say, the player observes to be in info set I_1. Then a GT player believes to be at the node h_0 in the history (R, X) with probability $\frac{1}{3}$ whereas a GDH player believes to be at h_0 in (R, X) with probability $\frac{1}{2}$. The names "Halving" and "Thirding" originate from this contrast but for a different example called Sleeping Beauty [Elga, 2000] (see appendix).*

3.3 Two Equilibrium Concepts

Any claims made in this section are proven in the appendix.

We start with the equilibrium concept that uses Causal Decision Theory and Generalized Thirding. Denote with $h \circ a$ the child node reached in Γ by following action $a \in A_h$ from player node $h \in \mathcal{H}_*$. Then $U(\mu \mid h \circ a)$ is the expected utility the player receives from being at h, playing a now, and playing according to μ afterwards.

Definition 9. *Let the player currently be at an info set I with $\text{Fr}(I \mid \mu) > 0$, and let $\alpha \in \Delta(A_I)$ be a mixed action. Then, the (CDT,GT)-expected utility of playing α now and according to μ otherwise is*

$$\text{EU}_{\text{CDT,GT}}(\alpha \mid \mu, I) := \sum_{h \in I} \mathbb{P}_{\text{GT}}(h \mid \mu, I) \cdot \left(\sum_{a \in A_I} \alpha(a) \cdot U(\mu \mid h \circ a) \right).$$

Note that the inner sum collapses to $U(\mu \mid h \circ a)$ if the considered mixed action α is a pure action a.

Definition 10. *We say a strategy μ^* of Γ is a (CDT,GT)-equilibrium if for all info sets $I \in \mathcal{I}_*$ with $\text{Fr}(I \mid \mu^*) > 0$ under μ^*, we have*
$\mu^*(\cdot \mid I) \in \text{argmax}_{\alpha \in \Delta(A_I)} \text{EU}_{\text{CDT,GT}}(\alpha \mid \mu^*, I)$.

Alternatively, we can use the easier-to-check condition that for all info sets $I \in \mathcal{I}_*$ with $\text{Fr}(I \mid \mu^*) > 0$ and all pure actions $a \in A_I$ with $\mu^*(a \mid I) > 0$, we have

$$a \in \underset{a' \in A_I}{\text{argmax}} \, \text{EU}_{\text{CDT,GT}}(a' \mid \mu^*, I). \tag{2}$$

Next, we introduce the equilibrium concept that uses Evidential Decision Theory and Generalized Double Halving.

Definition 11. *Let the player currently be at an info set I with $\mathbb{P}(I \mid \mu) > 0$, and let $\alpha \in \Delta(A_I)$ be a mixed action. Then, the (EDT,GDH)-expected utility of playing α now and according to μ otherwise is*

$$\text{EU}_{\text{EDT,GDH}}(\alpha \mid \mu, I) := \sum_{z \in \mathcal{Z}} \mathbb{P}_{\text{GDH}}(\text{hist}(z) \mid \mu_{I \mapsto \alpha}, I) \cdot u(z)$$

The GDH belief probabilities in Definition 11 are well-defined due to $\mathbb{P}(I \mid \mu) = \mathbb{P}(I \mid \mu_{I \mapsto \alpha})$.

Definition 12. *We say a strategy μ^* of Γ is a (EDT,GDH)-equilibrium if for all info sets $I \in \mathcal{I}_*$ with $\mathbb{P}(I \mid \mu^*) > 0$ under μ^*, we have*
$\mu^*(\cdot \mid I) \in \text{argmax}_{\alpha \in \Delta(A_I)} \text{EU}_{\text{EDT,GDH}}(\alpha \mid \mu^*, I)$.

For (EDT,GDH), it is not sufficient to only check for optimality of pure actions that are in the support of μ^*. For instance, take the game in Figure 1 and suppose the player enters the game with the strategy $\mu = (C, X)$. Say, the player observes to be in info set I_1. Then action C is optimal among pure actions {L, C, R}. But the player would strictly benefit from deviating to mixed action $\frac{1}{2}L + \frac{1}{2}R$.

