Imperfect-Recall Games: Equilibrium Concepts and Their Complexity

Emanuel Tewolde
Foundations of Cooperative AI Lab
Computer Science Department
Carnegie Mellon University
Pittsburgh, USA
emanueltewolde@cmu.edu

Brian Zhang
Computer Science Department
Carnegie Mellon University
Pittsburgh, USA

Caspar Oesterheld
Foundations of Cooperative AI Lab
Computer Science Department
Carnegie Mellon University
Pittsburgh, USA

Manolis Zampetakis
Department of Computer Science
Yale University
New Haven, USA

Tuomas Sandholm
Computer Science Department
Carnegie Mellon University
Pittsburgh, USA

Paul Goldberg
Department of Computer Science
University of Oxford
Oxford, UK

Vincent Conitzer
Foundations of Cooperative AI Lab
Computer Science Department
Carnegie Mellon University
Pittsburgh, USA

ABSTRACT

We investigate optimal decision making under imperfect recall, that is, when the agent(s) knows that it will forget information it once held before. An example is the absentminded driver game, as well as team games in which the members exhibit 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 $\exists\forall\mathbb{R}$.

KEYWORDS

Equilibrium Computation, Extensive-form Games, Computational Complexity, Imperfect Recall, Team Games

ACM Reference Format:
Emanuel Tewolde, Brian Zhang, Caspar Oesterheld, Manolis Zampetakis, Tuomas Sandholm, Paul Goldberg, and Vincent Conitzer. 2024. Imperfect-Recall Games: Equilibrium Concepts and Their Complexity. In *Appears at the 6th Games, Agents, and Incentives Workshop (GAIW-24). Held as part of the Workshops at the 22st International Conference on Autonomous Agents and Multiagent Systems., Auckland, New Zealand, May 2024*, IFAAMAS, 9 pages.

Appears at the 6th Games, Agents, and Incentives Workshop (GAIW-24). Held as part of the Workshops at the 22st International Conference on Autonomous Agents and Multiagent Systems., Abramowitz, Aziz, Dickerson, Hosseini, Mattei, Obraztsova, Rabinovich, Tsang, Wąs (Chairs), May 2024, Auckland, New Zealand. © 2024 Copyright held by the owner/author(s). . . . $ACM ISBN 978-x-xxxx-xxxx-x/YY/MM

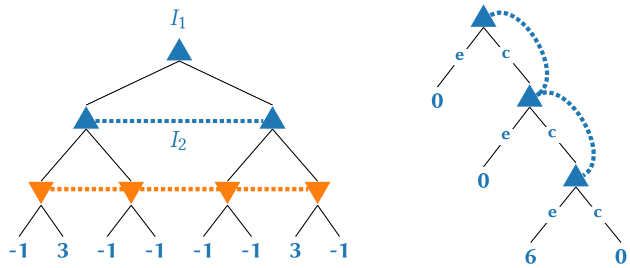
1 INTRODUCTION

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 games are already being used for state-of-the-art *abstraction* algorithms of larger games of perfect recall [2, 15, 29]. 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, imperfect-recall abstractions were a key component in the first superhuman AIs in no-limit Texas hold'em poker [3, 4].

Imperfect recall also naturally models settings in which forgetting is deliberate for other reasons, such as privacy of sensitive data [7, 33].

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 [5, 11, 28, 32]. 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 having learned yet what happened at the first node, so that effectively the



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

(b) Extended absent-minded 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.

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 another copy just knew [8].

