Imperfect-Recall Games: Equilibrium Concepts and Their Complexity

Emanuel Tewolde[1,4] , **Brian Hu Zhang**[1] , **Caspar Oesterheld**[1,4] ,
Manolis Zampetakis[2] , **Tuomas Sandholm**[1,5,6,7] , **Paul Goldberg**[3] and **Vincent Conitzer**[1,3,4]
[1]Carnegie Mellon University
[2]Yale University
[3]University of Oxford
[4]Foundations of Cooperative AI Lab (FOCAL)
[5]Strategic Machine, Inc., [6]Strategy Robot, Inc., [7]Optimized Markets, Inc.
emanueltewolde@cmu.edu, bhzhang@cs.cmu.edu, oesterheld@cmu.edu, manolis.zampetakis@yale.edu,
sandholm@cs.cmu.edu, paul.goldberg@cs.ox.ac.uk, conitzer@cs.cmu.edu

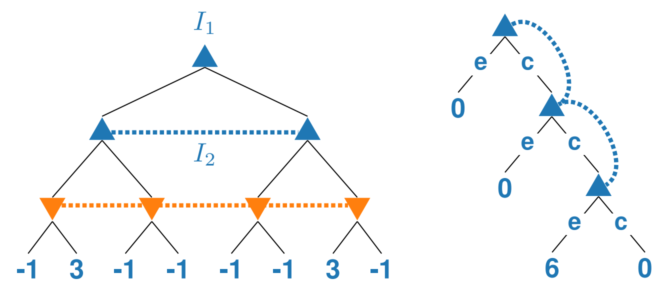
Abstract

We investigate optimal decision making under imperfect recall, that is, when an agent forgets information it once held before. An example is the absentminded driver game, as well as team games in which the members have limited communication capabilities. In the framework of extensive-form games with imperfect recall, we analyze the computational complexities of finding equilibria in multiplayer settings across three different solution concepts: Nash, *multiselves* based on evidential decision theory (EDT), and multiselves based on causal decision theory (CDT). We are interested in both exact and approximate solution computation. As special cases, we consider (1) single-player games, (2) two-player zero-sum games and relationships to maximin values, and (3) games without exogenous stochasticity (chance nodes). We relate these problems to the complexity classes P, PPAD, PLS, Σ_2^{P}, $\exists\mathbb{R}$, and $\exists\forall\mathbb{R}$.

1 Introduction

In game theory, it is common to restrict attention to games of *perfect recall*, that is, games in which no player ever forgets anything. At first, it seems that this assumption is even better motivated for AI agents than for human agents: humans forget things, but AI does not have to. However, we argue this view is mistaken: there are often reasons to design AI agents to forget, or to structure them so that they can be modeled as forgetful. Moreover, such forgetting-by-design follows predictable rules and is thereby easier to model formally than idiosyncratic human forgetting. Thus, games of imperfect recall are receiving renewed attention from AI researchers.

Imperfect recall is already being used for state-of-the-art *abstraction* algorithms for larger games of perfect recall [Waugh *et al.*, 2009; Ganzfried and Sandholm, 2014; Brown *et al.*, 2015]. The idea is that by forgetting unimportant aspects of the past, the AI can afford to conduct equilibrium-approximation computations with a game model that has a more refined abstraction of the present. Indeed,



(a) Forgetful penalty shoot-out. This game has no Nash equilibrium.

(b) Extended absentminded driver.

Figure 1: Games with imperfect recall. P1's (▲) utility payoffs are labeled on each terminal node. If P2 (▼) is present, the game is zero sum. Infosets are joined by dotted lines.

imperfect-recall abstractions were a key component in the first superhuman AIs in no-limit Texas hold'em poker [Brown and Sandholm, 2018, 2019].

Imperfect recall also naturally models settings in which forgetting is deliberate for other reasons, such as privacy of sensitive data [Conitzer, 2019; Zhang and Sandholm, 2022]. Conitzer provides the example of an AI driving assistant designed to intervene whenever the human car driver makes a significant error. In such instances, the AI must assess the overall skill level of the human driver, despite not being allowed to store information about the individual.

It can also model *teams* of agents with common goals and limited ability to communicate. Each team, represented by one agent with imperfect recall, is then striving for some notion of optimality among team members [von Stengel and Koller, 1997; Celli and Gatti, 2018; Emmons *et al.*, 2022; Zhang *et al.*, 2023]. Highly distributed agents are similarly well-described by imperfect recall: such an agent may take an action at one node based on information at that node, and then need to take another action at a second node without yet having learned yet what happened at the first node. Thus, effectively, the distributed agent has forgotten what it knew before. Finally, a single agent can be instantiated multiple times in the same environment, where one copy does not know what

Multi-player

	Nash (D)	EDT (D)	CDT (S)
exact	$\exists\mathbb{R}$-hard and in $\exists\forall\mathbb{R}$ (Thms. 1 & 3)		—
1/exp	Σ_2^{P}-complete (Thms. 2 and 4)		PPAD-complete (Thm. 6)
1/poly			

Single-player

	Optimal (D)	EDT (S)	CDT (S)
exact	$\exists\mathbb{R}$-complete [Gimbert *et al.*, 2020]	—	—
1/exp	NP-complete [Koller and Megiddo, 1992 Tewolde *et al.*, 2023]	PLS-complete (Thm. 5*)	CLS-complete [Tewolde *et al.*, 2023]
1/poly		P (Cor. 22*)	P (Cor. 17)

Table 1: Summary of complexity results. New results from this paper are shown with a light green background. (S) stands for search problem, which is when we ask for a solution strategy profile. In multi-player, (D) stands for deciding whether such an equilibrium even exists. In single-player, Optimal (D) decides whether some target utility can be achieved. Citations are given for results found in the literature. All of our hardness results even hold for highly restricted game instances. *: The number of actions per infoset is constant for these results. '—': No results exist for these settings to our knowledge. Also note the technical complication that arises here from the fact that there exist single-player games in which every exact EDT or CDT equilibrium involves irrational values [Tewolde *et al.*, 2023].

another copy just knew [Conitzer and Oesterheld, 2023]. For example, we might want to test goal-oriented AI agents in simulation to ensure that they will later act in a trustworthy fashion in the real world [Kovarík *et al.*, 2023, 2024]. Then, the AI agent will have to act in the real world without knowing how it acted in simulation.

Perfect recall is a common technical assumption in game theory because it implies many simplifying properties, such as polynomial-time solvability of single-player and two-player zero-sum settings [Koller and Megiddo, 1992]. In multi-player settings with *imperfect* recall, Nash equilibria may not exist anymore [Wichardt, 2008]; in fact, we show that deciding existence is computationally hard. To give an illustrative running example, consider a variation of Wichardt's game in Figure 1a, which we call the forgetful (soccer) penalty shoot-out. The shooter (P1) decides whether to shoot left or right, once before the whistle, and once again right before kicking the ball. At the second decision point, P1 has forgotten which direction they chose previously. P1 only succeeds in shooting in any direction if she chooses that direction at both decision points. Upon succeeding, it becomes a matching pennies game with the goalkeeper (P2) who chooses to jump left or right to block the ball. A similar analysis to the one of matching pennies implies that in a potential Nash equilibrium, none of the two players can play one side more often than the other. However, both players randomizing $50/50$ at each infoset is not a Nash equilibrium either: P1 is not best responding to P2 because she could instead deterministically shoot towards one side to avoid miscoordination with herself altogether which would achieve a payoff of 1 instead of 0.

