Reasoning about Causality in Games*

Lewis Hammond[1], James Fox[1], Tom Everitt[2], Ryan Carey[1],
Alessandro Abate[1], and Michael Wooldridge[1]

[1]University of Oxford
[2]DeepMind

Abstract

Causal reasoning and game-theoretic reasoning are fundamental topics in artificial intelligence, among many other disciplines: this paper is concerned with their intersection. Despite their importance, a formal framework that supports both these forms of reasoning has, until now, been lacking. We offer a solution in the form of *(structural) causal games*, which can be seen as extending Pearl's causal hierarchy to the game-theoretic domain, or as extending Koller and Milch's multi-agent influence diagrams to the causal domain. We then consider three key questions:

  i) How can the (causal) dependencies in games – either between variables, or between strategies – be modelled in a uniform, principled manner?

 ii) How may causal queries be computed in causal games, and what assumptions does this require?

iii) How do causal games compare to existing formalisms?

To address question i), we introduce *mechanised games*, which encode dependencies between agents' decision rules and the distributions governing the game. In response to question ii), we present definitions of predictions, interventions, and counterfactuals, and discuss the assumptions required for each. Regarding question iii), we describe correspondences between causal games and other formalisms, and explain how causal games can be used to answer queries that other causal or game-theoretic models do not support. Finally, we highlight possible applications of causal games, aided by an extensive open-source Python library.

Contents

Notation

Symbol	Object	Page
$\mathbf{Anc}_V$	Ancestors of V	7
$c_{\mathcal{R}}\mathcal{G}$	Condensed $\mathcal{R}$-Relevance Graph	63
$\mathbf{Ch}_V$	Children of V	7
$\mathbf{Desc}_V$	Descendants of V	7
do	Do Operator	8
$dom(V)$	Domain of V	7
E_V	Exogenous Variable for V	9
$\mathscr{E}$	Edges	7
$\mathcal{E}$	Extensive-Form Game	11
$\mathbf{Fa}_V$	Family of V	7
$\mathcal{G}$	Graph	7
$\mathcal{I}$	Intervention	9
J	Intervention Set	34
$\mathcal{M}$	Model	7
$\mathcal{M}(\zeta_k)$	Perturbed Model	33
M_V	Mechanism Variable for V	14
$\mathsf{m}\mathcal{G}$	Mechanised Graph	15
$\mathsf{m}\mathcal{M}$	Mechanised Model	15
N	Agents	11
$\mathbf{Pa}_V$	Parents of V	7
Pr or P	Probability Distribution	7, 11
$\mathrm{Pr}^{\boldsymbol{\pi}}$ or P^σ	Probability Distribution Combining Pr with $\boldsymbol{\pi}$ or P with σ	11, 12
$\mathcal{Q}$	Queries	18
$\mathcal{R}$	Rationality Relations	15
$\mathcal{R}(\mathsf{m}\mathcal{M})$	$\mathcal{R}$-Rational Outcomes of $\mathcal{M}$	16
r_D	Rationality Relation for D	15
$\mathsf{r}_{\mathcal{R}}\mathcal{G}$	$\mathcal{R}$-Relevance Graph	17
V	Variable	7
v	Value of V	7
$\boldsymbol{V}$	Variables	7
$\boldsymbol{v}$	Values of $\boldsymbol{V}$	7
δ	Kronecker Delta Function	9
Δ	Probability Simplex	7
Θ_V	Parameter Variable for V	14
$\boldsymbol{\theta}$	Parameters	7
μ^i	Mixed Policy for Agent i	29
$\boldsymbol{\pi}^i$	(Behavioural) Policy for Agent i	12
$\dot{\boldsymbol{\pi}}^i$	Pure (Behavioural) Policy for Agent i	29
$\boldsymbol{\pi}$	(Behavioural) Policy Profile	12
Π_D	Decision Rule Variable for D	14
σ	(Behavioural) Strategy Profile	11
$\perp_{\mathcal{G}}$	d-Separated in $\mathcal{G}$	7
$\perp\!\!\!\perp$	Independent	7
$\prec$	Topological Ordering	63

1 Introduction

Causal reasoning and game-theoretic reasoning are core capabilities for intelligent systems, and as such, they are fundamental research topics in artificial intelligence (AI). Causal reasoning is concerned with identifying causal relationships and estimating the effects of interventions. Game-theoretic reasoning is concerned with strategic behaviour: how rational decision-makers interact, taking into account others' incentives. Whilst formal treatments of causality [73, 75, 82, 91] and the game-theoretic foundations of multi-agent systems [71, 88, 99] have individually led to many recent applications in AI, our present concern is with techniques that *combine* causal and strategic reasoning.

Models that support both these kinds of reasoning offer a wide range of possible applications including analysing incentives [22, 57], fairness [49, 55, 67, 102], and blame and intention [28, 39]. As systems of multi-agent systems become increasingly ubiquitous and sophisticated, the problem of how to formally reason about these notions becomes increasingly acute. A framework that supports causal analysis of systems containing multiple self-interested agents would therefore appear to be of great importance [22, 28, 76].

Causal questions are often more challenging to answer in multi-agent settings; one must consider not only the causal dependencies present in the environment, but also the dependencies between agents' strategies. Similarly, reasoning strategically about what the effects of an action *would be*, or what other agents *would have done* under different circumstances, naturally leads one to consider both interventions and counterfactual possibilities when analysing games. These causal concepts, however, are typically left implicit in game-theoretic models.

The central purpose of this paper is to introduce a unifying framework for modelling games that supports both causal and game-theoretic reasoning. This framework – *(structural) causal games* – can be interpreted in two different ways. In one sense, it lifts Koller and Milch's (henceforth, K&M) multi-agent influence diagrams (MAIDs) [51] from the probabilistic models at level one of Pearl's 'causal hierarchy' [73] to causal models that support both interventions and counterfactuals, corresponding to levels two and three of the hierarchy, respectively. Building on K&M's graphical representation also means we may employ existing game-theoretic concepts such as *strategic relevance* [51] and *sufficient recall* [64], which can be elicited purely from the structure of the game. Alternatively, causal games can be interpreted as generalising the models of the causal hierarchy to the game-theoretic domain by introducing decision variables that lack a distribution until chosen by the corresponding agent, and a set of real-valued utility variables representing the payoff of each agent. Causal games thus support both game-theoretic and causal queries, as well as combinations of the two. It is our hope that this framework serves as a foundation for further work at the intersection of causality and game theory.

1.1 Contributions

In answering the three key questions introduced above, we make the following contributions.

i) How can the (causal) dependencies in games – either between variables, or between strategies – be modelled in a uniform, principled manner?

- We introduce *mechanised MAIDs* in Section 3, which allow us to cleanly model the dependencies of decision rules on each other and on the parameters of the game.

- We also generalise K&M's notion of *strategic relevance* to $\mathcal{R}$-*relevance* to enable modelling many different decision-making principles.

- Furthermore, we derive sound and complete graphical criteria for detecting relevant variables when agents are playing best responses to one another.

ii) How may causal queries be computed in such models, and what assumptions does this require?

- We generalise Pearl's causal hierarchy of models to the game-theoretic domain by introducing (structural) causal games in Section 4.

- We then describe how these models can be used to answer conditional, interventional, and counterfactual queries. By quantifying over the equilibria in the game (leading to *first order* queries such as "is it the case that in every Nash equilibrium, setting variable X to value x would increase agent i's expected payoff?") and taking into account causal effects due to rational agents who adapt their strategies to changes in the environment, such queries strictly generalise those in other causal models.

- In essence, the assumptions required to model a given problem as a causal game are the same as for their non-game-theoretic counterparts. To mechanise the causal game we further require assumptions about the decision-making principles that agents use play the game.

iii) How do the models we propose compare to existing formalisms?

- We introduce subgames and two classic equilibrium refinements – subgame perfectness and trembling hand perfectness – to MAIDs in Section 5, and provide a detailed comparison between MAIDs and EFGs (including several equivalence results) in Section 6.

- We also show that often more subgames can be found in a MAID than in the corresponding EFG, ruling out more non-credible threats, and making the computation of equilibria more efficient.

- Finally, we discuss a range of applications in which causal games subsume prior work in Section 7, and (dis)advantages with respect to other work in Section 8.

A previous conference paper contained preliminary results on subgames and equilibrium refinements in MAIDs, as well as their connections to EFGs [42], but did not contain any discussion of causality, which is the main focus of this paper. Similarly, a previous tool paper introduced an earlier version of our codebase [27], which implements many of the concepts in this paper though does not contain any theoretical work, which is the emphasis of the present paper.

We conclude this section with a review of other related work. Before the primary exposition of our results, we also provide the relevant background material on causal models, EFGs, and MAIDs, in Section 2. Proofs, further examples, and details of our codebase are relegated to Appendices A, B, and C, respectively.

1.2 Related Work

Causal games build on Pearl's hierarchy of causal models [73], and work on influence diagrams (IDs) – a kind of graphical model used to capture single-agent decision-making scenarios [45]. Later works have explicitly considered *causal* IDs, or CIDs [18, 22]. Indeed, Heckerman and Shachter's proposal to unify causal modelling and decision analysis via CIDs [43] can be viewed as a precursor to our proposal to unify causal modelling and game theory via causal games. None of these theories, however, place emphasis on games or strategic interactions between multiple agents, which is our setting of interest.

In contrast, multi-agent IDs (MAIDs) generalise IDs [51, 64], and were originally developed by K&M as a way to efficiently represent and solve games. One useful feature of MAIDs (which causal games inherit) is that their graphical structure encodes what information is, or is not, relevant for making an optimal decision. Thus in order to find an equilibrium, it is often possible to remove some of the edges (i.e. the *ignorable information*) prior to solving the game [64].

Many other graphical models of games, often inspired by MAIDs, have been introduced [48], including networks of influence diagrams (NIDs) [30], interactive dynamic influence diagrams (I-DIDs) [20, 77], and temporal action graph games [47]. Like MAIDs however, most of these works are motivated by computational concerns, and none incorporate rigorous causal reasoning.

Modelling causal relationships in game-theoretic scenarios often leads to cyclic dependencies, such as when one agent's best response depends on the other's, and vice versa. This can result in multiple solutions, a feature that is also captured by generalised or cyclic causal models [8, 38], chain graphs [59], probabilistic relational models [33, 62], and credal networks [14], among others. In mechanised (structural) causal games, these solutions correspond to different equilibria induced by (serial) *relations* representing the potentially non-deterministic decision-making processes of agents in the game (such as when an agent selects a decision rule using an arg max operation, for example). This is essential for modelling equilibria in games, where agents may be indifferent between decision rules. The result is a relational causal model with cycles – a formalism first explored concurrently with this work [1], but without any reference to the game-theoretic scenarios that are our focus.

Perhaps the most similar work to this paper is on *settable systems*,[1] which are partly inspired by generalised structural causal models and can be used to capture optimisation, equilibria, and learning [95, 96]. In order to deal with cycles, settable systems duplicate intervenable variables into a 'response' and 'setting' variable, so as to indicate which side of the structural function each occurs on. In these models, the emphasis is on the causal analysis of optimisation procedures at a relatively low level of abstraction, meaning that the algorithms used by agents to select actions, or the procedures used to select one equilibrium from many, are explicitly instantiated. In contrast, causal games represent the causal dependencies arising from the fact that agents select their decision rules rationally and non-deterministically, leading us to ask *first-order* causal queries. Settable systems concentrate on problems such as how to capture the data and attributes of a machine learning process, whereas we focus on problems such as identifying subgames and equilibrium refinements. Despite these differences, settable systems are nonetheless a useful comparator for causal games.

The concurrency and semantics community is another that has produced work at the intersection of games and causality. Much of the most influential recent work on the foundations of denotational semantics uses distributed games that are based on *event structures* – a partially ordered model of discrete events [69] – for deterministic [81] and probabilistic [98] concurrent games. Other related approaches to concurrent games use simpler mathematical formalisms [11, 101]. Though containing the same primitive concepts, these works are motivated primarily by the problem of deriving formal, low-level semantics for programs or probabilistic systems, whereas we are interested in more high-level models of strategic interactions that can be applied to a wide range of scenarios and disciplines.

Indeed, closely related causal models have been used (often in the context of analysing AI systems) to define notions of blame and intent (both in the single-agent and multi-agent settings) [28, 39], harm [80], incentives to control or respond to certain variables [22, 57], [49, 55, 67, 102], social influence [46], and reasoning patterns such as manipulation and signalling [76]. We show in Sections 4 and 7 that causal games subsume the models in these works and allow for even richer concepts. Aside from analysing AI systems, one other relevant domain of application is in economic analysis and mechanism design. For example, Toulis and Parkes use a behavioural causal model to determine long-term effects of policy interventions on multi-agent economies [92] – though their emphasis is on dynamical systems and behavioural analysis – and some

---

[1]We note that Gonzalez-Soto et al. also define a kind of causal game [35], though theirs is essentially a Bayesian game where the outcomes are computed using an unknown causal model. This is a simpler and more restricted setup than both settable systems and our causal games.

regulators such as the UK's Financial Conduct Authority include an informal 'causal chain' in their cost-benefit analyses when proposing policy analyses [25]. We provide a more formal case study of this second example in Section 7.

2 Background

We assume a basic familiarity with both probabilistic graphical models and game theory, though for completeness, we briefly review Pearl's causal hierarchy [73] and two game-theoretic models: extensive form games (EFGs) [54, 94] and multi-agent influence diagrams (MAIDs) [51]. Readers familiar with these models may safely skip these sections.