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 [19]. In multi-player settings with *imperfect* recall, Nash equilibria may not exist anymore [30]; in fact, we show that deciding the existence is computationally hard. To give an illustrative 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. In similar reasoning to matching pennies, in a potential Nash equilibrium, none of the two players can play one side more often than the other. However, if both players mix 50/50 at each infoset, this is not a Nash equilibrium either: P1 is not best responding to P2 because she could deterministically shoot towards one side instead to avoid miscoordination with herself altogether, and 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 [10] or the absentminded driver game of Figure 1b [25]. While this literature continues to this day, from it, 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 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. Our 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 [14] and Piccione and Rubinstein [25]. In this section, we follow the exposition of Tewolde et al. [27], 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. Infoset I_2 of P1, on the other hand, 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 this example, 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 all her earlier actions. We note that any node $h \in \mathcal{H}$ uniquely corresponds to a history path $\text{hist}(h)$ in the game tree, consisting of alternating nodes and actions from

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 [16]	—	—
1/exp	NP-complete [19 27]	PLS-complete (Thm. 5*)	CLS-complete [27]
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. Indeed, there exist single-player games in which every exact EDT or CDT equilibrium involves irrational values [27], so it is not even clear how to define these search problems.

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 $\text{hist}(h)$. Then, formally, a game has perfect recall if for all player $i \in \mathcal{N}$, 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 *mixed 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) to $\sqcup_{I \in \mathcal{I}^{(i)}} A_I$[1]. 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 mixed 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:

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

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 $h \notin \text{hist}(\bar{h})$, 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 the function $U^{(i)}$ that maps a profile μ to its expected utility from game start $U^{(i)}(\mu) := U^{(i)}(\mu \mid h_0)$. In Figure 1b, that is $U^{(1)}(\mu) = 6c^2e$.

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.

Any collection of N multivariate polynomials $p^{(i)} : \bigtimes_{i=1}^N \bigtimes_{j=1}^{\ell^{(i)}} \mathbb{R}^{m_j^{(i)}} \to \mathbb{R}$ is representable as an N-player game Γ with imperfect recall such that its expected utility functions satisfy $U^{(i)}(\mu) = p^{(i)}(\mu)$ on $\bigtimes_{i=1}^N \bigtimes_{j=1}^{\ell^{(i)}} \mathbb{R}^{m_j^{(i)}}$. This can be found in the appendix.

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 and some (utility) 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 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. These 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. 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 $\epsilon > 0$.

Remark 2. *By default, $\epsilon > 0$ will be given in binary, in which case we require* inverse-exponential *(1/exp) precision.*

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

[1] *Mixed* strategies will not be the focus of this paper. A mixed strategy is a probability distribution over all pure strategies. In the presence of imperfect recall, mixed strategies are not realization-equivalent to behavioral strategies [20]. Mixed strategies require the agent to coordinate her actions across infosets (*e.g.*, access to a correlation device); as this would imply a form of memory, it does not fit the motivation of this paper.

Complexity Classes. We now give short descriptions of the complexity classes, and refer to the appendix for details and references. The classes are listed in roughly increasing order of hardness. P: Decision problems solvable in polytime. CLS: Search problems in both PLS and PPAD. PLS: Search problems expressible as local optimization over a combinatorial search space. PPAD: Search problems expressible as fixed-point problems, *e.g.*, Nash equilibria. NP: Decision problems where solutions can be verified in polytime. Σ_2^{P}: Decision problems where solutions can be verified in polytime with access to a SAT oracle. $\exists\mathbb{R}$ and $\exists\forall\mathbb{R}$ are analogues of NP and Σ_2^{P} respectively for real-valued variables.

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 [23].

Definition 3. *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 Section 1. If there is only one single player, however, finding a Nash equilibrium 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:

Proposition 4 (16). *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 5 (19; 27). Opt-D *is* NP*-complete.*

Technically, Koller and Megiddo establish hardness for the *exact* decision problem, but their proof also implies – via the PCP theorem [17] – NP-hardness for absolute constant $\epsilon < 1/8$.

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 can equivalently minimize that utility.

Koller and Megiddo [19] 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 are interested in behavioral strategies instead:

Definition 6. *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. [16] prove that deciding $\underline{U} \geq t$ is in $\exists\forall\mathbb{R}$ and is hard for $\exists\mathbb{R}$. For approximation, we have:

Lemma 7 (32). *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.*

For 2p0s games Γ, there is a tight connection between the existence of Nash equilibria and the 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 8. *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.