Indeed, many of our intuitions fail for imperfect-recall games – to the point that a significant body of work in philosophy and game theory addresses conceptual questions about probabilistic reasoning and decision making in imperfect-recall games, such as in the Sleeping Beauty problem [Elga, 2000] or the absentminded driver game of Figure 1b [Piccione and Rubinstein, 1997]. From this literature, several distinct and coherent ways to approach games of imperfect recall have emerged. We will discuss these in detail in Section 4.

In this paper, we study the computational complexity of solving imperfect-recall extensive-form games. We focus on three solution concepts: (1) Nash equilibria where players play mutual best response strategies (or simply optimal strategies in single-player domains), (2) multiselves equilibria based on evidential decision theory, in which each infoset plays a best-response action to all other infosets and players, and (3) multiselves equilibria based on causal decision theory, in which each infoset plays a *Karush-Kuhn-Tucker (KKT)* point action for the current strategy profile. The latter two are relaxations of the first. Sections 2 and 4 cover preliminaries on imperfect-recall games and on multiselves equilibria, respectively. Sections 3 and 5 analyze the computation of Nash equilibria and of multiselves equilibria, respectively, in various setting. Our complexity results for these are summarized in Table 1. Last but not least, Section 6 shows that games with imperfect recall stay computationally equally hard even in the absence of exogenous stochasticity (*i.e.*, chance nodes).

2 Imperfect-Recall Games

We first define 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] and Piccione and Rubinstein [1997]. In this section, we follow the exposition of Tewolde *et al.* [2023], with the addition of introducing multi-player notation.

Definition 1. *An extensive-form* game with imperfect recall*, denoted by* Γ*, 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*. We denote the terminal nodes in* $\mathcal{H}$ *as* $\mathcal{Z}$ *and the set of actions available at a nonterminal node* $h \in \mathcal{H} \setminus \mathcal{Z}$ *as* A_h*.*
2. *A set of* $N+1$ *players* $\mathcal{N} \cup \{c\}$*, for* $N \in \mathbb{N}$*, and an assignment of nonterminal nodes to a player that shall choose an action at that node. Player* c *stands for* chance *and represents exogenous stochasticity that chooses an action. With* $\mathcal{H}^{(i)}$ *we denote all nodes associated to player* $i \in \mathcal{N}$*.*
3. *A fixed distribution* $\mathbb{P}^{(c)}(\cdot \mid h)$ *over* A_h *for each chance node* $h \in \mathcal{H}^{(c)}$*, with which an action is determined at* h*.*
4. *For each* $i \in \mathcal{N}$*, a* utility function $u^{(i)} : \mathcal{Z} \to \mathbb{R}$ *that specifies the payoff that player* i *receives from finishing the game at a terminal node.*
5. *For each* $i \in \mathcal{N}$*, a partition* $\mathcal{H}^{(i)} = \sqcup_{I \in \mathcal{I}^{(i)}} I$ *of player* i*'s decision nodes into information sets (*infosets*). We require* $A_h = A_{h'}$ *for all nodes* h, h' *of the same infoset. Therefore, infoset* I *has a well-defined action set* A_I*.*

Imperfect Recall. Nodes of the same infoset are assumed to be indistinguishable to the player during the game even though the player is always aware of the full game structure. This may happen even in perfect-recall games due to *imperfect information*, that is, when it is unobservable to the player what another player (or chance) has played. This effect is present in Figure 1a for P2. In contrast, infoset I_2 of P1 exhibits *imperfect recall* because once arriving there, the player has forgotten information about the history of play that she once held when leaving I_1, namely whether she chose left or right back then. In Figure 1b, the player is unable to recall whether she has been in the same situation before or not. This phenomenon is a special kind of imperfect recall called *absentmindedness*. The *degree of absentmindedness* of an infoset shall be defined as the maximum number of nodes of the same game trajectory that belong to that infoset. In Figure 1b, it is 3. The *branching factor* of a game is the maximum number of actions at any infoset.

In contrast to that, games with *perfect* recall have every infoset reflect that the player remembers the sequence of infosets she visited and the actions she took. We note that any node $h \in \mathcal{H}$ uniquely corresponds to a history path hist(h) in the game tree, consisting of alternating nodes and actions from root h_0 to h. Let $\exp^{(i)}(h)$ be the experienced sequence of infosets visited and actions taken by player i on the path hist(h). Then, formally, a game has perfect recall if for all players $i \in \mathcal{N}$, all infosets $I \in \mathcal{I}^{(i)}$, and all nodes $h, h' \in I$, we have $\exp^{(i)}(h) = \exp^{(i)}(h')$.

Strategies. Let $\Delta(A_I)$ denote the set of probability distributions over the actions in A_I. These will also be referred to as *randomized actions*. A (behavioral) *strategy* $\mu^{(i)} : \mathcal{I}^{(i)} \to \sqcup_{I \in \mathcal{I}^{(i)}} \Delta(A_I)$ of a strategic player i assigns to each of her infosets I a probability distribution $\mu^{(i)}(\cdot \mid I) \in \Delta(A_I)$. Upon reaching I, the player draws an action randomly from $\mu^{(i)}(\cdot \mid I)$. A *pure* strategy maps deterministically[1] to $\sqcup_{I \in \mathcal{I}^{(i)}} A_I$. A strategy profile, or *profile*, $\mu = (\mu^{(i)})_{i \in \mathcal{N}}$ specifies a behavioral strategy for each player. We may write $\left(\mu^{(i)}, \mu^{(-i)}\right)$ to emphasize the influence of $i \in \mathcal{N}$ on μ. Denote the strategy set of player $i \in \mathcal{N}$ with $\mathcal{S}^{(i)}$, and the set of profiles with $\mathcal{S}$.

For a computational analysis, we identify a randomized action set $\Delta(A_I)$ with the simplex $\Delta^{|A_I|-1}$, where $\Delta^{n-1} := \{x \in \mathbb{R}^n : x_k \geq 0 \, \forall k, \sum_{k=1}^n x_k = 1\}$. Therefore, the strategy sets are Cartesian products of simplices:

$$\mathcal{S} \equiv \times_{i \in \mathcal{N}} \times_{I \in \mathcal{I}^{(i)}} \Delta^{|A_I|-1} \text{ and } \mathcal{S}^{(i)} \equiv \times_{I \in \mathcal{I}^{(i)}} \Delta^{|A_I|-1}.$$

Reach Probabilities and Utilities. Let $\mathbb{P}(\bar{h} \mid \mu, h)$ be the probability of reaching node $\bar{h} \in \mathcal{H}$ given that the current game state is at $h \in \mathcal{H}$ and that the players are playing profile μ. It evaluates as 0 if $h \notin$ hist$(\bar{h})$, and as the product of probabilities of the actions on the path from h to $\bar{h}$ otherwise. The expected utility payoff of player $i \in \mathcal{N}$ at node $h \in \mathcal{H} \setminus \mathcal{Z}$ if profile μ is being followed henceforth is $U^{(i)}(\mu \mid h) := \sum_{z \in \mathcal{Z}} \mathbb{P}(z \mid \mu, h) \cdot u^{(i)}(z)$. We overload notation by defining $\mathbb{P}(h \mid \mu) := \mathbb{P}(h \mid \mu, h_0)$ for root h_0 of Γ, and by defining the function $U^{(i)}$ as $U^{(i)}(\mu) := U^{(i)}(\mu \mid h_0)$, mapping a profile μ to its expected utility from game start. In Figure 1b, this is $U^{(1)}(\mu) = 6c^2e$ – or, to follow our notation more precisely, $U^{(1)}(\mu) = 6\mu^{(1)}(c \mid I)^2 \mu^{(1)}(e \mid I)$.