Finally, observe that in games without absentmindedness, the following notions coincide: CDT and EDT, GT and GDH, and Definitions 9 and 11. In particular, both equilibrium concepts coincide:

Lemma 13. *In games without absentmindedness, a strategy μ is a (CDT,GT)-equilibrium if and only if it is an (EDT,GDH)-equilibrium.*

3.4 Equilibria from the Ex-Ante Perspective

Recall from Section 2.2 that the (ex-ante) strategy utility function U of Γ is a polynomial function from $\times_{i=1}^{\ell} \mathbb{R}^{m_i}$ to $\mathbb{R}$. In this section, we give characterizations for (CDT,GT)- and (EDT,GDH)-equilibria in terms of U, as presented by Oesterheld and Conitzer [2022] and Piccione and Rubinstein [1997]. We reprove these results in the appendix since our setup and end goal differs slightly.

Polynomial U is continuously differentiable in $\mu \in \times_{i=1}^{\ell} \mathbb{R}^{m_i}$. For $i \in [\ell]$ and $j \in [m_i]$, let $\nabla_{ij} U$ stand for the partial derivative in direction (i, j), that is, the linear change of U at a point μ if you infinitesimally increase its $\mu(a_j \mid I_i)$ value.

Lemma 14. *Let I_i be an info set, $a_j \in A_{I_i}$ an action, and $\mu \in \times_{i=1}^{\ell} \Delta(A_{I_i})$ a strategy. Then:*
1. *$\nabla_{ij} U(\mu) = 0$ if $\mathrm{Fr}(I_i \mid \mu) = 0$, and*
2. *$\nabla_{ij} U(\mu) = \mathrm{Fr}(I_i \mid \mu) \cdot \mathrm{EU}_{\mathrm{CDT,GT}}(a_j \mid \mu, I_i)$ otherwise.*

Note that an infinitesimal increase of $\mu(a_j \mid I_i)$ means in a game-theoretic sense that the decision for action a_j is made slightly more probable at *every* node of info set I_i. This resembles an EDT type of deviation power but restricted to small deviations that stay close to the current action profile μ. Then, Lemma 14 says that a CDT deviation – rescaled by $\mathrm{Fr}(I_i \mid \mu)$ – accurately captures the linear (=dominant) effect of such a "local EDT deviation".

Lemma 15. *Strategy $\mu \in \times_{i=1}^{\ell} \Delta(A_{I_i})$ of Γ is an (EDT,GDH)-equilibrium if and only if for all $i \in [\ell]$:*
$$\mu_{i\cdot} \in \underset{y \in \Delta(A_{I_i})}{\mathrm{argmax}} \ U(\mu_{1\cdot}, \ldots, \mu_{i-1\cdot}, y, \mu_{i+1\cdot}, \ldots, \mu_{\ell\cdot}).$$

One possible interpretation of Lemma 15 is that (EDT,GDH)-equilibria of Γ are exactly the Nash equilibria of an ℓ-player simultaneous and identical-interest game G: Each player i shall have the continuous action space $\Delta(A_{I_i})$ and the (single) utility payoff, a function of the chosen action profile $(\mu_{i\cdot})_{i=1}^{\ell} \in \times_{i=1}^{\ell} \Delta(A_{I_i})$, shall be the polynomial U.

3.5 Computational Considerations

One of our main results addresses the complexity of finding a (CDT,GT)-equilibrium. There are problem instances where all (CDT,GT)-equilibria take on irrational numbers even though the game is easy to encode (see appendix).

Therefore, we relax our search to ϵ-approximate (CDT,GT)-equilibria where ϵ is an inverse-exponential numerical precision parameter.