Throughout this paper, we use capital letters V for random variables, lowercase letters v for their instantiations, and bold letters $\boldsymbol{V}$ and $\boldsymbol{v}$ respectively for sets of variables and their instantiations. We let $dom(V)$ denote the domain of V (where by default we assume that $dom(V)$ is finite) and abuse notation somewhat by writing $dom(\boldsymbol{V}) \coloneqq \bigtimes_{V \in \boldsymbol{V}} dom(V)$. $\mathbf{Pa}_V$ denotes the parents of a variable V in a graphical representation and $\mathbf{pa}_V$ the instantiation of $\mathbf{Pa}_V$. We also define $\mathbf{Ch}_V$, $\mathbf{Anc}_V$, $\mathbf{Desc}_V$, and $\mathbf{Fa}_V \coloneqq \mathbf{Pa}_V \cup \{V\}$ as the children, ancestors, descendants, and family of V, respectively (where note that neither $\mathbf{Anc}_V$ nor $\mathbf{Desc}_V$ contain V, by convention). As with $\mathbf{pa}_V$, their instantiations are written in lowercase. We use $\Delta(dom(\boldsymbol{V}))$ to denote the set of all probability distributions over the values of $\boldsymbol{V}$, and therefore write $\Delta(dom(\boldsymbol{A}) \mid dom(\boldsymbol{B})) \coloneqq \bigtimes_{\boldsymbol{b} \in dom(\boldsymbol{B})} \Delta(dom(\boldsymbol{A}))$ to express the set of all conditional probability distributions (CPDs) over $\boldsymbol{A}$ given the values of $\boldsymbol{B}$. Unless otherwise indicated, we use superscripts to indicate an agent $i \in N = \{1, \ldots, n\}$ and subscripts to index the elements of a set; for example, the decision variables belonging to agent i are denoted $\boldsymbol{D}^i = \{D_1^i, \ldots, D_m^i\}$.

2.1 Causal Models

Pearl's *causal hierarchy* consists of three kinds of model: associational, interventional, and counterfactual [73]. Each step up the hierarchy demands stricter assumptions. The lowest level, *association*, pertains to correlations between variables that allow for *predictions* about a system. For this, it is sufficient to use observational data to construct a joint probability distribution over all of the variables in that system, which can then be represented graphically as a *Bayesian network* (BN). On the second level, we wish to reason about the effects of *interventions* – deliberate alterations made to the variables from outside the system. This requires the edges of the BN to reflect causal, not just associational, relationships, giving rise to a *causal Bayesian network* (CBN). The final level of the hierarchy is concerned with counterfactual questions – asking what would have happened had something been different, given that we made a certain observation – which corresponds to conditioning (as in associational queries), then intervening. Answering such questions requires knowledge of the underlying deterministic relationships between the variables, typically characterised using a *structural causal model* (SCM).

Association

Definition 1 (73)**.** A **Bayesian network (BN)** over a set of random variables $\boldsymbol{V}$ with joint distribution $\Pr(\boldsymbol{V})$ is a structure $\mathcal{M} = (\mathcal{G}, \boldsymbol{\theta})$ where $\mathcal{G} = (\boldsymbol{V}, \mathscr{E})$ is a directed acyclic graph (DAG) with vertices $\boldsymbol{V}$ and edges $\mathscr{E}$ that is **Markov compatible** with $\Pr$, meaning that $\Pr(\boldsymbol{v}; \boldsymbol{\theta}) = \prod_{V \in \boldsymbol{V}} \Pr(v \mid \mathbf{pa}_V; \theta_V)$. We drop the parameters $\boldsymbol{\theta} = \{\theta_V\}_{V \in \boldsymbol{V}}$ of the CPDs from our notation where unambiguous.

We can use the *d-separation* graphical criterion to identify the set of conditional independencies that any Markov compatible joint distribution over a DAG $\mathcal{G}$ must satisfy [74].

Definition 2 (73)**.** A **path** p in a DAG $\mathcal{G} = (\boldsymbol{V}, \mathscr{E})$ is a sequence of unrepeated adjacent variables in $\boldsymbol{V}$. A path p is said to be **blocked** by a set of variables $\boldsymbol{Y} \subset \boldsymbol{V}$ if and only if p

contains either:

- A *chain* $X \to W \to Z$ or $X \leftarrow W \leftarrow Z$, or a *fork* $X \leftarrow W \to Z$, and $W \in \boldsymbol{Y}$; or

- A *collider* $X \to W \leftarrow Z$ and $W \notin \mathbf{Anc_Y} \cup \{Y\}$.

For disjoint sets $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$, the set $\boldsymbol{Y}$ **d-separates** $\boldsymbol{X}$ from $\boldsymbol{Z}$, denoted $\boldsymbol{X} \perp_{\mathcal{G}} \boldsymbol{Z} \mid \boldsymbol{Y}$, if every path in $\mathcal{G}$ from a variable in $\boldsymbol{X}$ to a variable in $\boldsymbol{Z}$ is blocked by a variable in $\boldsymbol{Y}$. Otherwise, $\boldsymbol{X}$ is said to be **d-connected** to $\boldsymbol{Z}$ given $\boldsymbol{Y}$, denoted $\boldsymbol{X} \not\perp_{\mathcal{G}} \boldsymbol{Z} \mid \boldsymbol{Y}$.

For example, in the graph $\mathcal{G}$ shown in in Figure 1a, we have that $A \not\perp_{\mathcal{G}} B \mid C$ due to the active path $A \leftarrow D \to B$, but that $A \perp_{\mathcal{G}} B \mid D$, as conditioning on D blocks the connection along the aforementioned path as well as along the path $A \leftarrow C \to D \to B$.

If $\boldsymbol{X} \perp_{\mathcal{G}} \boldsymbol{Z} \mid \boldsymbol{Y}$ in $\mathcal{G}$, then $\boldsymbol{X}$ and $\boldsymbol{Z}$ are probabilistically independent conditional on $\boldsymbol{Y}$ in the sense that $\Pr(\boldsymbol{x} \mid \boldsymbol{y}, \boldsymbol{z}) = \Pr(\boldsymbol{x} \mid \boldsymbol{y})$, written $\boldsymbol{X} \perp\!\!\!\perp \boldsymbol{Z} \mid \boldsymbol{Y}$, in every distribution Pr that is Markov compatible with $\mathcal{G}$ and for which $\Pr(\boldsymbol{y}, \boldsymbol{z}) > 0$. Conversely, if $\boldsymbol{X} \not\perp_{\mathcal{G}} \boldsymbol{Z} \mid \boldsymbol{Y}$, then $\boldsymbol{X}$ and $\boldsymbol{Z}$ are dependent conditional on $\boldsymbol{Y}$ in at least one distribution Markov compatible with $\mathcal{G}$ [93].
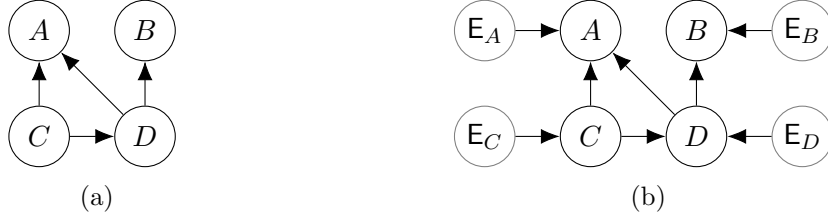

Figure 1: (a) A (C)BN with a Markov compatible joint distribution $\Pr(a, b, c, d) = \Pr(a \mid c, d) \Pr(b \mid d) \Pr(d \mid c) \Pr(c)$. (b) An SCM representing the same system. Each variable V is now associated with an exogenous parent $\mathsf{E}_V \in \mathbf{E}$.

A second well-established graphical criterion will also be a useful auxiliary result in Section 3.

Definition 3 (86)**.** Given a DAG $\mathcal{G}$, a variable V is a **requisite probability node** for $\Pr(\boldsymbol{x} \mid \boldsymbol{y})$ if there exist two parameterisations $\boldsymbol{\theta} \neq \boldsymbol{\theta}'$ of $\mathcal{G}$ for BNs $\mathcal{M}$ and $\mathcal{M}'$ differing only on θ_V such that $\Pr(\boldsymbol{x} \mid \boldsymbol{y}; \boldsymbol{\theta}) \neq \Pr(\boldsymbol{x} \mid \boldsymbol{y}; \boldsymbol{\theta}')$.

Lemma 1 (31)**.** *Given a BN $\mathcal{M}$, a variable V is a requisite probability node for the query $\Pr(\boldsymbol{X} \mid \boldsymbol{Y})$ if and only if $\mathsf{M}_V \not\perp_{\mathsf{m}_\perp \mathcal{G}} \boldsymbol{X} \mid \boldsymbol{Y}$.*

Intervention

Associational models such as BNs are, in general, insufficient to answer questions about interventions as they do not describe how the joint distribution changes in response; a causal model, or level two model, such as a causal Bayesian network (CBN), is required. The graph underlying a CBN differs from that of a BN only in its causal interpretation: the directed edges now represent the fact that intervening on a variable cannot affect those causally 'upstream' of it. The simplest form of intervention, a *hard intervention* $\mathrm{do}(\boldsymbol{Y} = \boldsymbol{y})$, sets the values of variables $\boldsymbol{Y}$ to some $\boldsymbol{y}$. We denote the resulting joint distribution by $\Pr_{\boldsymbol{y}}(\boldsymbol{V})$ or, equivalently, $\Pr(\boldsymbol{V}_{\boldsymbol{y}})$ [73].[2]

Definition 4 (73)**.** A **causal Bayesian network (CBN)** is a BN $\mathcal{M} = (\mathcal{G}, \boldsymbol{\theta})$ such that $\mathcal{G}$ is

---

[2]This is also known as an atomic [13], structural [21], surgical [73], or independent [52] intervention.

Markov compatible with $\Pr_{\boldsymbol{y}}$ for every $\boldsymbol{Y} \subseteq \boldsymbol{V}$ and $\boldsymbol{y} \in dom(\boldsymbol{Y})$, and that:

$$\Pr_{\boldsymbol{y}}(v \mid \mathbf{pa}_V) = \begin{cases} 1 & \text{when } V \in \boldsymbol{Y} \text{ and } v \text{ is consistent with } \boldsymbol{y}, \\ \Pr(v \mid \mathbf{pa}_V) & \text{when } V \notin \boldsymbol{Y} \text{ and } \mathbf{pa}_V \text{ is consistent with } \boldsymbol{y}. \end{cases}$$

By the Law of Total Probability, $\Pr_{\boldsymbol{y}}(v \mid \mathbf{pa}_V) = 0$ when v is inconsistent with $\boldsymbol{y}$. When $\mathbf{pa}_V$ is inconsistent with $\boldsymbol{y}$, then conditioning on a zero-probability event means that $\Pr_{\boldsymbol{y}}(v \mid \mathbf{pa}_V)$ is undefined. More generally, a *soft intervention*, specified using a partial distribution $\mathcal{I}$ over variables $\boldsymbol{Y}$ replaces each CPD $\Pr(Y \mid \mathbf{Pa}_Y)$ with a new CPD $\mathcal{I}(Y \mid \mathbf{Pa}_Y^*; \theta_Y^*)$ for each $Y \in \boldsymbol{Y}$, where $\mathbf{Pa}_Y^*$ may differ from $\mathbf{Pa}_Y$.[3] Any intervention $\mathcal{I}$ on the set of variables $\boldsymbol{Y}$ leads to a new joint distribution:

$$\Pr_{\mathcal{I}}(\boldsymbol{v}) \coloneqq \prod_{Y \in \boldsymbol{Y}} \mathcal{I}(y \mid \mathbf{pa}_Y^*) \cdot \prod_{V \in \boldsymbol{V} \setminus \boldsymbol{Y}} \Pr(v \mid \mathbf{pa}_V).$$

We represent an intervention $\mathcal{I}$ on $\boldsymbol{Y}$ graphically by outlining each variable $Y \in \boldsymbol{Y}$, replacing each variable name Y with $Y_{\mathcal{I}}$, and removing or adding edges from parent variables as necessary, if $\mathbf{Pa}_Y^* \neq \mathbf{Pa}_Y$. More generally, we use $\Pr(\boldsymbol{V}_{\mathcal{I}})$ and $\Pr_{\mathcal{I}}(\boldsymbol{V})$ interchangeably, and denote the new graph and model as $\mathcal{G}_{\mathcal{I}}$ and $\mathcal{M}_{\mathcal{I}}$ respectively. When $\mathcal{I}$ is a hard intervention, we simply sever all incoming edges to $\boldsymbol{Y}$, setting their values to $\boldsymbol{y}$, and write $\boldsymbol{V}_{\boldsymbol{y}}$, $\mathcal{G}_{\boldsymbol{y}}$, and $\mathcal{M}_{\boldsymbol{y}}$ respectively. Examples of hard and soft interventions are shown in Figures 2a and 2b respectively.

Figure 2: (a) The graph $\mathcal{G}_d$ representing the CBN in Figure 1a after a hard intervention $\mathrm{do}(D = d)$. (b) A similar graph $\mathcal{G}_{\mathcal{I}}$ where $\mathcal{I}(B \mid \mathbf{Pa}_B^*)$ is a soft intervention with $\mathbf{Pa}_B^* = \mathbf{Pa}_B \cup \{C\}$.

Counterfactuals

In counterfactual queries, evidence about the actual state of the world informs us about a hypothetical scenario in which some variables have been modified. For instance, we might be interested in the probability of $\boldsymbol{x}$ in the scenario in which $\boldsymbol{y}$, given that (in fact) we observed $\boldsymbol{z}$, written $\Pr(\boldsymbol{x}_{\boldsymbol{y}} \mid \boldsymbol{z})$. To answer such questions in general, one must appeal to level three of the causal hierarchy, such as by using a structural causal model (SCM). In SCMs, variables are partitioned into exogenous and endogenous sets E and $\boldsymbol{V}$ respectively, where each endogenous variable $V \in \boldsymbol{V}$ is deterministically related to its parents via a structural function $f_V : dom(\boldsymbol{V} \setminus \{V\}) \times dom(\mathsf{E}) \to dom(V)$ that specifies the mechanism governing the values of the variable, and where all stochasticity is relegated to the distribution $\Pr(\mathsf{E}; \boldsymbol{\theta})$ over the exogenous variables.