Polynomials Each summand $\mathbb{P}(z \mid \mu, h) \cdot u^{(i)}(z)$ in $U^{(i)}(\mu \mid h)$ is a monomial in μ times a scalar, and the expected utility function $U^{(i)}$ is a polynomial function in the profile μ. All these polynomials $U^{(i)}$ can be constructed in polynomial time (polytime) in the encoding size of Γ.

One might also ask how general those polynomial utility functions may be. Indeed, imperfect-recall games can be very expressive. We give a polytime construction in the full version of this paper that, given a collection of N multivariate polynomials $p^{(i)} : \times_{i=1}^N \times_{j=1}^{\ell^{(i)}} \mathbb{R}^{m_j^{(i)}} \to \mathbb{R}$, yields an associated N-player game Γ with imperfect recall such that its expected utility functions satisfy $U^{(i)}(\mu) = p^{(i)}(\mu)$ on $\times_{i=1}^N \times_{j=1}^{\ell^{(i)}} \mathbb{R}^{m_j^{(i)}}$.

Approximate Solutions The solution concepts we investigate will have a definition of the abstract form "Strategy μ is a *solution* if for all $y \in Y$ we have $f(\mu) \geq f_\mu(y)$" for some set Y of alternatives and some utility/objective functions f and f_μ. Then, we call a strategy μ an ϵ-solution if $\forall \, y \in Y : f(\mu) \geq f_\mu(y) - \epsilon$.

Computational Considerations In this paper, we discuss *decision problems* and *search problems*. The former ask for a yes/no answer; the latter ask for a solution point. The input to these computational problems may be a game Γ, a precision parameter $\epsilon > 0$, and/or a target value t. Values in Γ, as well as ϵ and t are assumed to be rational. We assume that a game Γ is represented by its game tree structure, which has size $\Theta(|\mathcal{H}|)$, and by a binary encoding of its chance node probabilities and its utility payoffs. If there is a target t, then it shall be given in binary as well.

If there is no precision parameter ϵ, then we are dealing with problems involving *exact* solutions. In our settings, such problems are usually beyond NP because equilibria may require irrational probabilities and may therefore not be representable in finite bit length. In fact, Tewolde *et al.* [2023][Figure 6] give a simple single-player example in which the unique equilibrium takes on irrational values. That is, in part, why we will also be interested in approximations up to a small precision error $\epsilon > 0$. Here, we mean 'small' relative to the range of utility payoffs, which – by shifting and rescaling utilies – we can w.l.o.g. assume to be $[0, 1]$.

Remark. *By default, $\epsilon > 0$ will be given in binary, in which case we require* inverse-exponential *(*1/exp*) precision.*

Here, the term 'inverse-exponential' indicates that $1/\epsilon$ can be exponentially larger than the tree size $|\mathcal{H}|$. Occasionally,

[1]Other work has also considered *mixed* strategies, that is, probability distributions over all pure strategies. In the presence of imperfect recall, mixed strategies are not realization-equivalent to behavioral strategies [Kuhn, 1953]. Mixed strategies require the agent to coordinate her actions across infosets (*e.g.*, access to a correlation device): For example, in contrast to our introductory discussion on the forgetful penalty shoot-out (Figure 1a), this game does admit a Nash equilibrium *in mixed strategies* since P1 can now choose to kick left twice in a row 50% of the time and to kick right twice in a row the other 50% of the time. As this would imply a form of memory, it does not fit the motivation of this paper.

we may instead require *inverse-polynomial* (1/poly) precision, which is when ϵ is given in unary, or require constant precision, which is when ϵ is fixed to a constant > 0. Naturally, 1/exp precision is hardest to achieve.

Complexity Classes We give a brief overview of the complexity classes appearing in this paper, and refer to the full version of this paper for references and more details. The subset relationships of the complexities classes we present here are believed to be strict. P describes the decision problems that can be solved in polytime. NP describes the decision problems that can be solved in non-deterministic polytime. Σ_2^{P} describes the decision problems that can be solved in non-deterministic polytime if given oracle access to an NP solver, such as a SAT oracle. We have $\mathsf{P} \subseteq \mathsf{NP} \subseteq \Sigma_2^{\mathsf{P}} \subseteq$ PSPACE. NP and Σ_2^{P} are classes for decision problems that can be formulated as one over discrete variables (w.l.o.g. Boolean variables). Their counterparts for real-valued decision problems are the *first-order-of-the-reals* classes $\exists\mathbb{R}$ and $\exists\forall\mathbb{R}$: A $\exists\mathbb{R}$ problem asks whether a sentence of the form $\exists x_1 \dots \exists x_n F(x_1, \dots, x_n)$ is true, where the x_i represent real-valued variables and F represents a quantifier-free formula of (in-)equalities of real polynomials in rational coefficients. $\exists\forall\mathbb{R}$ is defined analogously, except for sentences of the form $\exists x \in \mathbb{R}^{n_1} \forall y \in \mathbb{R}^{n_2} F(x, y)$. We have $\mathsf{NP} \subseteq \exists\mathbb{R} \subseteq$ PSPACE $\cap \exists\forall\mathbb{R}$.

The complexity classes FP and FNP are the search problem analogues of P and NP. The landscape between FP and FNP is rich, and total NP search problems are those problems in FNP for which one knows that each problem instance admits a solution. The complexity classes in it can be characterized by the natural, but exponential-time method with which one can show that each problem instance admits a solution. For the class PPAD the method is that of a fixed point argument, as is the case, *e.g.*, for the existence of a Nash equilibrium. For the class PLS the method is that of a local optimization argument on a directed acyclic graph. For the class PLS the method is that of a CLS a local optimization argument on a bounded polyhedral (continuous) domain. We have P "$\subseteq$" CLS $=$ PPAD $\cap$ PLS and PPAD, PLS "$\subseteq$" NP.

3 Nash Equilibria and Optimal Play

In this section, we present our computational results for the classic and most important solution concept in game theory – the Nash equilibrium [Nash, 1950].

Definition 2. *A profile μ is said to be a* Nash equilibrium *(in behavioral strategies) for game Γ if for all player $i \in \mathcal{N}$, and all alternative strategies $\pi^{(i)} \in \mathcal{S}^{(i)}$, we have*

$$U^{(i)}(\mu^{(i)}, \mu^{(-i)}) \geq U^{(i)}(\pi^{(i)}, \mu^{(-i)}).$$

In a Nash equilibrium, no player has any utility incentives to deviate unilaterally to another strategy. Nash showed that any finite perfect-recall game admits at least one Nash equilibrium. In contrast, some finite imperfect-recall games have no Nash equilibrium, as discussed in the introduction. If there is only a single player, however, finding a Nash equilibrium – *i.e.*, finding an *optimal* strategy – reduces to maximizing a polynomial utility function over a compact strategy space. Such a solution is guaranteed to exist, and its value is unique. Therefore, one may ask instead whether some target value t can be achieved in a given game. In Figure 1b, this would result in the $\exists\mathbb{R}$-sentence $\exists e, c : \; 6c^2 e \geq t \wedge c \geq 0 \wedge e \geq 0 \wedge c + e = 1$. This is an easier task than *finding* an optimal strategy. Nonetheless, we have:

Proposition 3 (Gimbert *et al.*, 2020)**.** *Deciding whether a single-player game with imperfect recall admits a strategy with value $\geq t$ is $\exists\mathbb{R}$-complete.*

For approximation, consider problem OPT-D that asks to distinguish between whether $\exists \mu \in S : U^{(1)}(\mu) \geq t$ and whether $\forall \mu \in S : U^{(1)}(\mu) \leq t - \epsilon$.