Definition 16. *An instance of the problem* (CDT,GT)-EQUILIBRIUM *consists of a single-player extensive-form game Γ with imperfect recall and a precision parameter $\epsilon > 0$ encoded in binary. A solution consists of a strategy μ for Γ that satisfies for all $I \in \mathcal{I}_*$ with $\mathrm{Fr}(I \mid \mu) > 0$:*

$$\mathrm{EU}_{\mathrm{CDT,GT}}\big(\mu(\cdot \mid I) \mid \mu, I\big) \geq \max_{a' \in A_I} \mathrm{EU}_{\mathrm{CDT,GT}}(a' \mid \mu, I) - \epsilon.$$

There is also an alternative notion of being close to an equilibrium, called ϵ-well-supported (CDT,GT)-equilibrium. It instead requires condition (2) to be satisfied up to ϵ precision. An analysis of when both approximation concepts are polynomial-time related can be found in the appendix.

We will give hardness results and restricted membership results for (CDT,GT)-EQUILIBRIUM for the class CLS (Continuous Local Search). CLS was introduced by Daskalakis and Papadimitriou [2011] who noted that it contains various important problems of continuous local optimization that belong both to PPAD and PLS. PPAD [Papadimitriou, 1994] is well-known as the class that captures the complexity of many problems of Nash equilibrium computation ([Daskalakis *et al.*, 2009; Chen *et al.*, 2009] and much subsequent work), while PLS (Polynomial Local Search [Johnson *et al.*, 1988]) represents the complexity of many problems of discrete local optimization. Recently, Fearnley *et al.* [2023] showed that CLS is *equal* to the intersection of PPAD and PLS, indicating that CLS-hardness is quite a reliable notion of computational difficulty. In addition, the hardness of CLS can also be based on the cryptographic assumption of indistinguishability obfuscation [Hubáček and Yogev, 2017]. Fearnley *et al.* [2023] also showed that a version of the KKT point search problem is CLS-complete. Subsequent work has established CLS-completeness of mixed Nash equilibria of congestion games [Babichenko and Rubinstein, 2021] and solutions to a certain class of contests [Elkind *et al.*, 2022]. We can characterize CLS through any of its complete problems.

We will mainly be interested in KKT points. Consider a general non-linear maximization problem

$$\max_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad Bx + b \leq 0, \, Cx + c = 0 \qquad (3)$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable, $B \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $C \in \mathbb{R}^{\ell \times n}$, and $c \in \mathbb{R}^\ell$, and the domain is bounded. A point $x \in \mathbb{R}^n$ is then said to be a KKT point for (3) if there exist KKT multipliers $\tau_1, \ldots, \tau_m, \kappa_1, \ldots, \kappa_\ell \in \mathbb{R}$ such that $Bx + b \leq 0$ and $Cx + c = 0$, $\forall j \in [m] : \tau_j \geq 0$, $\forall j \in [m] : \tau_j = 0$ or $B_{j\cdot}x + b_j = 0$, and

$$\nabla f(x) = \sum_{j=1}^{m} \tau_j \cdot (B_{j\cdot})^T + \sum_{i=1}^{\ell} \kappa_i \cdot (C_{i\cdot})^T = 0.$$

The KKT conditions are necessary first-order conditions for a point to be a local optimum of (3). Furthermore, feasible stationary points satisfy the KKT conditions.
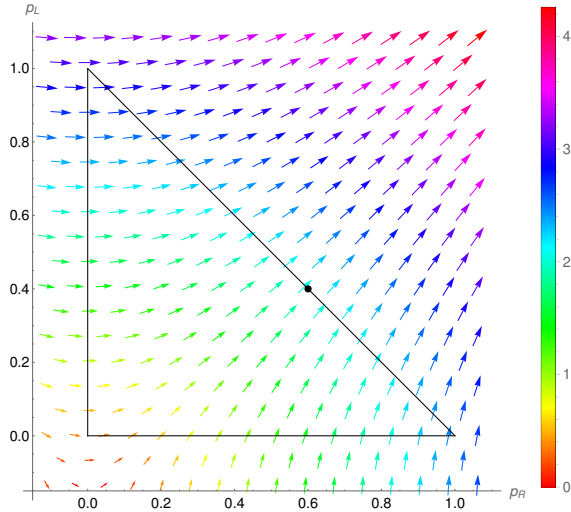
Figure 2: A plot of the gradient vector field ∇U of the strategy utility function U. The underlying game is the one from Figure 1 except where the player nodes of info set I_2 are replaced by chance nodes that choose actions X and Y with chance probabilities $\frac{1}{2}$ each. Then the player only has to choose a mixed action for info set I_1. The plot is in 2D for visualization convenience: The x-axis and y-axis represent the probabilities put on action R and L respectively. The action simplex $\Delta(\{L, C, R\})$ becomes a right triangle with the point $(0, 0)$ corresponding to pure action C. The gradient coloring represents the vector length.