In this paper, we make the simplifying assumption that all SCMs are *Markovian*, meaning that each variable V has exactly one exogenous parent E_V and the exogenous variables are independent. We also depart from convention by describing SCMs as a particular form of CBN (which, in turn, are a particular form of BN), using deterministic distributions $\Pr(V \mid \mathbf{Pa}_V) = \delta\big(V, f_V(\mathbf{Pa}_V)\big)$ for each endogenous variable V. This equivalent formulation will prove useful for avoiding unnecessary repetition and notation when introducing causal models of games in Section 4.

---

[3]A soft intervention is also known as a parametric [21] or dependent [52] intervention, and is referred to as conditional or stochastic when deterministic or stochastic, respectively [13]. Hard interventions are a special case in which each $\mathcal{I}(Y \mid \mathbf{Pa}_Y^*) = \delta(Y, y)$ for some $y \in dom(Y)$, where δ is the Kronecker delta function.

Definition 5 ([73]). *A (Markovian)* **structural causal model (SCM)** *is a CBN* $\mathcal{M} = (\mathcal{G}, \boldsymbol{\theta})$ *where* $\mathcal{G} = (\mathbf{E} \cup \boldsymbol{V}, \mathscr{E})$ *is a DAG over exogenous variables* $\mathbf{E} = \{\mathsf{E}_V\}_{V \in \boldsymbol{V}}$ *and endogenous variables* $\boldsymbol{V}$, *where* $\mathbf{Pa}_V \cap \mathbf{E} = \{\mathsf{E}_V\}$. *The parameters* $\boldsymbol{\theta}$ *assign deterministic distributions* $\Pr(v \mid \mathbf{pa}_V; \theta_V)$ *to each endogenous variable and a stochastic distribution* $\Pr(\mathbf{E}; \boldsymbol{\theta}) = \prod_{\mathsf{E} \in \mathbf{E}} \Pr(\mathsf{E}; \theta_\mathsf{E})$ *to the exogenous variables.*

Using such a model, we can evaluate a general counterfactual query $\Pr(\boldsymbol{x}_\mathcal{I} \mid \boldsymbol{z})$ by following three steps [73]:

1. Update $\Pr(\mathbf{e})$ to $\Pr(\mathbf{e} \mid \boldsymbol{z})$ by conditioning on observation $\boldsymbol{z}$ ('abduction');

2. Apply the intervention $\mathcal{I}(\boldsymbol{Y} \mid \mathbf{Pa}_Y^*)$ to the variables $\boldsymbol{Y}$ ('action');

3. Return the marginal distribution $\Pr(\boldsymbol{x})$ in this modified model ('prediction').

By convention, exogenous variables are viewed as beyond the realm of observation and intervention. Further, we assume that each $\mathcal{I}(Y \mid \mathbf{Pa}_Y^*)$ is a deterministic function of its parents. Soft, stochastic interventions may be modelled by adding a new exogenous variable $\mathsf{E}_Y^* \in \mathbf{Pa}_Y^*$.

Note that by marginalising out the exogenous variables in an SCM, we can form a standard (C)BN with a joint distribution $\Pr(\boldsymbol{v}) = \sum_{\mathbf{e}} \Pr(\boldsymbol{v}, \mathbf{e})$. This CBN is Markov compatible with the SCM's graph $\mathcal{G}$ when restricted to variables in $\boldsymbol{V}$. Moreover, any CBN is also a BN. Thus, each model in the causal hierarchy can be seen to generate those on lower levels; every query that can be answered in a lower level model can also be answered in a higher level model.

2.2 Game-Theoretic Models

We now review two important formalisms for representing sequential strategic decision-making scenarios, extensive form games (EFGs) and multi-agent influence diagram (MAIDs). Both of these models are illustrated using the following example of a signalling game [90].

Example 1 (Job Market). *A worker, who is either hard-working or lazy, is hoping to be hired by a firm. They have the option of pursuing a university education, but know that they will suffer from three years of studying, especially if they are lazy. The firm prefers hard-workers, but is using an automated hiring system that can only observe the worker's education, not their temperament.*

Figure 3: (a) An EFG representing Example 1. (b) A MAID representing the same game.

Extensive Form Games

The material on EFGs is required only for Section 6, and so may be safely skipped or referred back to later, depending on the reader's preferences.

Definition 6 (54). An **extensive form game (EFG)** is a structure $\mathcal{E} = (N, T, P, A, \lambda, I, U)$, where:

- $N = \{1, \ldots, n\}$ is a set of agents.

- $T = (\boldsymbol{V}, \mathscr{E})$ is a game tree with nodes $\boldsymbol{V}$ that are partitioned into sets $\boldsymbol{V}^0, \boldsymbol{V}^1, \ldots, \boldsymbol{V}^n, \boldsymbol{L}$ where $R \in \boldsymbol{V}$ is the root of T, $\boldsymbol{L}$ are the leaves of T, $\boldsymbol{V}^0$ are chance nodes, and $\boldsymbol{V}^i$ are the decision nodes controlled by agent $i \in N$. The nodes are connected by edges $\mathscr{E}$.

- $P = \{P_1, \ldots, P_{|\boldsymbol{V}^0|}\}$ is a set of probability distributions $P_j(\mathbf{Ch}_{V_j^0})$ over the children of each chance node V_j^0.

- A is a set of actions, where $A_j^i \subseteq A$ denotes the set of actions available at $V_j^i \in \boldsymbol{V}^i$.

- $\lambda : \mathscr{E} \to A$ is a labelling function mapping each edge (V_j^i, V_l^k) to an action $a \in A_j^i$.

- $I = \{I^1, \ldots, I^n\}$ contains a collection of information sets $I^i \subset 2^{\boldsymbol{V}^i}$, which partition the decision nodes controlled by agent i. Each information set $I_j^i \in I^i$ is defined such that for all $V_k^i, V_l^i \in I_j^i$, the available actions $A_j^i := A_{V_k^i} = A_{V_l^i}$ are the same at both nodes.

- $U : \boldsymbol{L} \to \mathbb{R}^n$ is a utility function mapping each leaf node to a vector that determines the final payoff for each agent.

Figure 3a shows Example 1's signalling game in extensive form. Nature, as a chance node V^0, flips a biased coin at the root of the tree to decide whether the worker is hard-working h (with probability p) or lazy $\neg h$ (with probability $1 - p$). The worker's decision whether to go g or avoid $\neg g$ university is represented at nodes $V_1^1, V_2^1 \in \boldsymbol{V}^1$ and the automated hiring system's decisions are given by $V_1^2, V_2^2, V_3^2, V_4^2 \in \boldsymbol{V}^2$, each with the option to offer a job or reject the worker (j or $\neg j$). The two non-singleton information sets (marked by dotted black lines) represent the fact that the hiring system does not receive the worker's temperament as an input.

The payoffs for the worker and the firm are given by the first and second elements at the leaves of the tree respectively. The worker receives a payoff of 5 if they are given a job offer, but they incur a cost of 1 for going to university if they are hard-working and a cost of 2 for going to university if they are lazy. The firm receives a payoff of 3 if they hire a hard worker. If they offer a job to a lazy worker, they incur a cost of 2, and if they reject a hard worker, they incur an opportunity cost of 1.

Definition 7. Given an EFG $\mathcal{E} = (N, T, P, A, \lambda, I, U)$, a (behavioural) **strategy** σ^i for agent i is a set of probability distributions $\sigma_j^i(A_j^i)$ over the actions available to the agent at each of their information sets I_j^i. A strategy is **pure** when each $\sigma_j^i(a) \in \{0, 1\}$ and **fully stochastic** when $\sigma_j^i(a) > 0$, for all $a \in A_j^i$. A **strategy profile** $\sigma = (\sigma^1, \ldots, \sigma^n)$ is a tuple of strategies, and $\sigma^{-i} = (\sigma^1, \ldots, \sigma^{i-1}, \sigma^{i+1}, \ldots, \sigma^n)$ denotes the partial strategy profile of all agents other than i, hence $\sigma = (\sigma^i, \sigma^{-i})$.

Combining a strategy profile σ with the distributions in P defines a probability distribution P^σ over paths ρ in $\mathcal{E}$. For each path ρ beginning from R and terminating in a leaf node $\rho[\boldsymbol{L}] \in \boldsymbol{L}$, agent i receives utility $U(\rho[\boldsymbol{L}])[i]$ – the ith entry in the corresponding payoff vector. Agent i's expected utility under a strategy profile σ is therefore given by $\mathbb{E}_\sigma\big[U(\rho[\boldsymbol{L}])[i]\big]$.

Definition 8. A **subgame** of an EFG, $\mathcal{E} = (N, T, P, A, \lambda, I, U)$, is the game $\mathcal{E}$ restricted to a subtree $T' = (\boldsymbol{V}', \mathscr{E}')$ of T such that: for any information set I_j^i in $\mathcal{E}$, if there exists $V_k \in I_j^i \cap \boldsymbol{V}'$,

11

then $I_j^i \subseteq \boldsymbol{V}'$; and for any $V_k \in \boldsymbol{V}'$, if $(V_k, V_l) \in \mathscr{E}$, then $V_l \in \boldsymbol{V}'$ and $(V_k, V_l) \in \mathscr{E}'$.

In other words, a subtree of the original game tree forms a subgame if it is closed under information sets and descendants. Any EFG is trivially a subgame of itself, so a subgame on a strictly smaller subtree is called *proper*. In this paper, we denote subgames in EFGs by enclosing them within dashed boxes. For instance, the EFG in Figure 3a has no proper subgames, but the EFG given later in Figure 8c has two.

Multi-Agent Influence Diagrams

Influence diagrams (IDs) generalise BNs to the decision-theoretic setting by adding decision and utility variables [45, 66], and multi-agent influence diagrams (MAIDs) generalise IDs by introducing multiple agents [51, 64]. MAIDs can therefore be viewed as a BN over a graph without parameters for the decision variables, although technically lie between levels one and two of Pearl's causal hierarchy as the decisions are effectively modelled as causal interventions [43], but paths between non-decision variables need not encode causal relationships [22].

Definition 9 (51). A **multi-agent influence diagram (MAID)** is a structure $\mathcal{M} = (\mathcal{G}, \boldsymbol{\theta})$ where $\mathcal{G} = (N, \boldsymbol{V}, \mathscr{E})$ specifies a set of agents $N = \{1, \ldots, n\}$ and a DAG $(\boldsymbol{V}, \mathscr{E})$ where $\boldsymbol{V}$ is partitioned into chance variables $\boldsymbol{X}$, decision variables $\boldsymbol{D} = \bigcup_{i \in N} \boldsymbol{D}^i$, and utility variables $\boldsymbol{U} = \bigcup_{i \in N} \boldsymbol{U}^i$. The parameters $\boldsymbol{\theta} = \{\theta_V\}_{V \in \boldsymbol{V} \setminus \boldsymbol{D}}$ define the CPDs $\Pr(V \mid \mathbf{Pa}_V; \theta_V)$ for each non-decision variable such that for *any* parameterisation of the decision variable CPDs, the resulting joint distribution over $\boldsymbol{V}$ induces a BN.

Figure 3b shows a MAID representing Example 1. Chance variables, such as whether the worker's temperament is hard-working or lazy (T), are denoted by white circles. Decision and utility variables are represented using squares and diamonds respectively. The worker's decision (D^1) and utility (U^1) variables are displayed in red, and the firm's (D^2 and U^2) in blue. Instead of information sets, the fact that an agent is unaware of the value of a certain variable when making a decision is represented by a missing edge between the two (e.g., the absence of an edge $T \to D^2$). The parameters $\boldsymbol{\theta}$ define conditional distributions for variables T, U^1, and U^2, in accordance with the values shown in Figure 3a.

Definition 10. Given a MAID $\mathcal{M} = (\mathcal{G}, \boldsymbol{\theta})$, a **decision rule** π_D for $D \in \boldsymbol{D}$ is a CPD $\pi_D(D \mid \mathbf{Pa}_D)$ and a **partial policy profile** $\pi_{\boldsymbol{D}'}$ is a set of decision rules π_D for each $D \in \boldsymbol{D}' \subseteq \boldsymbol{D}$, where we write $\pi_{-\boldsymbol{D}'}$ for the set of decision rules for each $D \in \boldsymbol{D} \setminus \boldsymbol{D}'$. A (behavioural) **policy** $\boldsymbol{\pi}^i$ refers to $\boldsymbol{\pi}_{\boldsymbol{D}^i}$, and a (full, behavioural) **policy profile** $\boldsymbol{\pi} = (\boldsymbol{\pi}^1, \ldots, \boldsymbol{\pi}^n)$ is a tuple of policies, where $\boldsymbol{\pi}^{-i} := (\boldsymbol{\pi}^1, \ldots, \boldsymbol{\pi}^{i-1}, \boldsymbol{\pi}^{i+1}, \ldots, \boldsymbol{\pi}^n)$. A decision rule is **pure** if $\pi_D(d \mid \mathbf{pa}_D) \in \{0, 1\}$ and **fully stochastic** if $\pi_D(d \mid \mathbf{pa}_D) > 0$ for all $d \in dom(D)$ and each **decision context** $\mathbf{pa}_D \in dom(\mathbf{Pa}_D)$; this holds for a policy (profile) if it holds for all decision rules in the policy (profile).