Proposition 4 (Koller and Megiddo, 1992; Tewolde *et al.*, 2023)**.** OPT-D *is* NP*-complete.*

Technically, Koller and Megiddo establish hardness for the *exact* decision problem. We shall merely add the observation that their proof also implies NP-hardness of the approximate problem; and via the PCP theorem [Håstad, 2001], even for a constant precision $\epsilon < 1/8$.

3.1 Two-Player Zero-Sum Games

A *two-player zero-sum (2p0s)* game is a two-player game where $U^{(2)} = -U^{(1)}$. In that case utilities can be given in terms of P1, and P2 simply minimizes that utility.

Koller and Megiddo [1992] prove Σ_2^{P}-completeness of deciding in 2p0s games with imperfect recall whether the max-min value in pure-strategy play exceeds some utility target $\geq t$. We will consider behavioral strategies instead.

Definition 5. *In a 2p0s game Γ, the (behavioral)* max-min value *and* min-max value *are defined as*

$$\underline{U} := \max_{\mu^{(1)} \in S^{(1)}} \min_{\mu^{(2)} \in S^{(2)}} U^{(1)}(\mu^{(1)}, \mu^{(2)}),$$
$$\bar{U} := \min_{\mu^{(2)} \in S^{(2)}} \max_{\mu^{(1)} \in S^{(1)}} U^{(1)}(\mu^{(1)}, \mu^{(2)}).$$

Gimbert *et al.* [2020] prove that deciding $\underline{U} \geq t$ is in $\exists\forall\mathbb{R}$ and is $\exists\mathbb{R}$-hard. For approximation, we know the following.

Lemma 6 (Zhang *et al.*, 2023)**.** *It is Σ_2^{P}-complete to distinguish $\underline{U} \geq 0$ from $\underline{U} \leq -\epsilon$ in 2p0s games with imperfect recall. Hardness holds even with no absentmindedness and* 1/poly *precision.*

To leverage this result in the subsequent sections, we will first show a tight connection between the existence of Nash equilibria in a 2p0s game Γ, and Γ's min-max and max-min values. Define the *duality gap* of Γ as the difference

$$\Delta := \bar{U} - \underline{U} \geq 0.$$

In Figure 1a the duality gap is $1 - 0 = 1$.

Proposition 7. *Let Γ be a 2p0s game with imperfect recall. If $\Delta \leq \epsilon$ then Γ admits an ϵ-Nash equilibrium. Conversely, if Γ admits an ϵ-Nash equilibrium, then $\Delta \leq 2\epsilon$.*

In particular, there is an equivalence between Nash equilibrium existence and vanishing duality gap. This result is not specific to behavioral strategies in imperfect-recall games; it holds for any family of strategies in any 2p0s game.

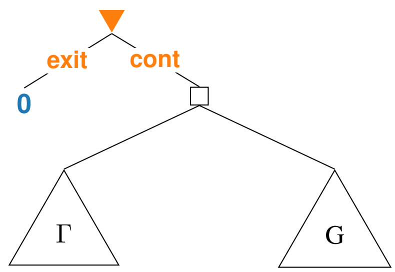


Figure 2: Game construction used to prove hardness of deciding equilibrium existence. We use boxes for chance nodes, at which chance plays uniformly at random. Γ is a placeholder game. G is a game with no equilibrium; Section 3.2 for example uses Figure 1a.

3.2 Deciding Nash Equilibrium Existence

We observe that the existence of a Nash equilibrium can be formulated as "there exists a profile μ such that for all other profiles π the condition of Definition 2 are satisfied for all $i \in \mathcal{N}$". This puts the exact and approximate decision problems in $\exists\forall\mathbb{R}$ and Σ_2^{P} respectively. For an intuitive idea of our upcoming hardness results, consider the game in Figure 2 where subgame G shall be that of Figure 1a and where subgame Γ is a game in which it is hard to decide what utility P1 can guarantee himself. Then a profile cannot be a Nash equilibrium if P2 is supposed to continue at the root node, because in that case G is reached with positive probability and the players cannot be in equilibrium in that subgame as we have discussed in the introduction. Note that exiting at the root node yields P2 a utility of 0, and best-responding to P1 in subgame G also yields P2 a utility of ≤ 0 (recall that P2 is the minimizer). Thus, for a profile to be a Nash equilibrium in the overall game, P2 must exit at the root node as a best response, which is the case exactly if P1 cannot achieve a utility of at least 0 in the subgame Γ. Using the problem instances of Proposition 3 for the subgame Γ, we obtain

Theorem 1. *Deciding if a game with imperfect recall admits a Nash equilibrium is $\exists\mathbb{R}$-hard and in $\exists\forall\mathbb{R}$. Hardness holds even for 2p0s games where on player has a degree of absentmindedness of 4 and the other player has perfect recall.*

Next, for the approximate case, we use the problem instances of Lemma 6 for the subgame Γ. Define NASH-D to ask to distinguish between whether an exact Nash equilibrium exists or whether no ϵ-Nash equilibrium exists.

Theorem 2. NASH-D *is Σ_2^{P}-complete. Hardness holds for 2p0s games with no absentmindedness and* 1/poly *precision.*

With Proposition 7, this immediately implies

Corollary 8. *It is Σ_2^{P}-complete to distinguish $\Delta = 0$ from $\Delta \geq \epsilon$ in 2p0s games. Hardness holds for 2p0s games with no absentmindedness and* 1/poly *precision.*

Later in this paper, Theorem 4 will imply another Σ_2^{P}-hardness for NASH-D but with different restrictions.

3.3 A Naïve Algorithm for Nash Equilibria

For game Γ, let $|\Gamma|$ denote its representation size and $m := \sum_{i \in \mathcal{N}} \sum_{I \in \mathcal{I}^{(i)}} |A_I|$ its the total number of pure actions.

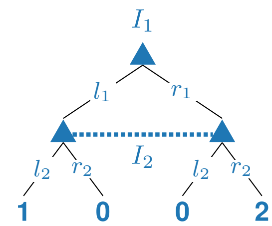


Figure 3: A single-player game with imperfect recall where miscoordinating actions with yourself is punished most.

Proposition 9. NASH-D *is solvable in time* $\text{poly}\left(|\Gamma|, \log\frac{1}{\epsilon}, (m \cdot |\mathcal{H}|)^{m^2}\right)$.

In fact, our algorithm *finds* an ϵ-Nash equilibrium whenever an exact Nash equilibrium exists. The idea is similar to that one of Lipton and Markakis [2004][Theorem 2] for multi-player normal-form games: Namely, we iteratively subdivide the strategy space, and repeatedly decide with first-order-of-the-reals solvers whether a Nash equilibrium exists in this smaller region. Those solvers also give rise to the exponential time dependence on m. In particular, the algorithm becomes polytime if m is bounded by a constant. This observation will aid us towards a PLS-membership proof in Theorem 5. Also note that such a bound on m will not restrict the size of the game tree since the degree of absentmindedness can still grow arbitrarily (cf. Figure 1b).

4 Introducing Multiselves Equilibria

Section 3 shows strong obstacles to finding Nash equilibria in games with imperfect recall. In light of these limitations, we relax the space of solutions and turn to the *multiselves* approach (cf. the agent-form [Kuhn, 1953]), which we review in this section. This approach argues that, whenever a player finds herself in an infoset, she has no influence over which actions she chooses at other infosets. Therefore, at a multiselves equilibrium μ, each player will play the best randomized action at each of their infosets, assuming that they themselves play according to μ at other infosets and assuming all other players also play according to μ.