There is no KKT point in the interior of the (projected) simplex because the gradient does not vanish there. The KKT points on the boundary of the simplex are those where the gradient is directed perpendicular outwards of the boundary constraint (except corner points, whose gradient only needs to lie in the positive cone of the boundary constraint directions). Thus $(0.6 \cdot L + 0.4 \cdot R)$ is the only KKT point. A game-theoretic analysis of the underlying game Γ also yields $(0.6 \cdot L + 0.4 \cdot R)$ as the only (CDT,GT)-equilibrium.

4 Main Results

To our knowledge, the results of this section are all novel unless explicitly stated otherwise. All proofs can be found in the appendix.

4.1 Complexity of the Search Problems

First, we use Lemma 14 to give a characterization of (CDT,GT)-equilibria in terms of ex-ante utility. For that, recall the ex-ante maximization problem (1).

Theorem 1. *Strategy* $\mu \in \times_{i=1}^{\ell} \Delta(A_{I_i})$ *of* Γ *is a (CDT,GT)-equilibrium if and only if* μ *is a KKT point of (1).*

We visualize Theorem 1 in Figure 2. This result also reveals a method to find (CDT,GT)-equilibria, namely by applying Gradient Descent on U. Note that in continuous optimization, there can be KKT points that are not locally optimal. An analogous effect can also happen in games with imperfect recall: In the game of Figure 1, strategy $\mu = (C, X)$ is a (CDT,GT)-equilibrium. But it is not a local optimum because for any $\epsilon > 0$ a shift from $\mu(\cdot \mid I_1) = C$ to $\epsilon \cdot (\frac{1}{2}L + \frac{1}{2}R) + (1 - \epsilon) \cdot C$ would yield the player ex-ante utility $5 \cdot \frac{\epsilon}{2} \cdot \frac{\epsilon}{2} > 0$. However, from a (CDT,GT) standpoint,

the player should be satisfied with her choice at I_1. In terms of the original definition of CDT, this is because deviating exactly *once* in I_1 never suffices to attain a utility of 5. In terms of Lemma 14 and Theorem 1, the issue is that the first order effect of increasing the probabilities of R and L is 0.

The three solution concepts considered in this paper form an inclusion hierarchy, a result shown by Oesterheld and Conitzer [2022] [cf. Piccione and Rubinstein, 1997]:

Lemma 17. *An ex-ante optimal strategy of a game* Γ *is also an (EDT,GDH)-equilibrium. An (EDT,GDH)-equilibrium is also a (CDT,GT)-equilibrium.*

In particular, any single-player extensive-form game Γ *with imperfect recall admits an (EDT,GDH)-equilibrium and a (CDT,GT)-equilibrium.*

The implication chain of Lemma 17 does not hold in the reverse direction: Consider the game in Figure 1. Then strategy $\mu = (C, X)$ is a (CDT,GT)-equilibrium, but not an (EDT,GDH)-equilibrium. Moreover, strategy $\mu' = (R, Y)$ is an (EDT,GDH)-equilibrium with ex-ante utility 1. This is not ex-ante optimal because strategy $\mu'' = (\frac{1}{2}L + \frac{1}{2}R, X)$ achieves the (optimal) ex-ante utility $5/4$.