By combining $\boldsymbol{\pi}$ with the partial distribution $\Pr$ over the chance and utility variables, we obtain a joint distribution:

$$\Pr^{\boldsymbol{\pi}}(\boldsymbol{x}, \boldsymbol{d}, \boldsymbol{u}) := \prod_{V \in \boldsymbol{V} \setminus \boldsymbol{D}} \Pr(v \mid \mathbf{pa}_V) \cdot \prod_{D \in \boldsymbol{D}} \pi_D(d \mid \mathbf{pa}_D),$$

over all the variables in $\mathcal{M}$; inducing a BN. The expected utility for an agent i given a policy profile $\boldsymbol{\pi}$ is defined as the expected sum of their utility variables in this BN, $\sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{\boldsymbol{\pi}}[U]$. This allows a Nash equilibrium (NE) [68] to be defined, which identifies outcomes of a game where every agent is simultaneously playing a best-response.[4]

---

[4]In Section 5.1, we build upon what's already known about NE in MAIDs, by explaining the difference between mixed policies and behavioural policies in MAIDs and clarifying when an NE is guaranteed to exist.

Definition 11 (51)**.** A policy profile $\boldsymbol{\pi}$ is a **Nash equilibrium (NE)** in a MAID if, for every agent $i \in N$, $\boldsymbol{\pi}^i \in \arg\max_{\hat{\boldsymbol{\pi}}^i \in dom(\boldsymbol{\Pi}^i)} \sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{(\hat{\boldsymbol{\pi}}^i, \boldsymbol{\pi}^{-i})}[U]$.

K&M also define *strategic relevance (s-relevance)* [51] as a concept to infer whether the choice of a decision rule can affect the optimality of another decision rule. They further show how *s*-relevance can be determined using a graphical criterion (*s-reachability*). In what follows, we use capital letters for variables Π_D and $\boldsymbol{\Theta}$ which may take different values π_D and $\boldsymbol{\theta}$. A more detailed introduction to this notation is provided in Section 3.1.

Definition 12 (51)**.** Let $D_k, D_l \in \boldsymbol{D}$ be decision nodes in a MAID $\mathcal{M}$. Π_{D_l} is **strategically relevant** (or *s-relevant*) to Π_{D_k} if there exist two joint distributions over $\boldsymbol{V}$ parameterised by $\boldsymbol{\Theta}$ and policy profiles $\boldsymbol{\pi}$ and $\boldsymbol{\pi}'$ respectively, and a decision rule π_{D_k}, such that:

- $\pi_{D_k} \in \arg\max_{\hat{\pi}_D \in dom(\Pi_D)} \sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{(\hat{\pi}_D, \boldsymbol{\pi}_{-D})}[U]$;

- $\boldsymbol{\pi}$ differs from $\boldsymbol{\pi}'$ only at Π_{D_l};

- $\pi_{D_k} \notin \arg\max_{\hat{\pi}_D \in dom(\Pi_D)} \sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{(\hat{\pi}_D, \boldsymbol{\pi}'_{-D})}[U]$, and neither does any decision rule $\tilde{\pi}_D$ that agrees with π_{D_k} on all $\mathbf{pa}_{D_k}$ such that $\Pr^{\boldsymbol{\pi}'}(\mathbf{pa}_{D_k}) > 0$.[5]

Proposition 1 (51)**.** *$\Pi_{D'}$ is s-relevant to Π_D if and only if $\Pi_{D'} \not\perp_{\mathcal{G}'} \boldsymbol{U}^i \cap \boldsymbol{Desc}_D \mid D, \boldsymbol{Pa}_D$ (Π_D is s-reachable from $\Pi_{D'}$), where $\mathcal{G}'$ is the same as $\mathcal{G}$ with an additional variable $\Pi_{D'}$ and edge $\Pi_{D'} \to D'$.*

Both *s*-relevance and *s*-reachability only consider which other *decision rules* matter under a particular assumption about the *rationality* of agents (corresponding to a notion of subgame perfectness). In Section 3, we generalise this idea to also consider the parameterisation of non-decision variables, which is key to reasoning about causality in games. We also generalise the concepts to other assumptions about agents' rationality. By considering the directed (but not necessarily acyclic) graph over all variables $\boldsymbol{\Pi_D} \coloneqq \{\Pi_D\}_{D \in \boldsymbol{D}}$ such that there is an edge $\Pi_{D'} \to \Pi_D$ if and only if $\Pi_{D'}$ is *s*-relevant to Π_D – called the (*s*-)relevance graph – K&M also introduce a weakening of the perfect recall assumption that is sufficient for the existence of an NE.

Definition 13 (51)**.** Agent i in a MAID $\mathcal{M}$ has **perfect recall** if there exists a topological ordering $D_1 \prec \cdots \prec D_m$ over $\boldsymbol{D}^i$ such that $\mathbf{Fa}_{D_j} \subset \mathbf{Pa}_{D_k}$ for any $1 \leq j < k \leq m$. $\mathcal{M}$ is said to have perfect recall if all agents in $\mathcal{M}$ have perfect recall.

Definition 14 (64)**.** Agent i in a MAID $\mathcal{M}$ has **sufficient recall** if the *s*-relevance graph restricted to just agent i's decision rules is acyclic. The MAID $\mathcal{M}$ is said to have sufficient recall if all agents have sufficient recall.[6]

Proposition 2 (51)**.** *Any MAID with sufficient recall has at least one NE in behavioural policies.*

3 Mechanised MAIDs and Relevance

Although MAIDs allow us to elegantly and concisely represent the dependencies between variables in a game, the edges in the graph only tell part of the story. Indeed, game-theoretic models are traditionally presented as objects to be *solved*, with the procedure used to produce solutions (i.e., policy profiles) left extrinsic to the game representation. However, the fact that agents act strategically means that agents' policies may be dependent on some of the parameters of the

---

[5]This final condition is included in order to ensure that π_{D_k} doesn't lead to poor decisions in decision contexts that occur with probability zero.

[6]Note that although sufficient recall is defined using the *s*-relevance graph, a similar criterion could be created for any set of rationality relations $\mathcal{R}$ and the resulting $\mathcal{R}$-relevance graph.

game (as well as other agents' policies) in ways that are crucial for causal reasoning, yet not explicitly represented by MAIDs.

In this section, we introduce *mechanised MAIDs*, which allow us to model these dependencies alongside the existing dependencies of the MAID. This representation will be fundamental for causal reasoning in games (in Section 4), as well as for the introduction of subgames in MAIDs (in Section 5.2). In the following subsections, we first define mechanised MAIDs before formally introducing the concept of *relevance* and corresponding graphical criteria.

3.1 Mechanised MAIDs

In order to explicitly capture the implicit dependencies between the decision rules and CPDs of a MAID $\mathcal{M} = (\mathcal{G}, \boldsymbol{\theta})$, a *mechanised MAID* $\mathsf{m}\mathcal{M}$ adds two elements: a (graphical) representation of the decision rules and parameters, and a description of how these depend on one another.

Mechanised Graphs

For the first addition, we extend $\mathcal{G}$ to form a *mechanised graph* $\mathsf{m}\mathcal{G}$. For each decision variable D, a new parent Π_D representing its decision rule is added, and for each non-decision variable V, a new parent Θ_V representing the parameters of its CPD. We call these additional parents *mechanism variables* $\mathbf{M}$, as they determine the mechanisms by which values of the variables in the game are set,[7] and the variables $\boldsymbol{V}$ in the original MAID *object-level variables*. To distinguish between different types of mechanism variable, we often refer to those for decisions $\boldsymbol{\Pi} = \mathbf{M}_{\boldsymbol{D}}$ as *decision rule variables* and those for non-decisions $\boldsymbol{\Theta} = \mathbf{M}_{\boldsymbol{V} \setminus \boldsymbol{D}}$ as *parameter variables*.

In the mechanised graph $\mathsf{m}\mathcal{G}$, we also add new edges $\mathscr{E}' \subseteq \bigcup_{D \in \boldsymbol{D}} \left((\mathbf{M} \setminus \Pi_D) \times \Pi_D \right)$ from other mechanism variables into decision rule variables. This represents the fact that agents typically select a decision rule π_D (i.e., the value of Π_D) based on both the parameterisation of the game (i.e., the values of $\boldsymbol{\Theta}$) and the selection of the other decision rules in the game $\boldsymbol{\pi}_{-D}$. For example, the worker and the firm's hiring system in Example 1 might want to select their decision rules π_{D^1} and π_{D^2} as a function of the probability p that a worker has a hard-working temperament $T = h$. This would imply edges $\Theta_T \to \Pi_{D^1}$ and $\Theta_T \to \Pi_{D^2}$.

In general, an agent might select each of their decision rules based on the value of any other mechanism variable – and so $\mathscr{E}'$ would be maximal (i.e., $\mathbf{Pa}_{\Pi_D} = \mathbf{M}_{\boldsymbol{V} \setminus \{D\}}$) – though typically some of these variables will not be relevant and thus $\mathscr{E}'$ will not be maximal, a notion we make precise in Section 3.2. A mechanised graph $\mathsf{m}\mathcal{G}$ for Example 1 is shown in Figure 4a, where decision rule and parameter variables are represented using black and white rounded squares respectively.

Mechanised Games

For the second addition to $\mathcal{M}$ – a description of how the decision rules and parameters depend on one another – we must specify how the values of the new mechanism variables are determined and how each of the CPDs of the original object-level variables are defined as a function of their new mechanism variable parent. Beginning with the latter, note that for any parameterisation $\mathbf{m}_{\boldsymbol{V} \setminus \boldsymbol{D}} = \boldsymbol{\theta} \in dom(\boldsymbol{\Theta})$ of the game and any policy profile $\mathbf{m}_{\boldsymbol{D}} = \boldsymbol{\pi} \in dom(\boldsymbol{\Pi})$, the resulting joint distribution over $\boldsymbol{V}$ can be written as:

$$\mathrm{Pr}^{\boldsymbol{\pi}}(\boldsymbol{v}; \boldsymbol{\theta}) = \mathrm{Pr}(\boldsymbol{v} \mid \mathbf{m}) \coloneqq \prod_{V \in \boldsymbol{V}} \mathrm{Pr}(v \mid \mathbf{pa}_V, \mathbf{m}_V),$$

---

[7]This name is partially inspired by the field of mechanism design, in which agents act by reporting their types $t \in T$ and the mechanism $f : T \to O$ determines how this type profile (which we might view as $\boldsymbol{d}$) is mapped to an outcome $o \in O$ of the game (which we might view as $\boldsymbol{v}$).

where $\Pr(v \mid \mathbf{pa}_V, \mathsf{m}_V)$ is simply the CPD $\Pr(v \mid \mathbf{pa}_V; \theta_V)$ defined by parameters $\theta_V = \mathsf{m}_V$ if V is a non-decision variable, or the decision rule $\pi_D(d \mid \mathbf{pa}_D)$ defined by $\pi_D = \mathsf{m}_V$ if V is a decision variable D. As such, given the parameters $\boldsymbol{\theta}$ of the MAID $\mathcal{M}$ and a policy profile $\boldsymbol{\pi}$, the distribution over $\boldsymbol{V}$ in $\mathsf{m}\mathcal{M}$ is identical to that in the original MAID.

Finally, we must specify how the values of the mechanism variables are determined. For the parameter variables, we simply set the distribution over each Θ_V to $\delta(\Theta_V, \theta_V)$, and let its domain be given by $\Delta(dom(V) \mid dom(\mathbf{Pa}_V))$. Intuitively, this may be viewed as the fact that any particular game induced over the graph corresponds to an instantiation of the parameter variables in the mechanised game, the values of which are known by all agents.[8]

In order to provide values for the decision rule variables, we introduce a set of *rationality relations* $\mathcal{R} = \{r_D\}_{D \in \boldsymbol{D}}$ that describe assumptions about how the agents choose decision rules. Concretely, each decision rule Π_D is governed by a *serial relation* $r_D \subseteq dom(\mathbf{Pa}_{\Pi_D}) \times dom(\Pi_D)$.[9] This accounts for the fact an agent may not deterministically choose a *single* decision rule π_D in response to some $\mathbf{pa}_{\Pi_D}$. In the remainder of the paper we abuse notation by also using $r_D(\mathbf{pa}_{\Pi_D})$ as a set-valued function to denote the *set* of all decision rules π_D such that $r_D(\mathbf{pa}_{\Pi_D}) = \pi_D$.[10]

Definition 15. Given a MAID $\mathcal{M} = (\mathcal{G}, \boldsymbol{\theta})$ and a set of rationality relations $\mathcal{R}$, a **mechanised MAID** is a structure $\mathsf{m}\mathcal{M} = (\mathsf{m}\mathcal{G}, \boldsymbol{\theta}, \mathcal{R})$, such that: the **mechanised graph** $\mathsf{m}\mathcal{G} = (N, \boldsymbol{V} \cup \boldsymbol{\mathsf{M}}, \mathsf{m}\mathscr{E})$ is a directed (possibly cyclic) graph over $\boldsymbol{V}$ and $\boldsymbol{\mathsf{M}}$ with edges $\mathsf{m}\mathscr{E} := \mathscr{E} \cup \{(\mathsf{M}_V, V)\}_{V \in \boldsymbol{V}} \cup \mathscr{E}'$, where $\mathscr{E}' \subseteq \bigcup_{D \in \boldsymbol{D}} \big((\boldsymbol{\mathsf{M}} \setminus \Pi_D) \times \Pi_D\big)$; and $\mathcal{R} = \{r_D\}_{D \in \boldsymbol{D}}$ is a set of **rationality relations**, where each $r_D \subseteq dom(\mathbf{Pa}_{\Pi_D}) \times dom(\Pi_D)$ is a serial relation.