3.2 Deciding Nash Equilibrium Existence

We find the following for multi-player settings.

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.*

We give a high-level description of the hardness reduction via Proposition 4: Its game instances Γ will be played by P1 and will form, together with the game of Figure 1a, a subgame $\tilde{\Gamma}$ as illustrated in the appendix. Then, a Nash equilibrium exists if and only if $\tilde{\Gamma}$ will never be reached (by P2) if and only if the target value can be achieved in Γ. Similar reasoning works for approximation. Consider problem Nash-D that asks 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.*

Corollary 9. *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.*

In a later section, we will show hardness for deciding existence of EDT equilibria. The results in that section will also imply hardness for Nash-D, with different restrictions.

3.3 A Naïve Algorithm

We now give a naïve algorithm of the problem Nash-D. For game Γ, let $|\Gamma|$ denote its representation size, N its number of players, L its maximum number of infosets per player, and M its maximum number of actions per infoset.

Proposition 10. Nash-D *is solvable in $\left(|\Gamma| + \log \frac{1}{\epsilon}\right)^{O(N^2 L^2 M^2)}$ time.*

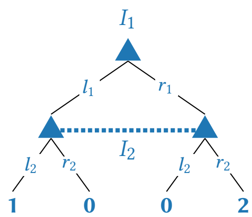


Figure 2: A single-player game with imperfect recall where miscoordinating actions with yourself is punished most.

In fact, our algorithm will *find* an ϵ-Nash equilibrium when an exact Nash equilibrium exists. The algorithm is simple, and rough bounds are provided only to highlight exponential dependencies. It will aid us in Theorem 5. The idea is to iteratively subdivide the strategy space, and to continuously decide with first-order-of-the-reals solvers whether a Nash equilibrium exists in this smaller region within exponential time. The algorithm is polytime if N, L, and M are constants. Such bounds do not restrict the size of the game tree since the degree of absentmindedness can grow arbitrarily.

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 [20]), 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 mixed 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 2. 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 best for her to play left now.

Multiselves equilibria can be arbitrarily bad in payoff in comparison to optimal strategies and Nash equilibria – an analogous phenomenon is well-known for team games, as we note in the appendix. 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.

[2]Among other aspects, the literature makes clear which decision theories (in combination with certain approaches to belief formation) avoid being Dutch-booked (money-pumped) [1, 24, 25].

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 mixed 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 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 $I \in \mathcal{I}^{(i)}$ where it plays according to $\alpha \in \Delta(A_I)$.

Definition 11. *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 mixed 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 mixed action. This is because the right hand side of the inequality represents the expected *ex-ante* utility of such an EDT deviation. We discuss the ex-ante perspective on multiselves equilibria in the appendix. The following result is known:

Proposition 12 (27). *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 take 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.

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

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

to CDT-deviate to pure action a at infoset I. This value is 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 [24, 25].

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

In other words, this is the first-order Taylor approximation of $U^{(i)}$ at μ for the space $\Delta(A_I)$. In the appendix, 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 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 mixed 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 mixed action.

Lemma 15 (21). *Any game Γ with imperfect recall admits a CDT equilibrium.*

Thus, let us 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 (27).

(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) Problem 1P-CDT-S *is* CLS*-complete.*

A comparison to the original formulation of Tewolde et al. is given in the appendix, since we work in multi-player settings now and in the ex-ante perspective. We shall also note a positive algorithmic result which they do not state, but which can be obtained analogously to Fearnley et al. [13, 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.

Proposition 18 (24). *All Nash equilibria are EDT equilibria. All EDT equilibria are CDT equilibria. In general, neither statement holds in reverse.*

This also implies that any single-player game admits both EDT and CDT equilibria since it admits an optimal strategy (= Nash equilibrium).