Consider Figure 3. The optimal strategy is to play (r_1, r_2). This is also a multiselves equilibrium. However, (l_1, l_2) is also a multiselves equilibrium, because if the player is at the top-level infoset I_1 and assumes that she will follow left at the bottom-level infoset I_2, then it is best for her to go left now. On the other hand, if the player is at I_2 and assumes that she played left at I_1, then it is again best for her to play left now.

Multiselves equilibria can be arbitrarily bad in payoff in comparison to optimal strategies and Nash equilibria, as can be seen by replacing the payoff of 2 in Figure 3 with some $\lambda \to \infty$. This phenomenon is due to miscoordination across infosets, and it arises in the same manner across teams in team games: The corresponding normal-form game $\begin{pmatrix} \lambda, \lambda & 0, 0 \\ 0, 0 & 1, 1 \end{pmatrix}$ shows that Nash equilibria can be arbitrarily worse relative to Pareto-optimal profiles.

In games with absentmindedness it becomes controversial how to apply the multiselves idea. Specifically, how should a player reason about implications of a choice at the current decision point for her action choices at past and future decision points *within the same infoset*, and – as a consequence – compute incentives to deviate? That is, in considering deviating, will the player assume they would perform the same deviation at other nodes in the same infoset, or that the deviation is a one-time-only event? We will handle this question using two well-motivated[2] decision theories that correspond to these two cases: evidential decision theory and causal decision theory. We will see that Nash equilibria are multiselves equilibria under both decision theories.

That this section is accompanied with an extensive section in the full version of this paper that – beyond proving the statements made in this section – also introduces some additional observations and lemmas needed for the development of our main results.

4.1 Evidential Decision Theory (EDT)

Suppose a game Γ is played with profile μ, and a player i arrives in one of her infosets $I \in \mathcal{I}^{(i)}$. EDT postulates that if that player deviates to a randomized action $\alpha \in \Delta(A_I)$ at the current node, then she will have also deviated to α whenever she arrived in I in the past, and that she will also deviate to α whenever she arrives in I again in the future. This is because EDT argues that the choice to play α now is evidence for the player playing the same α in the past and future.

We denote the behavioral strategy that results from an EDT deviation as $\mu^{(i)}_{I \mapsto \alpha}$. It plays according to $\mu^{(i)}$ at every infoset except for at $I \in \mathcal{I}^{(i)}$ where it plays according to $\alpha \in \Delta(A_I)$.

Definition 10. *We call μ an* EDT equilibrium *for game Γ if for all players $i \in \mathcal{N}$, all her infosets $I \in \mathcal{I}^{(i)}$, and all randomized actions $\alpha \in \Delta(A_I)$, we have*

$$U^{(i)}(\mu) \geq U^{(i)}(\mu^{(i)}_{I \mapsto \alpha}, \mu^{(-i)}).$$

In an EDT equilibrium, no player has an incentive to deviate at an infoset in an EDT fashion to another randomized action. This is because the right hand side of the inequality represents the expected *ex-ante* utility of such an EDT deviation. We give an extensive discussion on the ex-ante perspective for multiselves equilibria in the full version of this paper. Regarding equilibrium computation, the following result is known:

Proposition 11 (Tewolde *et al.*, 2023)**.** *Unless* NP = ZPP*, finding an ϵ-EDT equilibrium in a single-player game for* 1/poly *precision is not in* P.

4.2 Causal Decision Theory (CDT)

Say, again, game Γ is played with profile μ, and a player i arrives in one of her infosets $I \in \mathcal{I}^{(i)}$. Then CDT postulates that the player can deviate to an action $\alpha \in \Delta(A_I)$ at the current node without violating that she has been playing according to $\mu^{(i)}$ at past arrivals in I, or that she will be playing according to $\mu^{(i)}$ at future arrivals in I. This is in addition to assuming that all other players follow $\mu^{(-i)}$ as usual. The intuition behind CDT is that the player's choice to deviate from $\mu^{(i)}$ at the current node does not *cause* any change in her behavior at any other node of the same infoset I.

Example 12. *Recall Figure 1b in which – as the story goes – the absentminded driver has to exit a highway at the second highway exit to find home. Say the player enters the game with μ = 'e' (exit), and upon arriving in the infoset, considers deviating to 'c' (continue) at this point of time. EDT then argues that the player will always continue on the highway and arrive at the third "0" payoff of the game. CDT, on the other hand, argues that the player will continue on the highway once – or more precisely, continue at the root node since that is the only decision node she could possibly be at given her strategy μ – and then exit the highway at its second exit.*

For node $h \in \mathcal{H}^{(i)}$ and pure action $a \in A_h$, let ha denote the child node reached if player i plays a at h. Consequently, $U^{(i)}(\mu \mid ha)$ is the expected utility of player i from being at h, playing a, and everyone following profile μ afterwards. When at an infoset $I \in \mathcal{I}^{(i)}$, the player does not know at which node of I she currently is. Therefore, when computing her utility incentives for a CDT-style deviation to a, she scales each node by the probability of reaching that node under profile μ. This yields utility incentives

$$\textstyle\sum_{h \in I} \mathbb{P}(h \mid \mu) \cdot U^{(i)}(\mu \mid ha).$$

to CDT-deviate to pure action a at infoset I. This value is known to be equal to the partial derivative $\nabla_{I,a} U^{(i)}(\mu)$ of utility function $U^{(i)}$ w.r.t. to action a of $I \in \mathcal{I}^{(i)}$ at point μ [Piccione and Rubinstein, 1997; Oesterheld and Conitzer, 2022]. Hence, we can formulate the following definition.

Definition 13. *Given a profile μ in game Γ, a player $i \in \mathcal{N}$ determines her (ex-ante) utility from CDT-deviating at infoset $I \in \mathcal{I}^{(i)}$ to randomized action $\alpha \in \Delta(A_I)$ as*

$$U^{(i)}_{\text{CDT}}(\alpha \mid \mu, I) := U^{(i)}(\mu) + \textstyle\sum_{a \in A_I} (\alpha(a) - \mu(a \mid I)) \cdot \nabla_{I,a} U^{(i)}(\mu).$$

In other words, this is the first-order Taylor approximation of $U^{(i)}$ at μ in the subspace $\Delta(A_I)$. In the full version of this paper, we illustrate on a simple game that the ex-ante CDT-utility may yield unreasonable utility payoffs for values α far away from $\mu(\cdot \mid I)$. Moreover, if $\alpha \neq \mu(\cdot \mid I)$, we observe that the resulting behavior of the deviating player cannot be captured by a *behavioral strategy* anymore that the player could have chosen from the beginning. That is because the player is now acting differently at different nodes of the same infoset.

Definition 14. *A profile μ is said to be a* CDT equilibrium *for game Γ if for all player $i \in \mathcal{N}$, all her infosets $I \in \mathcal{I}^{(i)}$, and all alternative randomized actions $\alpha \in \Delta(A_I)$, we have*

$$U^{(i)}(\mu) = U^{(i)}_{\text{CDT}}\big(\mu^{(i)}(\cdot \mid I) \bigm| \mu, I\big) \geq U^{(i)}_{\text{CDT}}(\alpha \mid \mu, I).$$

Therefore, no player has any utility incentives to deviate at an infoset in a CDT fashion to another randomized action.