As an example, let us suppose that each agent in Example 1 plays a *best response* with respect to the other. In this case, the values of Π_{D^1} and Π_{D^2} are determined by relations $\mathcal{R}^{\mathrm{BR}} = \{r_{D^1}^{\mathrm{BR}}, r_{D^2}^{\mathrm{BR}}\}$ such that:

$$\pi_D \in r_D^{\mathrm{BR}}(\mathbf{pa}_{\Pi_D}) \quad \Leftrightarrow \quad \pi_D \in \operatorname*{arg\,max}_{\hat{\pi}_D \in dom(\Pi_D)} \sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{(\hat{\pi}_D, \boldsymbol{\pi}_{-D})}[U], \tag{1}$$

for each $D \in \boldsymbol{D}^i$, where note that the expectation – despite the fact that the notation above includes object-level variables – is defined in terms of the *mechanism variables* $\mathbf{pa}_{\Pi_D} = \mathsf{m}_{\boldsymbol{V} \setminus \{D\}}$ and π_D. Note that for games in which each agent has only one decision rule, this gives rise to an NE, as defined in Definition 11. In other words, the values of the mechanism variables determine the CPDs of the object-level variables, which in turn define the joint distribution over the object-level variables, and hence any expected quantities in the game. Many other kinds of game-theoretic equilibria can be naturally defined in terms of rationality relations, including subgame perfect equilibria and trembling hand perfect equilibria (introduced in Section 5.3).[11]

It will often be visually useful to restrict a mechanised MAID to just the mechanism variables, as can be seen in Figure 4b. We can then view a mechanised graph as simply the composition of the original MAID and this graph of mechanism variables, via the addition of edges $\mathsf{M}_V \to V$ for each $V \in \boldsymbol{V}$.

Rational Outcomes

If all of the rationality relations $\mathcal{R}$ in a MAID are satisfied by $\boldsymbol{\pi}$, then we say that $\boldsymbol{\pi}$ is an $\mathcal{R}$-*rational outcome* of the game. For example, the $\mathcal{R}^{\mathrm{BR}}$-rational outcomes of the MAID $\mathcal{M}$

---

[8]This simply amounts to a common prior assumption. Though such an assumption can be relaxed, doing so introduces additional complexities that we do not address in this work.

[9]A serial relation $r \subseteq A \times B$ is one such that for any $a \in A$ there exists at least one $b \in B$ such that $(a, b) \in r$.

[10]Whether we refer to some $r_D(\mathbf{pa}_{\Pi_D}) \in dom(\Pi_D)$ or $r_D(\mathbf{pa}_{\Pi_D}) \subseteq dom(\Pi_D)$ will typically be unambiguous.

[11]It would also be possible to define $\mathcal{R}$ such that, for example, every agent randomises their actions at every decision, or chooses the action that minimises their expected utility. Whilst it would be hard to view such behaviour as 'rational' in the game-theoretic sense, one may think of $\mathcal{R}$ as defining the *degree(s)* of rationality in a game.
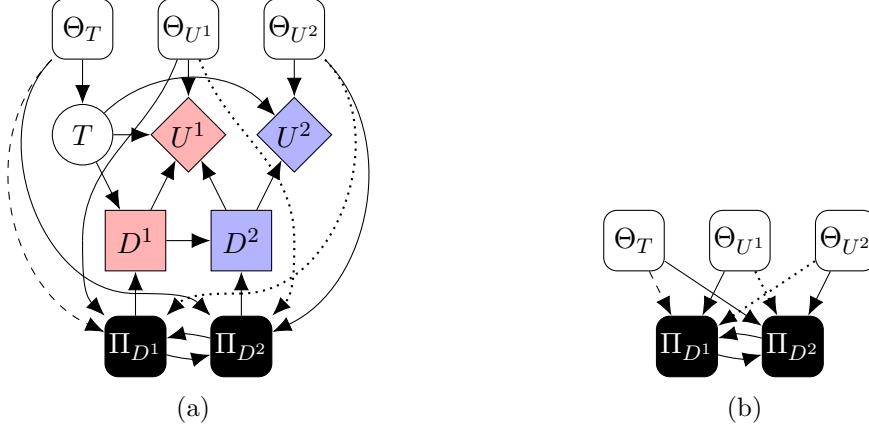
Figure 4: (a) A mechanised graph $\mathsf{m}\mathcal{G}$ representing Example 1. Dotted edges represent those that are present in neither the $\mathcal{R}^{\mathrm{BR}}$-minimal nor the s-minimal mechanised graph. Dashed edges represent just those that are not present in the s-minimal mechanised graph. (b) The $\mathcal{R}^{\mathrm{BR}}$-relevance graph is formed by removing the dotted edges, and the s-relevance graph by further removing the dashed edge.

representing Example 1 are simply the NEs of $\mathcal{M}$. Note that $\mathcal{R}$-rationality is not merely a convenience when reasoning about games, rather, it performs a necessary role by fully characterising the process by which decision rules are generated. As we will see later, the nature of the chosen rationality relations also allows us to deduce various facts about the game, purely from its graphical structure. In essence, such results hinge on the links between how agents choose their decision rules, and how the object-level variables depend on one another.

Definition 16. Given a mechanised MAID $\mathsf{m}\mathcal{M}$, we say that any $\pi_D \in r_D(\mathbf{pa}_{\Pi_D})$ is an $\mathcal{R}$-**rational response** to $\mathbf{pa}_{\Pi_D}$, and that a (partial) policy profile $\boldsymbol{\pi_{D'}}$ is $\mathcal{R}$-**rational** if $\pi_D \in r_D(\mathbf{pa}_{\Pi_D})$ for every $D \in \boldsymbol{D'}$. The set of (full) $\mathcal{R}$-rational policy profiles in $\mathsf{m}\mathcal{M}$ are the $\mathcal{R}$-**rational outcomes** of the game (where we tend to drop $\mathcal{R}$ from the terms above when unambiguous) and is denoted by $\mathcal{R}(\mathsf{m}\mathcal{M})$.

In general, the $\mathcal{R}$-rational outcomes in $\mathsf{m}\mathcal{M}$ therefore define a *set* of distributions $\{\mathrm{Pr}^{\boldsymbol{\pi}}\}_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M})}$ over the variables $\boldsymbol{V}$ where $\mathrm{Pr}^{\boldsymbol{\pi}}(\boldsymbol{v}; \boldsymbol{\theta}) = \mathrm{Pr}(\boldsymbol{v} \mid \mathbf{m})$ for $\mathbf{m}_D = \boldsymbol{\pi}$ and $\mathbf{m}_{V \backslash D} = \boldsymbol{\theta}$, which can be extended over the mechanised MAID as $\mathrm{Pr}(\boldsymbol{v}, \mathbf{m}) = \mathrm{Pr}(\boldsymbol{v} \mid \mathbf{m})\delta(\mathbf{M}, \mathbf{m})$. We can thus view a mechanised MAID as a set of BNs induced by the $\mathcal{R}$-rational outcomes.[12] The key point here is that rationality relations form a principled, formal, and highly general way to model the inherent non-determinism in a game, in order to render the model suitable for causal reasoning.

Note that Definition 16 is related to the notion of a *solution* in cyclic causal models [1, 8]. In both formalisms, a solution corresponds to a joint probability distribution consistent with all cyclic relationships, and in general a model may have many or no solutions. This, in turn, will impact the answers to the causal queries that we might wish to ask (as we explain further in Section 4). Similarly, we can view the rational responses at a decision variable as a *credal set* [60], and therefore the resultant model as a form of credal network [14], where the imprecise probabilities arise from the fact that there may be more than one rational outcome in a game.

Remark 1. *Often, agents might only be boundedly rational [89]. For example, due to computational costs, agents might seek only to satisfice (rather than optimise) their expected utility, leading to ϵ-approximate variations of equilibria concepts in which each agent cannot improve their expected utility by more than $\epsilon > 0$ by deviating [16]. These bounded rationality conditions*

---

[12] We continue, however, to highlight different variable types in our diagrams (rather than drawing each variable as a chance variable) for clarity of exposition.

can also be captured using rationality relations.

3.2 Relevance

A natural question to ask is which edges in the mechanised graph $\mathsf{m}\mathcal{G}$ are necessary. In other words, which elements of $\mathbf{Pa}_{\Pi_D}$ are necessary for computing the rational response $r_D(\mathbf{pa}_{\Pi_D})$? In the mechanised game shown in Figure 4a, for example, we don't need to know θ_{U^2} in order to compute the set of best responses π_{D^1} if we are given π_{D^2}. Whenever a mechanism variable M_V is found to be irrelevant to Π_D in this sense, we may remove the edge $\mathsf{M}_V \to \Pi_D$ to make this independence explicit. Removing all irrelevant edges results in a subgraph $\mathsf{m}_\mathcal{R}\mathcal{G}$ of $\mathsf{m}\mathcal{G}$ that is *minimal* with respect to $\mathcal{R}$.

In the extreme case where all decision rules are chosen independently from all other mechanism variables, then all edges $\mathscr{E}'$ between mechanism variables can be pruned. We call this subgraph of $\mathsf{m}\mathcal{G}$ the *independent mechanised graph*[13] $\mathsf{m}_\perp\mathcal{G}$. Game-theoretic models thereby break the *independent causal mechanism* assumption, which states that mechanisms should be causally and probabilistically independent of each other [75].

Formally, we have the following definitions, which are similar in spirit to those proposed in earlier work [51, 64, 70], but differ in that they consider the case of non-deterministic rationality relations that encode the solutions of a game.

Definition 17. Given a mechanised MAID $\mathsf{m}\mathcal{M}$ with rationality relations $\mathcal{R}$, $\mathsf{M}_V \in \mathbf{Pa}_{\Pi_D}$ is $\mathcal{R}$-**relevant** to Π_D if there exists $\mathbf{pa}_{\Pi_D} \neq \mathbf{pa}'_{\Pi_D}$ such that $r_D(\mathbf{pa}_{\Pi_D}) \neq r_D(\mathbf{pa}'_{\Pi_D})$, where $\mathbf{pa}_{\Pi_D}$ and $\mathbf{pa}'_{\Pi_D}$ differ only on M_V.

Definition 18. Given a mechanised MAID $\mathsf{m}\mathcal{M}$ with rationality relations $\mathcal{R}$, we say that its mechanised graph $\mathsf{m}\mathcal{G}$ is $\mathcal{R}$-**minimal**, denoted by $\mathsf{m}_\mathcal{R}\mathcal{G}$, when it contains an edge $\mathsf{M}_V \to \Pi_D$ if and only if M_V is $\mathcal{R}$-relevant to Π_D. When $\mathsf{m}_\mathcal{R}\mathcal{G}$ is restricted to the mechanism variables $\mathbf{M}$, we refer to it as the $\mathcal{R}$-**relevance graph**, denoted $\mathsf{r}_\mathcal{R}\mathcal{G}$.

Definition 17 is a property of a *game* after mechanising it via rationality relations. Therefore, an immediate follow-up question is whether given some choice of $\mathcal{R}$, we can identify $\mathcal{R}$-relevance (and hence prune edges in order to find the $\mathcal{R}$-minimal graph) for any MAID $\mathcal{M} = (\mathcal{G}, \boldsymbol{\theta})$ simply by appealing to its underlying *graph* $\mathcal{G}$, and the form of $\mathcal{R}$. For many natural choices of $\mathcal{R}$ this is indeed the case, and we can derive sound and complete graphical criteria for identifying $\mathcal{R}$-relevance. For a given $\mathcal{R}$, we refer to these criteria as $\mathcal{R}$-*reachability*. For example, the graphical criteria defining $\mathcal{R}^{\mathrm{BR}}$-reachability in Proposition 3 enable us to remove the dotted edges $\Theta_{U^1} \to \Pi_{D^2}$ and $\Theta_{U^2} \to \Pi_{D^1}$ in Figure 4a.

Proposition 3. M_V *is* $\mathcal{R}^{BR}$-*relevant to* Π_D *if and only if* $\mathsf{M}_V \not\perp_{\mathsf{m}_\perp\mathcal{G}} \boldsymbol{U}^i \cap \boldsymbol{Desc}_D \mid D, \boldsymbol{Pa}_D$ *or* $\mathsf{M}_V \not\perp_{\mathsf{m}_\perp\mathcal{G}} \boldsymbol{Pa}_D$, *where if* $D \in \boldsymbol{D}^i$, *then* $\boldsymbol{U}^i \cap \boldsymbol{Desc}_D \neq \varnothing$.

By applying $\mathcal{R}$-reachability to $\mathsf{m}\mathcal{G}$, we can find the $\mathcal{R}$-minimal mechanised graph $\mathsf{m}_\mathcal{R}\mathcal{G}$ and thus the $\mathcal{R}$-relevance graph $\mathsf{r}_\mathcal{R}\mathcal{G}$. This latter object will be one that we make repeated use of, and can be viewed as a generalisation of K&M's concept of a relevance graph simpliciter, to include all mechanism variables (as opposed to just those for decision variables) and for use with any $\mathcal{R}$.[14] To avoid confusion, we refer to K&M's relevance graph as an *s-relevance* graph. By generalising *s*-relevance (Definition 12) and *s*-reachability (Proposition 1) to all mechanism

---

[13]Dawid [18] call this diagram the extended influence diagram, but they do not investigate the dependencies and relationships between mechanisms.

[14]Note that the direction of the edges of the relevance graphs in this paper are the same as in [51], but are reversed compared with [50] and [42]. K&M also refer to V being relevant to/reachable from D, whereas we phrase this in terms of M_V being relevant to/reachable from Π_D.

variables, we can see in Figure 4a, for example, that Θ_T is not s-relevant to Π_{D^1}, though it is $\mathcal{R}^{\mathrm{BR}}$-relevant.