The second part of Lemma 17 holds because ex-ante optimal strategies always exist. This is in contrast to the multiplayer setting where Nash equilibria may not exist in the presence of imperfect recall. Moreover, Lemma 17 implies that finding an ex-ante optimal strategy must be at least as hard as finding an (EDT,GDH)-equilibrium which must be at least as hard as finding a (CDT,GT)-equilibria. For the latter, we get the following classification:

Theorem 2. (CDT,GT)-EQUILIBRIUM *is CLS-hard. CLS-hardness holds even for games restricted to:*
(1.) a tree depth of 6 *and the player has* 2 *actions per info set,*
(2.) no absentmindedness and a tree depth of 6, *and*
(3.) no chance nodes, a tree depth of 5, *and only one info set.*
The problem is in CLS for the subclass of problem instances of (CDT,GT)-EQUILIBRIUM *where a lower bound on positive visit frequencies in* Γ *is easily obtainable.*

All CLS results of Theorem 2 also hold analogously for the search problem of an approximate well-supported (CDT,GT)-equilibrium. We prove (1.) by a reduction from finding a KKT point of a polynomial function over the hypercube. For (2.) and (3.), we reduce from finding a Nash equilibrium of a polytensor identical interest game. Both search problems we reduce from were shown to be CLS-complete by Babichenko and Rubinstein [2021].

The CLS membership in Theorem 2 implies that, unless NP = co-NP, the considered problem cannot be hard for the class NP [Megiddo and Papadimitriou, 1991]. We only prove CLS membership for those games Γ where we can construct a lower bound value $\lambda > 0$ that satisfies $\text{Fr}(I \mid \mu) = 0$ or $\text{Fr}(I \mid \mu) \geq \lambda$ for all strategies μ and info sets I in Γ. That is because if $\text{Fr}(I \mid \mu) > 0$ is too small, the approximation error may explode when transitioning from the ex-ante perspective $\nabla_{ij} U(\mu)$ to the *de se* perspective $\text{EU}_{\text{CDT,GT}}(a_j \mid \mu, I_i)$. Fortunately, such a lower bound exists and is polynomial-time computable for many well-known imperfect-recall games, such as the Absentminded Driver,

Sleepy Beauty type problems, or all the games used in the CLS-hardness results of Theorem 2. Thus, the computation of an approximate (CDT,GT)-equilibrium is CLS-complete in those games that admit such a lower bound on positive visit frequencies. Statement (2.) shows in particular that absent-mindedness is not the reason for CLS hardness. With Lemma 13, this implies

Corollary 18. *In games without absentmindedness where a lower bound on positive visit frequencies is easily obtainable, it is CLS-complete to find an ϵ-(EDT,GDH)-equilibrium.*

The authors are not aware of any complexity classification for the problem of *finding* an approximate (EDT,GDH)-equilibrium in games that may have absentmindedness – even though Lemma 15 gives a nice characterization of (EDT,GDH)-equilibria. Nonetheless, we are able to give conditional inapproximability results for (EDT,GDH)-equilibria with the next theorem.

4.2 Complexity of the Decision Problems

Next, we show that maximizing expected utility in an info set or maximizing over the space of equilibria is NP-hard. In the following problem formulations, any target value $t \in \mathbb{Q}$ shall be encoded in binary.