We will find in this paper that CDT equilibria are easier to compute than EDT equilibria. Indeed, Proposition 12 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. While this is the general case, we shall highlight two special cases.

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

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

5 COMPLEXITIES OF MULTISELVES EQUILIBRIA

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

5.1 EDT Equilibria

In multi-player settings, EDT equilibria may not exist. In the appendix, we illustrate an *absentminded* penalty shoot-out that is a variant of Figure 1a with no EDT equilibrium, and additionally parameterize the bottom left payoff with $\lambda \in \mathbb{R}$.

Lemma 21. *The parameterized absentminded penalty shoot-out has an EDT equilibrium if and only if $\lambda \geq 3$.*

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.*

The idea is to reduce from Proposition 4 again, as done in Theorem 1. However, this time we attach the single-player game to the bottom left of the absentminded kicker game.

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 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 7 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, albeit only over one simplex.

Next, we turn to the search problem. The algorithm of Proposition 10 also finds ϵ-EDT equilibria if adjusted 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.

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.

This is in contrast to the CLS-membership of 1P-CDT-S by Proposition 16. In the appendix, we discuss CLS as a subclass of PLS that is believed to be a proper subset based on conditional separations.

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

In the proofs, we use 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 would have to satisfy

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

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 PLS-complete problem MaxCut/Flip [26, 31]. 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 10 in order to compute an ϵ-best response mixed action of an infoset to the other infosets. This takes polytime if the number of actions per infoset is bounded. Without this restriction, we run into the impossibility result of Proposition 12.

5.2 CDT Equilibria

It is well-known that finding a Nash equilibrium in a normal-form game is PPAD-complete [6, 9]. This straightforwardly gives a complexity lower bound for CDT-S since a normal-form game can be cast to extensive form. Furthermore, we can show the following.

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 PPAD-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, its construction blows up in size in the order of factorials. Therefore, we modify the fixed point argument to one that works directly on CDT utilities, and to one whose Brouwer function and precision errors satisfy the conditions given in [12] for PPAD-membership.

We highlight the stark contrast between Theorems 4 and 6. 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

Our hardness results for single-player settings (Proposition 4, Proposition 5, and Theorem 5) so far rely on the presence of chance nodes. 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.

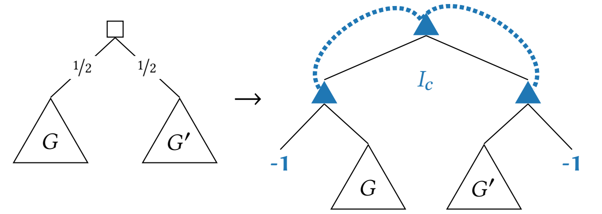


Figure 3: How to remove a chance node if it is located at the root. Starting with the game on the left, replace it with infoset I_c. Assuming w.l.o.g. that the subgames G and G' always yield positive payoffs, the player of I_c wants to randomize uniformly there, independent of the play in G and G'.

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. This is on top of previously possible restrictions* (e.g., *branching factor), except that the restrictions on the number of infosets and the degree of absentmindedness increase by one and to* $O(\log|\mathcal{H}|)$ *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 P1's own past actions in an identical situation. The idea of the proof is to first transform the game Γ to an equivalent game $\tilde{\Gamma}$ that has one chance node h_c at the root. Next, we replace h_c with an infoset I_c with absentmindedness. Figure 3 illustrates this for the case where h_c 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 will be the same in Γ and Γ'. Finally, there is a polynomial relationship between the equilibrium precision errors in the two games.

Next, recall Opt-D from Proposition 5 which asks whether an approximate target value can be achieved in a single-player game with imperfect recall. We improve Theorem 7 for Opt-D specifically with 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.*

In particular, this also holds for deciding whether *any* EDT equilibrium achieves an approximate target value. The proof reduces from the 2-MinSAT problem [18].