Our generalisation of the idea of s-relevance to different rationality relations parallels our development of various equilibrium refinements within MAIDs (introduced in Section 5.3), as opposed to the singular concept introduced by K&M. For instance, we shall see in Section 5.3 that s-relevance corresponds to the concept of subgame perfectness. Moreover, use of the full mechanised graph as opposed to simply its restriction to the decision rule variables is not only important for reasoning about game-theoretic notions such as equilibria and subgames, but will also be critical for reasoning about causality in games where agents may adapt their policies in response to interventions on any mechanism variable (not just those representing decision rules).

Generalising the Soundness and Completeness Results

We conclude this section by noting that it is possible to generalise the proof procedures for Propositions 3 and 1 to other choices of $\mathcal{R}$. The key step in doing so is to identify, for any decision variable D, a 'sufficient' (sound) but 'minimal' (complete) set of queries $\mathcal{Q}_D \subseteq \mathcal{Q}$ whose values, given some $\mathbf{m} = (\boldsymbol{\pi}, \boldsymbol{\theta})$, completely determine each r_D. Formally, let $\mathcal{Q}(\mathbf{m})$ be the set of probabilistic queries $\mathrm{Pr}^{\boldsymbol{\pi}}(\boldsymbol{x} \mid \boldsymbol{y}; \boldsymbol{\theta})$ over $\boldsymbol{X}, \boldsymbol{Y} \subseteq \boldsymbol{V}$ that can be formed in a MAID given $\mathbf{m} = (\boldsymbol{\pi}, \boldsymbol{\theta})$. Note that we can use such queries to express expected utilities, among many other things. For any decision variable D, we are looking for a sufficient but minimal subset of queries $\mathcal{Q}_D \subseteq \mathcal{Q}$ and a truth-valued function g_D such that:

$$\pi_D \in r_D(\mathbf{m}_{\boldsymbol{V} \setminus \{D\}}) \quad \Leftrightarrow \quad g_D\big(\mathcal{Q}_D(\mathbf{m}), dom(\boldsymbol{V})\big). \tag{2}$$

In this setting, asking whether M_V is $\mathcal{R}$-relevant to Π_D reduces to asking whether the choice of m_V may affect some $\mathrm{Pr}^{\boldsymbol{\pi}}(\boldsymbol{x} \mid \boldsymbol{y}) \in \mathcal{Q}_D$, which is, in turn, equivalent to asking whether V is a *requisite probability node* for $\mathrm{Pr}^{\boldsymbol{\pi}}(\boldsymbol{x} \mid \boldsymbol{y})$ [86], introduced in Definition 3. Whether a variable is a requisite probability node can be determined using a well-established graphical criterion [31], given as Lemma 1.

We can use this criterion to establish graphical criteria for $\mathcal{R}$-reachability, which, given a judicious choice of $\mathcal{Q}_D$, will be sound and complete. For example, the graphical criteria in Propositions 1 (s-reachability) and 3 ($\mathcal{R}^{\mathrm{BR}}$-reachability) correspond to choices:

$$\mathcal{Q}_D^s = \{\mathrm{Pr}^{\boldsymbol{\pi}}(\boldsymbol{u}^i \cap \mathbf{desc}_D \mid d, \mathbf{pa}_D)\},$$
$$\mathcal{Q}_D^{\mathrm{BR}} = \{\mathrm{Pr}^{\boldsymbol{\pi}}(\boldsymbol{u}^i \cap \mathbf{desc}_D \mid d, \mathbf{pa}_D), \mathrm{Pr}^{\boldsymbol{\pi}}(\mathbf{pa}_D)\},$$

respectively. With such a soundness and completeness result in hand, we can identify an $\mathcal{R}$-relevance graph as follows:

$$
\begin{array}{lll}
& \mathsf{r}_{\mathcal{R}}\mathcal{G} \text{ contains an edge } \mathsf{M}_V \to \Pi_D & \\
\Leftrightarrow & \mathsf{M}_V \text{ is } \mathcal{R}\text{-relevant to } \Pi_D & \text{by Definition 18} \\
\Leftrightarrow & r_D(\mathbf{pa}_{\Pi_D}) \text{ may vary with the value of } \mathsf{M}_V & \text{by Definition 17} \\
\Leftrightarrow & \mathcal{Q}_D(\mathbf{pa}_{\Pi_D}, \pi_D) \text{ may vary with the value of } \mathsf{M}_V & \text{by (2)} \\
\Leftrightarrow & V \text{ is a requisite probability node for some } \mathrm{Pr}^{\boldsymbol{\pi}}(\boldsymbol{x} \mid \boldsymbol{y}) \in \mathcal{Q}_D & \text{by Definition 3} \\
\Leftrightarrow & \mathsf{M}_V \not\perp_{\mathsf{m}_{\perp}\mathcal{G}} \boldsymbol{X} \mid \boldsymbol{Y} \text{ for some } \mathrm{Pr}^{\boldsymbol{\pi}}(\boldsymbol{x} \mid \boldsymbol{y}) \in \mathcal{Q}_D & \text{by Lemma 1}
\end{array}
$$

4 Causality in Games

Many types of queries may be of interest in game-theoretic scenarios. For instance, recalling Example 1, we could ask questions corresponding to:

1) Predictions, such as a) 'Given that the worker went to university, what is their wellbeing?' or b) 'Given that the worker always decides to go to university, what is their wellbeing?'

2) Interventions, such as a) 'Given that the worker is forced to go to university, what is their wellbeing?' or b) 'Given that the worker goes to university if and only if they are selected via a lottery system, what is their wellbeing?'

3) Counterfactuals, such as a) 'Given that the worker didn't go to university, what would be their wellbeing if they had?' or b) 'Given that the worker never decides to go to university, what would be their wellbeing if they always decided to go to university?'

Although these queries are ostensibly similar, they belong to different levels of the causal hierarchy. Predictions can be answered using (mechanised) MAIDs, which, as associational models, reside on level one. However, in order to reason about interventions and counterfactuals in games, we require models on levels two and three respectively. To this end, we introduce *causal games* (CGs) and *structural causal games* (SCGs). These can be viewed as generalising MAIDs to the causal setting, or CBNs and SCMs to the game-theoretic setting.

	SCM	SCIM	SCG		B	Bayesian
					C	Causal
level	CBN	CID	CG		G	Game
					I	Influence
	BN	ID	MAID		M	Model
					MA	Multi-Agent
		agents →		N	Network
					S	Structural

Figure 5: In this paper, we introduce CGs and SCGs. The causal hierarchy (associational, interventional, and counterfactual) forms the vertical axis and the number of agents (0, 1, and n) forms the horizontal axis. Note that all models in this diagram can also be mechanised.

The connections between all of these models and their various acronyms are displayed in Figure 5. Importantly, the models in each row are a special case of those in the row below (and the models in each column generalise those in the column to its left). As such, everything defined with respect to MAIDs (such as the mechanised games of Definition 15 and the equilibrium refinements we introduce in Section 5) are also well-defined in both CGs and SCGs.

Note that as well as asking queries regarding the object-level variables – such as queries 1a, 2a, and 3a – *after* some policy profile π has been chosen, mechanised games allow us to also ask queries regarding the mechanism variables – such as queries 1b, 2b, and 3b – *before* fixing a policy. While this distinction between 'post-policy' and 'pre-policy' queries is less significant in the case of predictions, we shall see in Section 4.2 that this difference corresponds to whether agents can adjust their policies in response to an intervention or not.

This distinction, which is critical in game-theoretic settings, does not apply to standard causal models, as they do not contain strategic, decision-making agents. Similarly, standard game-theoretic models do not explicitly represent this distinction either, as the dependencies between agents' decisions and other aspects of the game are not explicitly captured. Computationally, pre-policy interventions correspond to altering the original game then calculating the outcomes, and post-policy interventions correspond to calculating the outcomes of the original game, and then altering them. In a sense, these latter queries are a more natural analogue of existing work; in this work we unify both types of query using the same formalism.

In the remainder of this section, we show how the six types of causal queries represented by the examples above (each of which is written formally in Table 1) can be answered using causal

games. Doing so requires us to identify which level of the causal hierarchy a query belongs to, and to assess whether the query is pre-policy or post-policy. It is also simple to combine pre- and post-policy queries using mechanised games, though for clarity and brevity we do not do so here.

	1) Prediction	2) Intervention	3) Counterfactual
a) Post-policy	$\Pr^{\boldsymbol{\pi}}(u^1 \mid g)$	$\Pr^{\boldsymbol{\pi}}(u^1_g)$	$\Pr^{\boldsymbol{\pi}}(u^1_g \mid \neg g)$
b) Pre-policy	$\Pr(u^1 \mid \bar{\pi}_{D^1})$	$\Pr(u^1_{\hat{\pi}_{D^1}})$	$\Pr(u^1_{\tilde{\pi}_{D^1}} \mid \tilde{\pi}_{D^1})$

Table 1: Examples of the queries we may ask in causal models of games, corresponding to those listed at the top of this section. Recall that $dom(D^1) = \{g, \neg g\}$, indicating whether or not the worker goes to university, and that $\bar{\pi}_{D^1}, \hat{\pi}_{D^1}, \tilde{\pi}_{D^1}$ denote possible values of the decision rule variable Π_{D^1}. For example, $\bar{\pi}_{D^1}$ in query 1b is the decision rule in which the worker always decides to go to university, i.e., $\bar{\pi}_{D^1}(D_1 \mid T) = \delta(D_1, g)$. As introduced in Section 2, $\Pr(\boldsymbol{x_y})$ denotes the probability of $\boldsymbol{x}$ given a hard intervention ($\boldsymbol{Y} = \boldsymbol{y}$) and $\Pr(\boldsymbol{x_y} \mid \boldsymbol{z})$ represents the counterfactual probability of $\boldsymbol{x}$ had $\boldsymbol{y}$ been true, given that (in fact) $\boldsymbol{z}$ was true.

4.1 Predictions

Each policy profile $\boldsymbol{\pi}$ in a MAID $\mathcal{M}$ induces a BN with joint distribution $\Pr^{\boldsymbol{\pi}}(\boldsymbol{V}; \boldsymbol{\theta})$. We can therefore easily compute the probability of $\boldsymbol{x}$ given some observation $\boldsymbol{z}$, written $\Pr^{\boldsymbol{\pi}}(\boldsymbol{x} \mid \boldsymbol{z})$, under a given policy profile $\boldsymbol{\pi}$. Note that the distribution $\Pr^{\boldsymbol{\pi}}(\boldsymbol{V}; \boldsymbol{\theta})$ in the MAID can equivalently be written in the mechanised MAID as $\Pr(\boldsymbol{V} \mid \mathbf{m})$ with $\mathbf{m} = (\boldsymbol{\pi}, \boldsymbol{\theta})$.

However, in game-theoretic settings, we typically assume only that a *rational outcome* of the game will be chosen, not some unique $\boldsymbol{\pi}$. Moreover, we may not have any reason to favour one rational outcome over another, implying that we ought to evaluate queries with respect to a *set* of policy profiles. In the definition below, we therefore consider the distribution $\Pr^{\boldsymbol{\pi}}(\boldsymbol{x} \mid \boldsymbol{z})$ induced by each rational outcome that is consistent with the observation $\boldsymbol{z}$. Note that, as remarked in Section 3.1, there may be many or no such rational outcomes.

Definition 19. Given a mechanised MAID m$\mathcal{M}$ with rationality relations $\mathcal{R}$, the **answer to a conditional query** of the probability of $\boldsymbol{x}$ given observation $\boldsymbol{z}$ is given by the set $\Pr^{\mathcal{R}}(\boldsymbol{x} \mid \boldsymbol{z}) := \left\{ \Pr^{\boldsymbol{\pi}}(\boldsymbol{x} \mid \boldsymbol{z}) \right\}_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z})}$ where $\mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z}) := \{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M}) : \Pr^{\boldsymbol{\pi}}(\boldsymbol{z}) > 0\}$ is the set of **conditional rational outcomes**. In general, $\boldsymbol{Z} \subseteq \boldsymbol{V} \cup \mathbf{M}$ can include mechanism variables, and so we compute $\Pr^{\boldsymbol{\pi}}(\boldsymbol{x} \mid \boldsymbol{z})$ in $\mathcal{M}$ as $\Pr(\boldsymbol{x} \mid \boldsymbol{z}, \mathbf{m}')$ in m$\mathcal{M}$, where $\mathbf{M}' = \mathbf{M} \setminus \boldsymbol{Z}$ and $\mathbf{m}_D = \boldsymbol{\pi}$.

More generally, we can view queries in games as *first-order* queries defined over formulae in which $\boldsymbol{\pi}$ is a free variable, such as $\varphi(\boldsymbol{\pi}) \equiv \Pr^{\boldsymbol{\pi}}(\boldsymbol{x} \mid \boldsymbol{z})$. The answers to these queries are therefore only well-defined when this free variable becomes bound, or when considering a set of answers, as in Definition 19. For example, we can bind $\boldsymbol{\pi}$ by quantifying over it as in $\exists \boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z}). \left(\varphi(\boldsymbol{\pi}) \sim q \right)$ where $\sim \in \{<, \leq, =, \geq, >\}$ and $q \in [0, 1]$ (as is often done in first-order logics for reasoning about and verifying multi-agent systems [12, 56, 100]), or by returning bounds such as $\max_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z})} \varphi(\boldsymbol{\pi})$ (as is often done in credal networks [14]). Above, we quantify over $\mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z})$, but it is also possible to quantify over any desired set of policies, or even to posit a prior distribution $\Pr(\boldsymbol{\Pi})$ over policies. These queries strictly generalise those in BNs and models such as settable systems [95, 96], which effectively consider only a single instantiation of $\boldsymbol{\Pi}$.