[2] The debate around decision theories is related to the approach for belief formation (cf. the Sleeping Beauty problem [Elga, 2000]). Among other aspects, the literature has studied which combination of decision theories and belief formation avoid being Dutch-booked (money-pumped) [Piccione and Rubinstein, 1997; Briggs, 2010; Oesterheld and Conitzer, 2022].

CDT equilibria have received a more thorough treatment in the literature than EDT equilibria have.

Lemma 15 (Lambert *et al.*, 2019)**.** *Any game* Γ *with imperfect recall admits a CDT equilibrium.*

Thus, we shall define CDT-S as the search problem that asks for an ϵ-CDT equilibrium in the game (which always exists). Let 1P-CDT-S be its restriction to single-player games.

Proposition 16 (Tewolde *et al.*, 2023)**.**

1. *A profile* μ *is a CDT equilibrium of* Γ *if and only if for all player* $i \in \mathcal{N}$*, strategy* $\mu^{(i)}$ *is a KKT-point of* $\max_{\pi^{(i)} \in S^{(i)}} U^{(i)}(\pi^{(i)}, \mu^{(-i)})$.
2. *The problem* 1P-CDT-S *is* CLS*-complete.*

The original formulation of Tewolde *et al.* was not given for the multi-player setting and the ex-ante perspective. We compare it to the above formulation in the full version of this paper. Furthermore, we may also highlight a positive algorithmic implication which has not been stated before. It can be obtained analogously to [Fearnley *et al.*, 2023, Lemma C.4].

Corollary 17. 1P-CDT-S *for* 1/poly *precision is in* P.

4.3 Comparing the Solution Concepts

The three solution concepts form an inclusion hierarchy. This result is known for single-player settings and extends straightforwardly to multi-player settings.

Proposition 18 (Oesterheld and Conitzer, 2022)**.** *A Nash equilibrium is an EDT equilibrium. An EDT equilibrium is a CDT equilibrium.*

This also implies that any single-player game admits both EDT and CDT equilibria since it admits an optimal strategy (= Nash equilibrium). In general, neither statement in Proposition 18 holds in reverse. Indeed, we have seen in Figure 3 that multiselves equilibria may not be the optimal strategy. Moreover, the strategy μ described in Example 12 forms a CDT equilibrium but not an EDT equilibrium (an EDT deviation to a uniformly randomized action achieves positive utility).

We will find in this paper that CDT equilibria are easier to compute than EDT equilibria. Indeed, Proposition 11 and Corollary 17 already serve as the first evidence towards such a separation. We can also find a hint towards such an insight by considering the easier problem of *verifying* whether a given profile could be an equilibrium. For CDT, this can be done in polytime: since $U^{(i)}_{\text{CDT}}$ is linear in α, we do not actually need to check Definition 14 for all $\alpha \in \Delta(A_I)$, but it suffices to only check it for *pure* actions $a \in A_I$. For EDT equilibria, on the other hand, there is no simple-to-check characterization: $U^{(i)}(\mu^{(i)}_{I \mapsto \cdot}, \mu^{(-i)})$ is a polynomial function over $\Delta(A_I)$, for which no easy verification method is known. At least, this is true in general. As for special cases, we have:

Remark 19. *Without absentmindedness, deviation incentives of EDT and of CDT coincide, and so do the equilibrium concepts. Hence, complexity results such as Proposition 16 and Theorem 6 will apply to EDT equilibria as well.*

Remark 20. *If each player has only one infoset in total, then the EDT equilibria coincide with the Nash equilibria.*

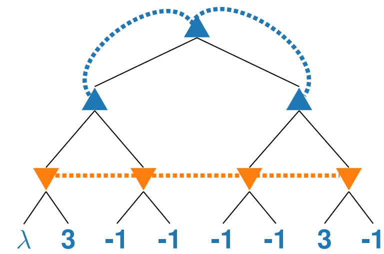


Figure 4: A variant of Figure 1a where P1 has one single infoset with absentmindedness. It is parametrized by the payoff $\lambda \in \mathbb{R}$ from P1 shooting left and P2 blocking left.

5 Complexities of Multiselves Equilibria

In this section, we present our computational results for multiselves equilibria.

5.1 EDT Equilibria

Consider the (parametrized) *absentminded* penalty shoot-out in Figure 4. It shows that in multi-player settings, EDT equilibria may not exist. Absentmindedness is crucial for such an example due to Remark 19 and Lemma 15.

Lemma 21. *Figure 4 has an EDT equilibrium if and only if* $\lambda \geq 3$.

The next result establishes $\exists\mathbb{R}$-hardness again by similar arguments to Theorem 1. Except in this construction, we attach the single-player game Γ from Proposition 3 to the bottom left of Figure 4. Note here that by an appropriate payoff shift in Γ, we can w.l.o.g. assume the target t for Γ to be 3.

Theorem 3. *Deciding whether a game with imperfect recall admits an EDT equilibrium is* $\exists\mathbb{R}$*-hard and in* $\exists\forall\mathbb{R}$*. Hardness holds even for 2p0s games where on player has a degree of absentmindedness of* 4 *and the other player has perfect recall.*

Now consider problem EDT-D that asks to distinguish between whether an exact EDT equilibrium exists or whether no ϵ-EDT equilibrium exists.

Theorem 4. EDT-D *is* Σ_2^{P}*-complete. Hardness holds for* 1/poly *precision and 2p0s games with one infoset per player and a degree of absentmindedness of* 4*.*

The technically involved proof casts the game construction for Theorem 1 to a game where each player only has one infoset, in order to use Remark 20. For that, we cannot reduce from Lemma 6 this time, but we reduce directly from the Σ_2^{P}-complete problem $\exists\forall$3-DNF-SAT. Moreover, we make use of the flexibility that EDT-utilities can represent arbitrary polynomial functions as long as they are only over a single simplex.

Next, we turn to the search problem. The algorithm of Proposition 9 can also find ϵ-EDT equilibria if we adjust for its equilibrium conditions. In single-player settings, however, we can do better since EDT equilibria are guaranteed to exist. Let 1P-EDT-S be the search problem that asks for an ϵ-EDT equilibrium. This problem was left open by Tewolde *et al.* [2023].

Theorem 5. 1P-EDT-S *is* PLS*-complete when the branching factor is constant. Hardness holds even when the branching factor and the degree of absentmindedness are* 2*.*

Before we touch on the proof idea, we shall highlight the contrast to the CLS-membership result for 1P-CDT-S in Proposition 16. CLS is believed to be a proper subset of PLS evidenced by conditional separations; see the full version of this paper.

Corollary 22. 1P-EDT-S *for* 1/poly *precision is in* P *when the branching factor is constant.*

The proofs first establish that 1P-EDT-S is computationally equivalent to the search problem that takes a polynomial function p over a product of simplices, and asks for an approximate "Nash equilibrium point" of it. In the special case where the branching factor is 2, the domain becomes the hypercube $[0,1]^\ell$, and an ϵ-Nash equilibrium x is characterized by the property

$$\forall j \in [\ell]\, \forall y \in [0,1]: \quad p(x) \geq p(y, x_{-j}) - \epsilon.$$

We show that this problem is PLS-complete. This result may be of independent interest for the optimization literature.

The PLS-hardness follows from a reduction from the PLS-complete problem MaxCut/Flip [Schäffer and Yannakakis, 1991; Yannakakis, 2003]. For the positive algorithmic results of PLS and P membership respectively, we show that ϵ-best-response dynamics converges to an ϵ-EDT equilibrium. We run a similar method to Proposition 9 in order to compute an ϵ-best response randomized action of an infoset to the other infosets. This takes polytime if the number of actions per infoset (= branching factor) is bounded. Without this restriction, we run into the impossibility result of Proposition 11.