Theorem 3. *The following problems are all NP-hard. Unless NP = ZPP, there is also no FPTAS for these problems.*
(1a.) Given Γ and $t \in \mathbb{Q}$, is there a (CDT,GT)-equilibrium of Γ with ex-ante utility $\geq t$?
(1b.) Given Γ, an info set I of Γ and $t \in \mathbb{Q}$, is there a (CDT,GT)-equilibrium μ with $\mathrm{Fr}(I \mid \mu) > 0$, and such that the player has a (CDT,GT)-expected utility $\geq t$ upon reaching I?
(1c.) Given Γ, an info set I of Γ and $t \in \mathbb{Q}$, is there a strategy μ of Γ with $\mathrm{Fr}(I \mid \mu) > 0$, and such that the player has a (CDT,GT)-expected utility $\geq t$ upon reaching I?
(2a.) Given Γ and $t \in \mathbb{Q}$, is there an (EDT,GDH)-equilibrium of Γ with ex-ante utility $\geq t$?
(2b.) Given Γ, an info set I of Γ and $t \in \mathbb{Q}$, is there an (EDT,GDH)-equilibrium μ with $\mathbb{P}(I \mid \mu) > 0$, and such that the player has an (EDT,GDH)-expected utility $\geq t$ upon reaching I?
(2c.) Given Γ, an info set I of Γ and $t \in \mathbb{Q}$, is there a strategy μ of Γ with $\mathbb{P}(I \mid \mu) > 0$, and such that the player has an (EDT,GDH)-expected utility $\geq t$ upon reaching I?
(3a.) Given Γ and $t \in \mathbb{Q}$, do all (EDT,GDH)-equilibria of Γ have ex-ante utility $\geq t$?
(3b.) Given Γ, an info set I of Γ and $t \in \mathbb{Q}$, do all (EDT,GDH)-equilibria μ with $\mathbb{P}(I \mid \mu) > 0$ yield the player an (EDT,GDH)-expected utility $\geq t$ upon reaching I?

All the results of Theorem 3 follow from Proposition 4. Therefore, NP-hardness and conditional inapproximability remain for problems of the form (-a.) even if we restrict the game instances as described in Proposition 4. The same holds for problems of the form (-b.) and (-c.) except that we have to add one info set and one tree depth level to the game instances. Hardness of decision problem (3a.) also relies on the observation that in games with one info set only, any

(EDT,GDH)-equilibrium is also ex-ante optimal (cf. Lemma 15).

From (3a.) we obtain in particular that, unless NP = ZPP, there is no FPTAS for the search problem of an (EDT,GDH)-equilibrium in games with imperfect recall[1]. To compare this to Theorem 2, we shall remark that the conditional inapproximability result for (EDT,GDH)-equilibria (and ex-ante optimal strategies) is obtained even for games where a lower bound on positive visit frequencies is easily obtainable.

The decision problems of the form (1-.), (2a.), and (2c.) are all members of the complexity class $\exists \mathbb{R}$, and therefore, in particular, in PSPACE. On one hand, this is because (CDT,GT)-expected utilities and (EDT,GDH)-expected utilities can be described as rational functions (fractions of polynomial functions). Furthermore, this is because the alternative definition (2) of a (CDT,GT)-equilibrium gives rise to polynomially many comparisons of polynomial functions, and, for (2a.), because ex-ante optimal strategies are (EDT,GDH)-equilibria.

5 Conclusion

Games of imperfect recall have traditionally often been considered a theoretical curiosity; it is hard to model settings with human actors as imperfect-recall games, because, while most of us frequently forget things, we do not *reliably* forget things according to well-specified rules. For AI agents, however, this is no longer true; moreover, because they can be instantiated many times, sometimes in simulation, one instantiation will generally not know what another knew earlier. All this motivates the *computational* study of games of imperfect recall, which we initiated here for the single-player case. We are aided in this endeavor by recent conceptual work that specifies and motivates several natural solution concepts, and we based our work on these. Standard polynomial-time algorithms such as ones based on the sequence form are known to no longer work in the presence of imperfect recall. In this paper we found various complexity-theoretic evidence that indeed, single-player imperfect-recall games are hard to solve. Some of this evidence is, intriguingly, based on the complexity class CLS whose careful study is only very recent. On the positive side, we also provided insights into solving such games by drawing close connections to several problems about maximizing polynomial functions.