By way of illustration, let us return to Example 1 and query 1a from earlier in this section. We can interpret this question as asking about the expected utility of the worker given the observation that they went to university, written $\mathbb{E}_{\boldsymbol{\pi}}[U^1 \mid g]$ for some policy $\boldsymbol{\pi}$. For the worked examples in this section, we assume that $p = \Pr(T = h) = \frac{1}{2}$ and that the automated hiring

system and the worker are playing best responses to one another, i.e., $\mathcal{R} = \mathcal{R}^{\mathrm{BR}}$. The various rational outcomes induced by this choice (i.e., the NEs of the game) are:

1. The worker always chooses $\neg g$. The hiring system chooses j if the worker chose $\neg g$, otherwise they choose j with any probability $q \in [0, 1]$.

2. The worker chooses $\neg g$ if $T = \neg h$, and otherwise chooses g with probability $\frac{1}{2}$. The hiring system chooses j if the worker chose g, otherwise they choose j with probability $\frac{4}{5}$.

3. The worker always chooses g. The hiring system chooses j if the worker chose g, otherwise they choose j with any probability $q \in [0, \frac{3}{5}]$.

The expected utilities for the worker under these rational outcomes are 5, 4, and $\frac{7}{2}$, respectively. In query 1a, the conditional rational outcomes $\mathcal{R}(\mathsf{m}\mathcal{M} \mid g)$ are the NEs consistent with observing $D^1 = g$. To answer the query, we must therefore compute $\mathrm{Pr}^{\mathcal{R}}(u^1 \mid g)$, which yields $\{\mathbb{E}_{\boldsymbol{\pi}}[U^1 \mid g]\}_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M}|g)} = \{\frac{7}{2}, 4\}$.

As noted above, Definition 19 can also be employed when the observations made are of the *mechanism variables*. For example, in query 1b we condition on the observation that the worker's strategy is given by $\delta(D^1, g)$, which we refer to as $\bar{\pi}_{D^1}$, i..e, the worker *always* decides to go to university. Based on this, we can use the mechanised MAID to compute $\mathrm{Pr}^{\mathcal{R}}(u^1 \mid \bar{\pi}_{D^1})$ and thus that $\{\mathbb{E}[U^1 \mid \bar{\pi}_{D^1}]\}_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M}|\bar{\pi}_{D^1})} = \{\frac{7}{2}\}$. Note that this set is distinct from the answer to query 1a as observations of mechanism and object-level variables provide us with different information, i.e., $\mathcal{R}(\mathsf{m}\mathcal{M} \mid \bar{\pi}_{D^1}) \neq \mathcal{R}(\mathsf{m}\mathcal{M} \mid g)$. In particular, observations of mechanism variables serve primarily to rule out certain rational outcomes by conditioning on decision rules (conditioning on *parameter* variables tells us nothing as they have deterministic distributions and no parents).

One advantage of computing predictions in MAIDs (as opposed to in EFGs, for instance) is that we may exploit the conditional independencies in the graph. For example, if we were interested in how likely a worker is to be hard-working given that they went to university and were hired, then $T \perp_{\mathcal{G}} D^2 \mid D^1$ implies that $\mathrm{Pr}^{\boldsymbol{\pi}}(h \mid g, j) = \mathrm{Pr}^{\boldsymbol{\pi}}(h \mid g)$ for any policy profile $\boldsymbol{\pi}$. When answering queries over object-level variables using mechanised MAIDs, we implicitly condition on the values of the mechanism variables to represent the fact the game and policy under consideration are fixed. For example, the query $\mathrm{Pr}^{\boldsymbol{\pi}}(h \mid g, j)$ in $\mathcal{M}$ is given by $\mathrm{Pr}(h \mid g, j, \mathbf{m})$ in $\mathsf{m}\mathcal{M}$, where $\mathbf{m}_D = \boldsymbol{\pi}$. Hence, although $T \not\perp_{\mathsf{m}\mathcal{G}} D^2 \mid D^1$, we do have $T \perp_{\mathsf{m}\mathcal{G}} D^2 \mid D^1, \mathbf{M}$, as expected.

4.2 Interventions

Interventional queries concern the effect of causal influences from outside a system. This becomes especially interesting in the case of games, when interventions affect not only the environment but also how the self-interested agents adapt their policies in response. In order to answer such queries, the edges in a MAID must reflect the causal structure of the world. This gives rise to the following definition, which can be viewed simply as a CBN without parameters for the decision variables. Note that as causal games are a form of MAID, they also support the associational queries introduced in the preceding section (just as CBNs may also be used to compute both interventional and associational queries).

Definition 20. A **causal game (CG)** $\mathcal{M} = (\mathcal{G}, \boldsymbol{\theta})$ is a MAID such that for *any* parameterisation of the decision variable CPDs $\boldsymbol{\pi}$, the induced model with joint distribution $\mathrm{Pr}^{\boldsymbol{\pi}}(\boldsymbol{V})$ is a CBN.

Unlike CBNs, CGs let us ask about the effect of an intervention *before* or *after* a policy profile has been selected, which we refer to as *pre-* and *post-policy* queries respectively. Asking about the effect of an intervention after a particular policy profile $\boldsymbol{\pi}$ has been selected (as in query 2a)

is simply the same as performing an interventional query on the CBN with joint distribution $\Pr^{\boldsymbol{\pi}}$. Asking about the effect of an intervention *before* a policy profile has been selected (as in query 2b) means that agents are made aware of the intervention before selecting their decision rules, and thus they may react to its effects. In other words, the intervention can be viewed as producing a slightly different game that the agents then (knowingly) play.

Our key observation is that pre-policy interventions can be modelled as interventions on the *mechanism variables* in the mechanised CG, which ensures that the effects are propagated through the processes via which agents select their decision rules. This is because the additional mechanism variables and their outgoing edges in a mechanised CG represent causal (though potentially non-deterministic) processes via which parameterisations for the object-level variables are selected. Post-policy interventions, in turn, can be modelled as standard interventions on object-level variables. We write $\mathsf{m}\mathcal{M}_\mathcal{I}$ for an intervention $\mathcal{I}$ which may contain both pre-policy and post-policy interventions.

This unification of pre- and post-policy interventions is one of the key benefits of mechanised models. Indeed, post-policy interventions, and pre-policy interventions on parameter variables, are defined exactly as in CBNs, while a pre-policy intervention on a decision rule variable Π_D corresponds to replacing $r_D : dom(\mathbf{Pa}_{\Pi_D}) \to dom(\Pi_D)$ by some new $r_D^* : dom(\mathbf{Pa}_{\Pi_D}^*) \to dom(\Pi_D)$, where we may have $\mathbf{Pa}_{\Pi_D}^* \neq \mathbf{Pa}_{\Pi_D}$. As for conditional queries, in our definition (which mirrors that introduced for cyclic causal models [8], but where the cyclic dependencies are governed by relations instead of functions [1]) we quantify over the set of rational outcomes that are consistent with the given intervention.

Definition 21. Given a mechanised CG $\mathsf{m}\mathcal{M}$ with rationality relations $\mathcal{R}$, the **answer to an interventional query** of the probability of $\boldsymbol{x}$ given intervention $\mathcal{I}$ on variables $\boldsymbol{Y}$ is given by the set $\Pr^{\mathcal{R}}(\boldsymbol{x}_\mathcal{I}) := \big\{ \Pr^{\boldsymbol{\pi}}(\boldsymbol{x}_\mathcal{I}) \big\}_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M}_\mathcal{I})}$ where $\mathcal{R}(\mathsf{m}\mathcal{M}_\mathcal{I})$ is the set of **interventional rational outcomes** in the mechanised MAID $\mathsf{m}\mathcal{M}_\mathcal{I}$ with rationality relations $\mathcal{R}^* := \{r_D^*\}_{\Pi_D \in \boldsymbol{Y}} \cup \{r_D\}_{\Pi_D \notin \boldsymbol{Y}}$ and parameters determined by $\Pr(\boldsymbol{\Theta}_\mathcal{I})$. Note that if $\mathcal{I}$ is fully post-policy, then the rational outcomes remain the same, i.e., $\mathcal{R}(\mathsf{m}\mathcal{M}_\mathcal{I}) = \mathcal{R}(\mathsf{m}\mathcal{M})$ when $\boldsymbol{Y} \subseteq \boldsymbol{V}$.

To illustrate these ideas, let us return to Example 1 and queries 2a and 2b. Query 2a concerns a *post-policy* intervention since the worker is forced to go to university unbeknownst to the firm's hiring system; in other words, the hiring system does not observe this fact before computing a decision rule. To compute the worker's expected utility, we must calculate $\Pr^{\mathcal{R}}(u_g^1)$ and thus perform a hard intervention $do(D^1 = g)$ in the mechanised game (shown in Figure 6a). As the set of rational outcomes does not change under a post-policy intervention, we have that $\Pr^{\mathcal{R}}(u_g^1) = \big\{ \Pr^{\boldsymbol{\pi}}(u_g^1) \big\}_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M})}$, which results in $\big\{ \mathbb{E}_{\boldsymbol{\pi}}[U_g^1] \big\}_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M})} = \{-\frac{3}{2}, \frac{7}{2}\}$. Note that unlike query 1a, the fact that $D^1 = g$ tells us nothing about the value of T, as it is causally upstream of the intervention on D^1. Therefore, the wellbeing of the worker may decrease because they may be sent to university even when they are lazy.

To answer query 2b, concerning a *pre-policy* intervention, we must compute $\Pr(u_{\hat{\pi}_{D^1}}^1)$ where $\hat{\pi}_{D^1}(g \mid h) = \hat{\pi}_{D^1}(g \mid \neg h) = \frac{1}{2}$ represents the aforementioned lottery system, which selects students to attend university randomly with probability $\frac{1}{2}$. This time, the firm modifies their hiring system after observing the intervention – denoted by $\mathcal{I}$ and shown in Figure 6b – but before the hiring system computes a decision rule, and so the new set of rational outcomes is given by $\mathcal{R}(\mathsf{m}\mathcal{M}_\mathcal{I}) = \{(\hat{\pi}_{D^1}, \pi_{D^2}) : \pi_{D^2} \in r_{D^2}(\mathbf{pa}_{\Pi_{D^2}})\}$. In other words, we set $\Pi_{D^1} = \hat{\pi}_{D^1}$ using a hard intervention and then allow the hiring system to best respond to this decision rule using r_{D^2}. Note that $\mathcal{R}(\mathsf{m}\mathcal{M}_\mathcal{I}) \neq \mathcal{R}(\mathsf{m}\mathcal{M} \mid \hat{\pi}_{D^1}) = \varnothing$, as there is no NE in the game that contains decision rule $\hat{\pi}_{D^1}$. The lottery system removes any signalling effect of going to university, resulting in an optimal policy for the hiring system of always offering a job to the worker and expected utility $\big\{ \mathbb{E}_{\boldsymbol{\pi}}[U_{\hat{\pi}_{D^1}}^1] \big\}_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M}_\mathcal{I})} = \{\frac{17}{4}\}$.
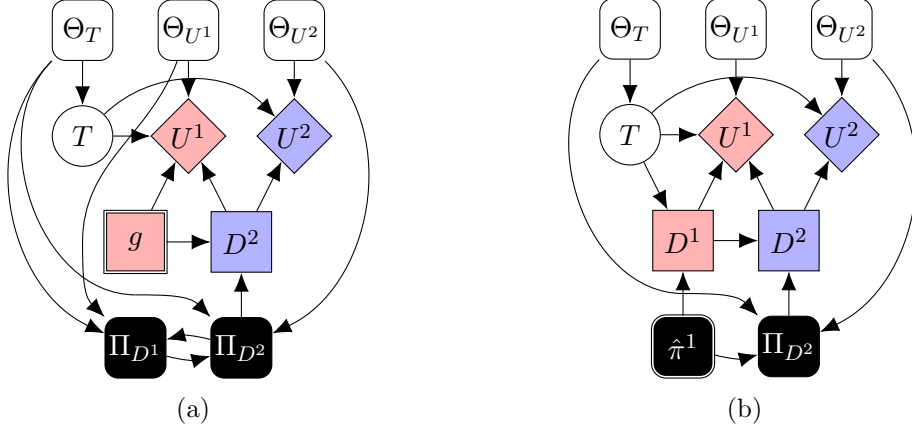
Figure 6: (a) A mechanised game showing the hard post-policy intervention $\mathrm{do}(D^1 = g)$, where incoming edges to D^1 are severed. (b) A mechanised game showing the hard pre-policy intervention $\mathrm{do}(\Pi_{D^1} = \hat{\pi}^1)$, where incoming edges to Π_{D^1} are severed.

Remark 2. *In previous work, soft interventions on object-level variables V have been modelled as hard interventions on its mechanism variable M_V [18]. While these can be viewed as essentially equivalent in the single-agent case,*[15] *possible dependencies between mechanism variables in mechanised games means that these two types of intervention may have markedly different effects in the multi-agent setting. The difference between pre- and post-policy interventions results from a difference in the* information *that is available to agents when they make their decisions, rather than to the chronology of play in a game (as is also the case for the structure of EFGs).*

4.3   Counterfactuals

The final type of question we investigate arises when we combine predictions and interventions, as in query 3a: 'Given that the worker didn't go to university, what would be their wellbeing if they had?'. Such questions are *counterfactual*, as they combine observations made in the actual world (in which the worker didn't go to university), with questions pertaining to a counterfactual world (where they did go to university). Answering these queries in games is significantly more nuanced and complex than those of the preceding subsections. To do so, we must first consign all stochasticity to a set of exogenous variables, one for each variable in the causal game. Just as in an SCM, each variable V is thus associated with an exogenous variable E_V, and is governed by a deterministic CPD $\mathrm{Pr}^{\boldsymbol{\pi}}(V \mid \mathbf{Pa}_V)$, where $\mathsf{E}_V \in \mathbf{Pa}_V$.