5.2 CDT Equilibria

How hard is CDT-S, now that we allow for many players? We can get PPAD-hardness straightforwardly because any normal-form game can be cast to extensive form, and because finding a Nash equilibrium in a normal-form game is PPAD-complete [Daskalakis *et al.*, 2009; Chen *et al.*, 2009]. Interestingly enough, we can also show PPAD-membership.

Theorem 6. CDT-S *is* PPAD*-complete. Hardness holds even for two-player perfect-recall games with one infoset per player and for* 1/poly *precision.*

For membership we investigate the existence proof of Lemma 15 by Lambert *et al.*. They first shows a connection to perfect-recall games with particular symmetries, and then give a Brouwer fixed point argument which resembles that of Nash's for symmetric games. However, the connection relies on a construction whose size blows up in the order of factorials, *i.e.*, super-polynomially. Therefore, we modify the fixed point argument to one that works directly on CDT utilities: In a game of imperfect recall, given a profile μ, define the advantage of a pure action a at infoset I of player i as

$$g^{(i)}_{I,a}(\mu) := U^{(i)}_{\text{CDT}}(a \mid \mu, I) - U^{(i)}(\mu)\,.$$

Intuitively, if the advantage of an action a over the current randomized action $\mu^{(i)}(\cdot \mid I)$ is large, then the player should increase its probability of play. Therefore, we may define the Brouwer function to map any profile μ to profile π defined as

$$\pi^{(i)}(a \mid I) := \frac{\mu^{(i)}(a \mid I) + \max\{0, g^{(i)}_{I,a}(\mu)\}}{1 + \sum_{a' \in I} \max\{0, g^{(i)}_{I,a'}(\mu)\}}\,.$$

Then we show that this forms a valid a Brouwer function whose fixed points are indeed CDT equilibria of the underlying game, and that the Brouwer function and precision errors satisfy the computational requirements developed by Etessami and Yannakakis [2010] to imply PPAD-membership.

The PPAD-membership result is a positive algorithmic result: it shows that we can find CDT equilibria with fixed point solvers and path-following methods, just as it is the case with Nash equilibria in normal-form games. In particular, we shall highlight the stark contrast to Theorem 4. Finding a CDT equilibrium sits well within in the landscape of total NP search problems, whereas even deciding whether an EDT equilibrium exists is already on higher levels of the polynomial hierarchy, let alone finding one.

6 The Insignificance of Exogenous Stochasticity

As of now, the hardness results for single-player settings rely on the presence of chance nodes; see Propositions 3 and 4 and Theorem 5. In this section, we investigate the complexity of games without chance nodes. Of course, one might choose to add players to the game to simulate nature, even in games of perfect recall. However, adding players may add significantly to the computational complexity of the game, cf. P vs PPAD for Nash equilibria in single-player vs two-player settings under perfect recall, or Proposition 16 vs Theorem 6 for CDT equilibria under imperfect recall. Interestingly enough, we can show that in the presence of imperfect recall, chance nodes do not affect the complexity.

Theorem 7. *All computational hardness results in this paper for the three equilibrium concepts* {*Nash, EDT, CDT*} *still hold even when the game has no chance nodes. They hold together with previously possible restrictions (*e.g.*, on the branching factor), except that the restrictions on the number of infosets and the degree of absentmindedness increase by one and to $\mathcal{O}(\log |\mathcal{H}|)$ respectively.*

In other words, all exogenous stochasticity can be replaced by one infoset (of an arbitrary player, say P1) with absentmindedness, *i.e.*, replaced by uncertainty that arises from forgetting one's past actions in an identical situation. The proof first transforms the game Γ to an equivalent game $\tilde{\Gamma}$ that only has a single chance node h_c that is located at the root. Next, we replace h_c with an infoset I_c with absentmindedness. We illustrate in Figure 5 how to do it with a chance node that uniformly randomizes over two actions. The resulting game Γ' has the same number of players and strategy sets as Γ, except for the additional infoset I_c for P1. In equilibrium, the induced conditional probability distribution over the children of h_c in Γ and the nonterminal "children" of I_c in Γ' will be the same. Finally, there will be a polynomial relationship between the equilibrium precision errors in Γ and Γ'.

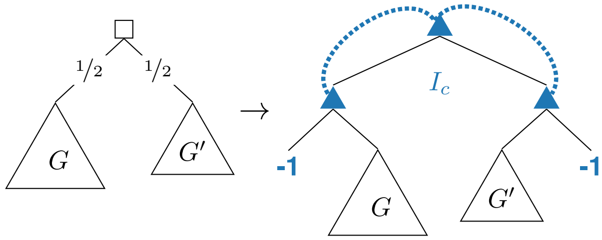


Figure 5: How to remove a chance node if it is located at the root. Starting with the game on the left, replace it with infoset I_c. Assuming w.l.o.g. that the subgames G and G' always yield positive payoffs, the player of I_c will want to randomize uniformly at I_c – independent of the play in G and G'.

Next, recall OPT-D from Proposition 4 which asks whether an approximate target value can be achieved in a single-player game with imperfect recall. We improve on Theorem 7 in the specific problem OPT-D via an independent proof.

Proposition 23. OPT-D *is* NP*-hard, even for games with no chance nodes, one infoset, a degree of absentmindedness of* 2*, and* 1/poly *precision.*

Due to Remark 20, this hardness result also holds for deciding whether *any* EDT equilibrium achieves an approximate target value. The proof reduces from the 2-MINSAT problem [Kohli *et al.*, 1994].

7 Conclusion

Historically, games of imperfect recall have received only limited attention, as it is not clear that they cleanly model any strategic interactions between humans. However, as we argued in the introduction, they are more practically significant in the context of AI agents. However, they also pose new challenges. Optimal decision making under imperfect recall is hard due to its close connections to polynomial optimization. This and previous work has shown this for the single-player setting. Moreover, it holds even more so in multi-player settings, where we established that even deciding whether a Nash equilibrium (*i.e.*, mutual best responses) exists is very hard. Therefore, we turned towards suitable relaxations that arose from the game theory and philosophy literature. We studied them, and their relationship to each other and to the Nash equilibrium concept, with a computational lens.

We find that CDT equilibria stay relatively easy to find, joining the complexity class of finding a Nash equilibrium in *perfect-recall* or *normal-form* games. This is because CDT defines the most local form of deviation, affecting only one decision node at a time. EDT equilibria show a more convoluted picture. In single-player settings, we relate it to polynomial local search via best-response dynamics. Furthermore, without absentmindedness, EDT and CDT equilibria coincide and hence become equally easy (Remark 19). *With* absentmindedness, on the other hand, the relevant decision problems for EDT equilibria (in single- or multi-player settings) tend to coincide in complexity with the analogous problems for Nash equilibria under *imperfect* recall.

One conclusion, however, has presented itself in all settings considered throughout this paper: (assuming well-accepted complexity assumptions), CDT equilibria are in general strictly easier to find and decide than EDT and Nash equilibria (Proposition 16 vs Theorem 5, Corollary 17 vs Proposition 11, and Theorem 6 vs Theorem 4). Does this imply that CDT-based reasoning is more suitable for computationally-bounded agents?

Finally, the computational differences between EDT equilibria and Nash equilibria have yet to be properly understood, that is, the differences between global optimization of polynomials over a single simplex versus a product of simplices. We leave this open for future work, with a particular interest in the search complexities of these two equilibrium concepts.