There remain many avenues for future work. What can be said about these computational problems for representation schemes other than the extensive form? Are there special cases of imperfect-recall games that can be solved more efficiently, whether they are single-player or multi-player? One may also ask whether our results give insight into the more conceptual questions. For example, to the extent that (CDT,GT)-equilibria are (under reasonable complexity assumptions) easier to compute than (EDT,GDH)-equilibria, does that provide support for using the former solution concept, at least for certain purposes? We hope that the work we have done in this paper can serve as a springboard for further research into this fascinating and important topic.

---

[1] Note that FPTAS even allow for an inverse-*polynomial* precision ϵ.

Acknowledgements

References

[Babichenko and Rubinstein, 2021] Yakov Babichenko and Aviad Rubinstein. Settling the complexity of nash equilibrium in congestion games. In *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1426–1437. ACM, 2021.

[Briggs, 2010] Rachael Briggs. Putting a value on Beauty. In Tamar Szabó Gendler and John Hawthorne, editor, *Oxford Studies in Epistemology: Volume 3*, pages 3–34. Oxford University Press, 2010.

[Camerer, 2003] Colin F. Camerer. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press, 2003.

[Canny, 1988] John Canny. Some algebraic and geometric computations in pspace. In *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, STOC '88, page 460–467, New York, NY, USA, 1988. Association for Computing Machinery.

[Chen *et al.*, 2009] Xi Chen, Xiaotie Deng, and Shang-Hua Teng. Settling the complexity of computing two-player Nash equilibria. *J. ACM*, 56(3):14:1–14:57, 2009.

[Chu and Halpern, 2001] Francis C. Chu and Joseph Y. Halpern. On the np-completeness of finding an optimal strategy in games with common payoffs. *Int. J. Game Theory*, 30(1):99–106, 2001.

[Conitzer and Oesterheld, 2023] Conitzer and Oesterheld. Foundations of cooperative ai. In *Thirty-Seventh AAAI Conference on Artificial Intelligence, AAAI 2023, February 7 - 14, 2022*. AAAI Press, 2023.

[Conitzer, 2015] Vincent Conitzer. A Dutch book against sleeping beauties who are evidential decision theorists. *Synthese*, 192(9):2887–2899, 2015.

[Daskalakis and Papadimitriou, 2011] Constantinos Daskalakis and Christos Papadimitriou. Continuous local search. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '11, page 790–804, USA, 2011. Society for Industrial and Applied Mathematics.

[Daskalakis *et al.*, 2009] Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The complexity of computing a Nash equilibrium. *SIAM J. Comput.*, 39(1):195–259, 2009.

[de Klerk, 2008] Etienne de Klerk. The complexity of optimizing over a simplex, hypercube or sphere: a short survey. *Central Eur. J. Oper. Res.*, 16(2):111–125, 2008.

[Draper and Pust, 2008] Kai Draper and Joel Pust. Diachronic Dutch Books and Sleeping Beauty. *Synthese*, 164(2):281–287, 2008.

[Elga, 2000] Adam Elga. Self-locating belief and the Sleeping Beauty problem. *Analysis*, 60(2):143–147, 2000.

[Elkind *et al.*, 2022] Edith Elkind, Abheek Ghosh, and Paul W. Goldberg. Simultaneous contests with equal sharing allocation of prizes: Computational complexity and price of anarchy. In *Algorithmic Game Theory - 15th International Symposium, SAGT*, volume 13584 of *Lecture Notes in Computer Science*, pages 133–150. Springer, 2022.

[Fearnley *et al.*, 2023] John Fearnley, Paul Goldberg, Alexandros Hollender, and Rahul Savani. The complexity of gradient descent: CLS = PPAD ∩ PLS. *J. ACM*, 70(1):7:1–7:74, 2023.

[Fudenberg and Tirole, 1991] Drew Fudenberg and Jean Tirole. *Game Theory*. MIT Press, October 1991.

[Gimbert *et al.*, 2020] Hugo Gimbert, Soumyajit Paul, and B. Srivathsan. A bridge between polynomial optimization and games with imperfect recall. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '20, page 456–464, Richland, SC, 2020. International Foundation for Autonomous Agents and Multiagent Systems.