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 that in the single-player setting—and 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 12, 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.

REFERENCES

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

[2] Noam Brown, Sam Ganzfried, and Tuomas Sandholm. 2015. Hierarchical Abstraction, Distributed Equilibrium Computation, and Post-Processing, with Application to a Champion No-Limit Texas Hold'em Agent. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2015, Istanbul, Turkey, May 4-8, 2015*, Gerhard Weiss, Pinar Yolum, Rafael H. Bordini, and Edith Elkind (Eds.). ACM, 7–15. http://dl.acm.org/citation.cfm?id=2772885

[3] Noam Brown and Tuomas Sandholm. 2018. Superhuman AI for heads-up no-limit poker: Libratus beats top professionals. *Science* 359, 6374 (2018), 418–424. https://doi.org/10.1126/science.aao1733 arXiv:https://www.science.org/doi/pdf/10.1126/science.aao1733

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

[5] Andrea Celli and Nicola Gatti. 2018. Computational Results for Extensive-Form Adversarial Team Games. In *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*, Sheila A. McIlraith and Kilian Q. Weinberger (Eds.). AAAI Press, 965–972. https://doi.org/10.1609/AAAI.V32I1.11462

[6] Xi Chen, Xiaotie Deng, and Shang-Hua Teng. 2009. Settling the complexity of computing two-player Nash equilibria. *J. ACM* 56, 3 (2009), 14:1–14:57. https://doi.org/10.1145/1516512.1516516

[7] Vincent Conitzer. 2019. Designing Preferences, Beliefs, and Identities for Artificial Intelligence. In *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence*. Honolulu, HI, USA, 9755–9759.

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

[9] Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. 2009. The Complexity of Computing a Nash Equilibrium. *SIAM J. Comput.* 39, 1 (2009), 195–259.

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

[11] Scott Emmons, Caspar Oesterheld, Andrew Critch, Vincent Conitzer, and Stuart Russell. 2022. For Learning in Symmetric Teams, Local Optima are Global Nash Equilibria. In *International Conference on Machine Learning, ICML 2022, 17-23 July 2022, Baltimore, Maryland, USA (Proceedings of Machine Learning Research, Vol. 162)*, Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvári, Gang Niu, and Sivan Sabato (Eds.). PMLR, 5924–5943. https://proceedings.mlr.press/v162/emmons22a.html

[12] Kousha Etessami and Mihalis Yannakakis. 2010. On the Complexity of Nash Equilibria and Other Fixed Points. *SIAM J. Comput.* 39, 6 (2010), 2531–2597. https://doi.org/10.1137/080720826

[13] John Fearnley, Paul Goldberg, Alexandros Hollender, and Rahul Savani. 2023. The Complexity of Gradient Descent: CLS = PPAD ∩ PLS. *J. ACM* 70, 1 (2023), 7:1–7:74. https://doi.org/10.1145/3568163

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

[15] Sam Ganzfried and Tuomas Sandholm. 2014. Potential-Aware Imperfect-Recall Abstraction with Earth Mover's Distance in Imperfect-Information Games. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27–31, 2014, Québec City, Québec, Canada*, Carla E. Brodley and Peter Stone (Eds.). AAAI Press, 682–690. https://doi.org/10.1609/AAAI.V28I1.8816

[16] Hugo Gimbert, Soumyajit Paul, and B. Srivathsan. 2020. A Bridge between Polynomial Optimization and Games with Imperfect Recall. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems* (Auckland, New Zealand) *(AAMAS '20)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 456–464.