Acknowledgments

We are grateful to Ioannis Anagnostides for the fruitful discussions during the development of this project, and to the anonymous reviewers for their valuable improvement suggestions for this paper. Emanuel Tewolde, Caspar Oesterheld, and Vincent Conitzer thank the Cooperative AI Foundation, Polaris Ventures (formerly the Center for Emerging Risk Research) and Jaan Tallinn's donor-advised fund at Founders Pledge for financial support. Caspar Oesterheld's work is also supported by an FLI PhD Fellowship. Paul Goldberg is supported by EPSRC Grant EP/X040461/1. The work of Prof. Sandholm's group is supported by the Vannevar Bush Faculty Fellowship ONR N00014-23-1-2876, National Science Foundation grants RI-2312342 and RI-1901403, ARO award W911NF2210266, and NIH award A240108S001. Brian Hu Zhang's work is supported in part by the CMU Computer Science Department Hans Berliner PhD Student Fellowship.

References

Sanjeev Arora and Boaz Barak. *Computational Complexity - A Modern Approach*. Cambridge University Press, 2009.

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.

Noam Brown and Tuomas Sandholm. Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science*, 359(6374):418–424, 2018.

Noam Brown and Tuomas Sandholm. Superhuman AI for multiplayer poker. *Science*, 365(6456):885–890, 2019.

Noam Brown, Sam Ganzfried, and Tuomas Sandholm. Hierarchical abstraction, distributed equilibrium computation, and post-processing, with application to a champion no-limit Texas hold'em agent. In Gerhard Weiss, Pinar Yolum, Rafael H. Bordini, and Edith Elkind, editors, *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2015, Istanbul, Turkey, May 4-8, 2015*, pages 7–15. ACM, 2015.

Joshua Buresh-Oppenheim and Tsuyoshi Morioka. Relativized NP search problems and propositional proof systems. In *Proceedings of the 19th IEEE Conference on Computational Complexity (CCC)*, pages 54–67, 2004.

Samuel R. Buss and Alan S. Johnson. Propositional proofs and reductions between NP search problems. *Annals of Pure and Applied Logic*, 163(9):1163–1182, 2012.

John Canny. Some algebraic and geometric computations in PSPACE. In *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, STOC '88, pages 460—-467, New York, NY, USA, 1988. Association for Computing Machinery.

Andrea Celli and Nicola Gatti. Computational results for extensive-form adversarial team games. In Sheila A. McIlraith and Kilian Q. Weinberger, editors, *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018*, pages 965–972. AAAI Press, 2018.

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.

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

Vincent Conitzer. Designing preferences, beliefs, and identities for artificial intelligence. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence*, pages 9755–9759, Honolulu, HI, USA, 2019.

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

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

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

Scott Emmons, Caspar Oesterheld, Andrew Critch, Vincent Conitzer, and Stuart Russell. For learning in symmetric teams, local optima are global Nash equilibria. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvári, Gang Niu, and Sivan Sabato, editors, *International Conference on Machine Learning, ICML 2022, 17-23 July 2022, Baltimore, Maryland, USA*, volume 162 of *Proceedings of Machine Learning Research*, pages 5924–5943. PMLR, 2022.

Kousha Etessami and Mihalis Yannakakis. On the complexity of Nash equilibria and other fixed points. *SIAM J. Comput.*, 39(6):2531–2597, 2010.

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

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

Sam Ganzfried and Tuomas Sandholm. Potential-aware imperfect-recall abstraction with earth mover's distance in imperfect-information games. In Carla E. Brodley and Peter Stone, editors, *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27–31, 2014, Québec City, Québec, Canada*, pages 682–690. AAAI Press, 2014.

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, pages 456–464, Richland, SC, 2020. International Foundation for Autonomous Agents and Multiagent Systems.

Johan Håstad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001.

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.

Rajeev Kohli, Ramesh Krishnamurti, and Prakash Mirchandani. The minimum satisfiability problem. *SIAM J. Discret. Math.*, 7(2):275–283, 1994.

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.

Vojtech Kovarík, Caspar Oesterheld, and Vincent Conitzer. Game theory with simulation of other players. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI 2023, 19th-25th August 2023, Macao, SAR, China*, pages 2800–2807. ijcai.org, 2023.

Vojtech Kovarík, Caspar Oesterheld, and Vincent Conitzer. Recursive joint simulation in games. *CoRR*, abs/2402.08128, 2024.

H. W. Kuhn. Extensive games and the problem of information. In Harold William Kuhn and Albert William Tucker, editors, *Contributions to the Theory of Games (AM-28), Volume II*, chapter 11, pages 193–216. Princeton University Press, Princeton, 1953.

Nicolas S. Lambert, Adrian Marple, and Yoav Shoham. On equilibria in games with imperfect recall. *Games Econ. Behav.*, 113:164–185, 2019.

Richard J. Lipton and Evangelos Markakis. Nash equilibria via polynomial equations. In Martin Farach-Colton, editor, *LATIN 2004: Theoretical Informatics, 6th Latin American Symposium, Buenos Aires, Argentina, April 5-8, 2004, Proceedings*, volume 2976 of *Lecture Notes in Computer Science*, pages 413–422. Springer, 2004.

Tsuyoshi Morioka. Classification of search problems and their definability in bounded arithmetic. Master's thesis, University of Toronto, 2001.

John F. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1):48–49, 1950.

John Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):286–295, 1951.

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.

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.

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

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.

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

Alejandro A. Schäffer and Mihalis Yannakakis. Simple local search problems that are hard to solve. *SIAM J. Comput.*, 20(1):56–87, 1991.

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.

R. H. Strotz. Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, 23(3):165–180, 1955.

Emanuel Tewolde and Vincent Conitzer. Game transformations that preserve nash equilibria or best-response sets. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI 2024, Jeju, South Korea, 3-9 August 2024*. ijcai.org, 2024.

Emanuel Tewolde, Caspar Oesterheld, Vincent Conitzer, and Paul W. Goldberg. The computational complexity of single-player imperfect-recall games. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence, IJCAI 2023, 19th–25th August 2023, Macao, SAR, China*, pages 2878–2887. ijcai.org, 2023.

Bernhard von Stengel and Daphne Koller. Team-maxmin equilibria. *Games and Economic Behavior*, 21(1-2):309–321, 1997.

Kevin Waugh, Martin Zinkevich, Michael Johanson, Morgan Kan, David Schnizlein, and Michael H. Bowling. A practical use of imperfect recall. In Vadim Bulitko and J. Christopher Beck, editors, *Eighth Symposium on Abstraction, Reformulation, and Approximation, SARA 2009, Lake Arrowhead, California, USA, 8-10 August 2009*. AAAI, 2009.

Philipp C. Wichardt. Existence of Nash equilibria in finite extensive form games with imperfect recall: A counterexample. *Games Econ. Behav.*, 63(1):366–369, 2008.

Mihalis Yannakakis. *Computational complexity*, pages 19–56. Princeton University Press, 2003.

Brian Hu Zhang and Tuomas Sandholm. Polynomial-time optimal equilibria with a mediator in extensive-form games. In Sanmi Koyejo, S. Mohamed, A. Agarwal, Danielle Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*, 2022.

Brian Hu Zhang, Gabriele Farina, and Tuomas Sandholm. Team belief DAG: generalizing the sequence form to team games for fast computation of correlated team max-min equilibria via regret minimization. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett, editors, *International Conference on Machine Learning, ICML 2023, 23-29 July 2023, Honolulu, Hawaii, USA*, volume 202 of *Proceedings of Machine Learning Research*, pages 40996–41018. PMLR, 2023.