[Hitchcock, 2004] Christopher Hitchcock. Beauty and the bets. *Synthese*, 139(3):405–420, 2004.

[Hubácek and Yogev, 2017] Pavel Hubácek and Eylon Yogev. Hardness of continuous local search: Query complexity and cryptographic lower bounds. In Philip N. Klein, editor, *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA*, pages 1352–1371. SIAM, 2017.

[Johnson *et al.*, 1988] David S. Johnson, Christos H. Papadimitriou, and Mihalis Yannakakis. How easy is local search? *Journal of Computer and System Sciences*, 37(1):79–100, 1988.

[Koller and Megiddo, 1992] Daphne Koller and Nimrod Megiddo. The complexity of two-person zero-sum games in extensive form. *Games and Economic Behavior*, 4(4):528–552, 1992.

[Kovarik *et al.*, 2023] Vojtech Kovarik, Caspar Oesterheld, and Vincent ConitzerYuta. Game theory with simulation of other players. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI 2023, Macao, 9-25 August 2023*. ijcai.org, 2023.

[Kroer and Sandholm, 2016] Christian Kroer and Tuomas Sandholm. Imperfect-recall abstractions with bounds in games. In *Proceedings of the Seventeenth ACM Conference on Economics and Computation (EC)*, pages 459–476, Maastricht, the Netherlands, 2016.

[Lanctot *et al.*, 2012] Marc Lanctot, Richard G. Gibson, Neil Burch, and Michael Bowling. No-regret learning in

extensive-form games with imperfect recall. In *Proceedings of the 29th International Conference on Machine Learning (ICML-12)*, Edinburgh, Scotland, UK, 2012.

[Megiddo and Papadimitriou, 1991] Nimrod Megiddo and Christos H. Papadimitriou. On total functions, existence theorems and computational complexity. *Theor. Comput. Sci.*, 81(2):317–324, apr 1991.

[Nisan *et al.*, 2007] Noam Nisan, Tim Roughgarden, Éva Tardos, and Vijay V. Vazirani, editors. *Algorithmic Game Theory*. Cambridge University Press, 2007.

[Oesterheld and Conitzer, 2022] Caspar Oesterheld and Vincent Conitzer. Can *de se* choice be *ex ante* reasonable in games of imperfect recall? https://www.andrew.cmu.edu/user/coesterh/DeSeVsExAnte.pdf, 2022. Working paper. Accessed: 2022-12-14.

[Papadimitriou, 1994] Christos H. Papadimitriou. On the complexity of the parity argument and other inefficient proofs of existence. *Journal of Computer and System Sciences*, 48(3):498–532, 1994.

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

[Renegar, 1992] James Renegar. On the computational complexity and geometry of the first-order theory of the reals. part i: Introduction. preliminaries. the geometry of semi-algebraic sets. the decision problem for the existential theory of the reals. *Journal of Symbolic Computation*, 13(3):255–299, 1992.

[Schaefer and Stefankovic, 2017] Marcus Schaefer and Daniel Stefankovic. Fixed points, nash equilibria, and the existential theory of the reals. *Theory Comput. Syst.*, 60(2):172–193, 2017.

[Shor, 1990] Peter W. Shor. Stretchability of pseudolines is NP-hard. In Peter Gritzmann and Bernd Sturmfels, editors, *Applied Geometry And Discrete Mathematics, Proceedings of a DIMACS Workshop, Providence, Rhode Island, USA, September 18, 1990*, volume 4 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 531–554. DIMACS/AMS, 1990.

[Waugh *et al.*, 2009] Kevin Waugh, Martin Zinkevich, Michael Johanson, Morgan Kan, David Schnizlein, and Michael H. Bowling. A practical use of imperfect recall. In *Eighth Symposium on Abstraction, Reformulation, and Approximation (SARA-09)*, Lake Arrowhead, CA, USA, 2009.