[17] Johan Håstad. 2001. Some optimal inapproximability results. *J. ACM* 48, 4 (2001), 798–859. https://doi.org/10.1145/502090.502098

[18] Rajeev Kohli, Ramesh Krishnamurti, and Prakash Mirchandani. 1994. The Minimum Satisfiability Problem. *SIAM J. Discret. Math.* 7, 2 (1994), 275–283. https://doi.org/10.1137/S0895480191220836

[19] Daphne Koller and Nimrod Megiddo. 1992. The complexity of two-person zero-sum games in extensive form. *Games and Economic Behavior* 4, 4 (1992), 528–552. https://doi.org/10.1016/0899-8256(92)90035-Q

[20] H. W. Kuhn. 1953. Extensive Games and the Problem of Information. In *Contributions to the Theory of Games (AM-28), Volume II*, Harold William Kuhn and Albert William Tucker (Eds.). Princeton University Press, Princeton, Chapter 11, 193–216. https://doi.org/doi:10.1515/9781400881970-012

[21] Nicolas S. Lambert, Adrian Marple, and Yoav Shoham. 2019. On equilibria in games with imperfect recall. *Games Econ. Behav.* 113 (2019), 164–185. https://doi.org/10.1016/J.GEB.2018.09.007

[22] John Nash. 1951. Non-Cooperative Games. *Annals of Mathematics* 54, 2 (1951), 286–295. http://www.jstor.org/stable/1969529

[23] John F. Nash. 1950. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences* 36, 1 (1950), 48–49. https://doi.org/10.1073/pnas.36.1.48 arXiv:https://www.pnas.org/doi/pdf/10.1073/pnas.36.1.48

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

[25] Michele Piccione and Ariel Rubinstein. 1997. On the Interpretation of Decision Problems with Imperfect Recall. *Games and Economic Behavior* 20, 1 (1997), 3–24. https://doi.org/10.1006/game.1997.0536

[26] Alejandro A. Schäffer and Mihalis Yannakakis. 1991. Simple Local Search Problems That are Hard to Solve. *SIAM J. Comput.* 20, 1 (1991), 56–87. https://doi.org/10.1137/0220004

[27] Emanuel Tewolde, Caspar Oesterheld, Vincent Conitzer, and Paul W. Goldberg. 2023. 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.* ijcai.org, 2878–2887. https://doi.org/10.24963/IJCAI.2023/321

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

[29] Kevin Waugh, Martin Zinkevich, Michael Johanson, Morgan Kan, David Schnizlein, and Michael H. Bowling. 2009. A Practical Use of Imperfect Recall. In *Eighth Symposium on Abstraction, Reformulation, and Approximation, SARA 2009, Lake Arrowhead, California, USA, 8-10 August 2009*, Vadim Bulitko and J. Christopher Beck (Eds.). AAAI. http://www.aaai.org/ocs/index.php/SARA/SARA09/paper/view/839

[30] Philipp C. Wichardt. 2008. Existence of Nash equilibria in finite extensive form games with imperfect recall: A counterexample. *Games Econ. Behav.* 63, 1 (2008), 366–369. https://doi.org/10.1016/J.GEB.2007.08.007

[31] Mihalis Yannakakis. 2003. *Computational complexity.* Princeton University Press, 19–56. http://www.jstor.org/stable/j.ctv346t9c.7

[32] Brian Hu Zhang, Gabriele Farina, and Tuomas Sandholm. 2023. Team Belief DAG: Generalizing the Sequence Form to Team Games for Fast Computation of Correlated Team Max-Min Equilibria via Regret Minimization. In *International Conference on Machine Learning, ICML 2023, 23-29 July 2023, Honolulu, Hawaii,*

USA (Proceedings of Machine Learning Research, Vol. 202), Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (Eds.). PMLR, 40996–41018. https://proceedings.mlr.press/v202/zhang23j.html

[33] Brian Hu Zhang and Tuomas Sandholm. 2022. Polynomial-Time Optimal Equilibria with a Mediator in Extensive-Form Games. In *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*, Sanmi Koyejo, S. Mohamed, A. Agarwal, Danielle Belgrave, K. Cho, and A. Oh (Eds.). http://papers.nips.cc/paper_files/paper/2022/hash/9d823334fdccb62a544fa7643cf0615d-Abstract-Conference.html