Definition 22. A (Markovian) **structural causal game (SCG)** $\mathcal{M} = (\mathcal{G}, \boldsymbol{\theta})$ is a causal game over exogenous and endogenous variables $\mathbf{E} \cup \boldsymbol{V}$ such that for *any* (deterministic) parameterisation of the decision variable CPDs $\dot{\boldsymbol{\pi}}$, the induced model with joint distribution $\mathrm{Pr}^{\dot{\boldsymbol{\pi}}}(\boldsymbol{V}, \mathbf{E})$ is an SCM.

An SCG can be seen as an SCM without parameters for the decision variables. Given a policy $\boldsymbol{\pi}$, we recover an SCM, as we explain in more detail below. Meanwhile, mechanised SCGs can be viewed as (a special case of) a relational causal model with cycles [1].

When we mechanise SCGs, although we introduce mechanism variables for the exogenous variables (as can be seen in Figure 7b), we view them and their object-level exogenous children as beyond the realm of observation and intervention, just as in SCMs. As such, mechanism variables of exogenous variables can largely be ignored. As in SCMs, interventional distributions $\mathcal{I}(V \mid \mathbf{Pa}_V^*)$ must be deterministic, and soft interventions may be defined by introducing a new exogenous variable E_V^* to $\mathbf{Pa}_V^*$ [13]. With these caveats in place, both pre-policy and

---

[15]To be precise, this is true when the agent has *sufficient recall*; see Definition 14.

post-policy predictions and interventions may be defined in this model as described in 4.1 and 4.2 respectively.
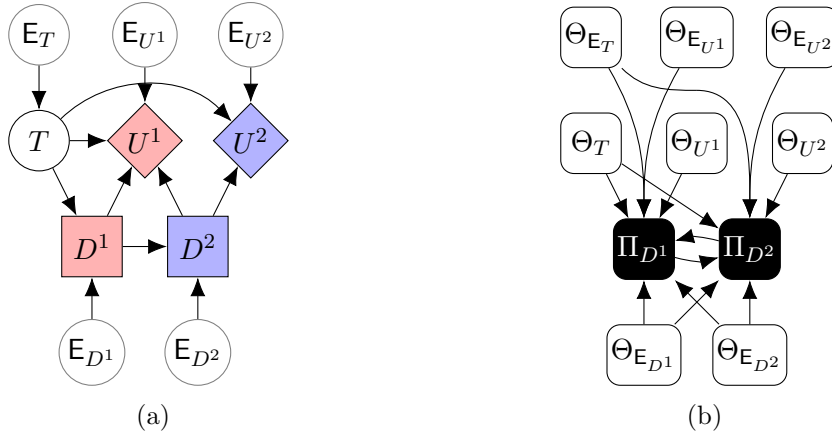

Figure 7: (a) A (Markovian) SCG representing Example 1. Note that we have included exogenous variables for U^1 and U^2, although as neither is stochastic, this is not strictly necessary. (b) The $\mathcal{R}^{\mathrm{BR}}$-relevance graph of this game.

While computing predictions and interventions in SCGs is therefore relatively straightforward, there are two main difficulties that arise when computing counterfactuals. The first is the choice of how to represent stochastic decision rules using structural functions and exogenous variables, and the second is the problem of updating our beliefs about the policy profile played in the counterfactual world given our evidence about the policy profile played in the actual world. We resolve each of these difficulties in turn.

Decision Rules as Structural Functions

We assume that agents play the same kind of game regardless of the level in the causal hierarchy at which we model them. In these games, (behavioural) play equates to selecting decision rules which stochastically sample a decision conditional on the value of some (non-exogenous) parents. In *structural* causal games, we represent these decision rules using structural functions and exogenous variables. One proposal would therefore be to view each agent as choosing both a 'structural decision rule' $\dot{\pi}_D(D \mid \mathbf{Pa}_D)$ and a distribution $\Pr(\mathsf{E}_D)$, with a shared mechanism parent Π_D for both D and E_D. This, however, leads to a different type signature for decision rules, and moreover leads to a formalism in which (pre-policy) interventions can be made upstream of stochastic variables, which are ruled out in SCMs. We therefore propose an equivalent formulation in which each agent controls only their decision variables and not their exogenous parents.

Unfortunately, while our assumptions about the rationality of the agents tell us what CPDs are assigned to their decision variables, they are insufficient for telling us what precise deterministic mechanisms the agents use to implement these CPDs (as a function of some stochastic exogenous variable). In fact, unless we chose to explicitly restrict the form of the mechanism, such as by stipulating that it belongs to some parametric class, there will typically be *infinitely* many deterministic functions that induce a particular distribution over a decision variable [6]. Without specifying such functions, it will not (in general) be possible to answer counterfactual queries in games, and yet the precise form of these functions may impact the answers to these queries [18].

In essence, the choice of how to represent a decision rule $\pi_D \in \Delta(dom(D) \mid dom(\mathbf{Pa}_D \setminus \{\mathsf{E}_D\}))$ using a stochastic exogenous variable E_D and a deterministic mechanism $\dot{\pi}_D \in \Delta(dom(D) \mid$

$dom(\mathbf{Pa}_D))$ is the choice of what part of the decision rule we assume remains fixed across counterfactual worlds (E_D) and what part may vary ($\dot{\pi}_D$). Assuming that we have no pre-existing knowledge about this representation, we propose to stay true to the spirit of behavioural policies by viewing each agent's randomisation as independent between both:

- *Decision rules*, in the sense that learning about an agent's random choice under one decision rule π_D is uninformative in settings where the agent is using a different decision rule π'_D;

- *Decision contexts*, in the sense that an agent's decision rule π_D can naturally be interpreted as independently sampling an action d *after* seeing an assignment $\overline{\mathbf{pa}}_D$ of the non-exogenous parents $\overline{\mathbf{Pa}}_D := \mathbf{Pa}_D \setminus \{\mathsf{E}_D\}$.

We formalise this assumption by representing each exogenous variable E_D as a set containing a variable $\mathsf{E}_D^{\pi_D, \overline{\mathbf{pa}}_D}$ for each $\pi_D \in \Delta(dom(D) \mid dom(\overline{\mathbf{Pa}}_D))$ and $\overline{\mathbf{pa}}_D \in dom(\overline{\mathbf{Pa}}_D)$, where $dom(\mathsf{E}_D^{\pi_D, \overline{\mathbf{pa}}_D}) = dom(D)$ (i.e., E_D is a random field with independently distributed elements). Given a stochastic decision rule $\pi_D(D \mid \overline{\mathbf{Pa}}_D)$, we may then define a canonical structural representation by setting:

$$
\begin{aligned}
\Pr(\mathsf{E}_D^{\pi_D, \overline{\mathbf{pa}}_D} = d) &:= \pi_D(d \mid \overline{\mathbf{pa}}_D), \\
\dot{\pi}_D(D = d \mid \overline{\mathbf{pa}}_D, \mathsf{e}_D) &:= \delta(D, \mathsf{e}_D^{\pi_D, \overline{\mathbf{pa}}_D}),
\end{aligned}
\tag{3}
$$

where note that $\dot{\pi}_D$ is effectively parameterised by π_D, i.e., we have $\dot{\pi}_D(D \mid \overline{\mathbf{Pa}}_D, \mathsf{E}_D; \pi_D)$. The joint distribution over E_D is simply the *product of probability distributions* over $\mathsf{E}_D^{\pi_D, \overline{\mathbf{pa}}_D}$ [7]. Proposition 4 below then follows immediately,[16] and means that we may continue to interpret decision rules, policy profiles, and rationality relations as we do in MAIDs and CGs, where each agent plays some $\pi_D \in r_D(\mathbf{pa}_{\Pi_D}) \subseteq \Delta(dom(D) \mid dom(\overline{\mathbf{Pa}}_D))$ and each π_D parameterises $\dot{\pi}_D$. Moreover, this additional structure allows us to generalise the definition of counterfactuals in SCMs to counterfactuals in games.

Proposition 4. *For distributions over E_D and D as governed by equations (3), there is a one-to-one correspondence between the set of stochastic decision rules $\Delta(dom(D) \mid dom(\overline{\mathbf{Pa}}_D))$ and the set of deterministic decision rules $dom(\dot{\Pi}_D) \subset \Delta(dom(D) \mid dom(\mathbf{Pa}_D))$. Moreover, given two such corresponding decision rules π_D and $\dot{\pi}_D$, then $\int_{dom(\mathsf{E}_D)} \dot{\pi}_D(d \mid \overline{\mathbf{pa}}_D, \mathsf{e}_D) \Pr(\mathsf{e}_D) \, d\mathsf{e}_D = \pi_D(d \mid \overline{\mathbf{pa}}_D)$.*

Counterfactual Rational Outcomes

The second difficulty of answering counterfactual queries in games arises due to the possible existence of multiple rational outcomes. If we have evidence that the equilibrium $\boldsymbol{\pi}$ was played in the actual world, how and to what extent should that inform us of the equilibrium $\boldsymbol{\pi}'$ played in the counterfactual world where the values of some mechanism variables may have changed?

To answer this question, we begin by introducing our approach to answering counterfactual queries in SCGs, which mirrors Pearl's approach for SCMs (described in Section 2.1). That is, we condition, intervene, and then compute the resulting distribution, though we generalise this to observations and interventions on both object-level and mechanism variables. In particular, we compute the set $\Pr^{\mathcal{R}}(\boldsymbol{x}_{\mathcal{I}} \mid \boldsymbol{z})$ as follows:

1. For every *actual* rational outcome $\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z})$, update $\Pr(\mathbf{e})$ to $\Pr^{\boldsymbol{\pi}}(\mathbf{e} \mid \boldsymbol{z})$ ('abduction');

---

[16]Though note that there are many constructions that would result in these properties; we merely present one particularly simple example.

2. Apply the intervention $\mathcal{I}$, on variables $\boldsymbol{Y}$, recomputing any rational responses to form $\boldsymbol{\pi}'$ and adding new exogenous variables $\mathbf{E}^*$ to form $\mathbf{E}' = \mathbf{E} \cup \mathbf{E}^*$ where required ('action');

3. Return each marginal distribution $\int_{dom(\mathbf{E}')} \Pr^{\boldsymbol{\pi}'}(\boldsymbol{x} \mid \mathbf{e}') \Pr(\mathbf{e}') \, d\mathbf{e}'$ in this modified model for each new *counterfactual* rational outcome $\boldsymbol{\pi}'$ ('prediction').

In the first step, we update our beliefs about the exogenous and decision rule variables in the actual world under $\boldsymbol{\pi}$; in the second, we apply an intervention $\mathcal{I}$ to form the counterfactual world (and recompute the rational outcomes based on this change); and in the third, we return the new set of distributions that are consistent with our beliefs from the first step and the results of the intervention made in the second step.

By reading these steps carefully, one notices a difficulty in SCGs that does not arise in an SCM: when we condition on $\boldsymbol{z}$ in the first step we obtain a set of policies $\mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z})$ that are rational in the *actual* world. Then in step two, we compute new rational responses $\boldsymbol{\pi}'$, and it is $\boldsymbol{\pi}'$ rather than $\mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z})$ that features in the *counterfactual* world of the final step. This raises the question of the extent to which knowledge of the rational policies $\mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z})$ should be used to compute $\boldsymbol{\pi}'$.

Let us first note that if $\mathcal{I}$ is a *post*-policy intervention, then this has no impact on the rational outcomes of the counterfactual world – they are simply the same as the actual world, given by $\mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z})$ – and hence there is no difficulty. The issue only arises in *pre*-policy counterfactuals, such as query 3b ('Given that the worker never decides to go to university, what would be their wellbeing if they always decided to go to university?'), where the intervention (in query 3b, on the worker's decision rule) means that the set of rational outcomes in the counterfactual world will be different from those in the actual world.

Because each policy profile $\boldsymbol{\pi}$ is made up of decision rules π_D, we can formalise this question by asking which decision rule variables $\boldsymbol{\Pi}(\mathcal{I}) \subseteq \boldsymbol{\Pi}$ are *invariant* to the intervention $\mathcal{I}$. Written in terms of the three-step process above, we must have $\boldsymbol{\pi}(\mathcal{I}) = \boldsymbol{\pi}'(\mathcal{I})$, i.e., for any invariant decision rule variable $\Pi \in \boldsymbol{\Pi}(\mathcal{I})$ then the counterfactual decision rule π_D' is equal to the actual decision rule π_D. Those that are not invariant must have their values recomputed in the new counterfactual model. For instance, as argued above, when $\mathcal{I}$ is post-policy, $\boldsymbol{\Pi}(\mathcal{I}) = \boldsymbol{\Pi}$. That is, none of the values of the decision rule variables need to be recomputed. How should we choose $\boldsymbol{\Pi}(\mathcal{I})$ when $\mathcal{I}$ is not (fully) post-policy?

There are multiple principles that could be invoked in order to make this choice. The simplest – let us call it the *simplicity principle* – is to recompute the values of *all* decision rule variables, i.e., $\boldsymbol{\Pi}(\mathcal{I}) = \varnothing$. In other words, the intervention means that 'all bets are off' after the intervention is made, and the actual rational outcomes $\mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z})$ have no bearing on the counterfactual rational outcomes. Under this principle, computing pre-policy counterfactuals is reminiscent of the approach taken in existing work on cyclic causal models [8], where two sets of solutions are induced by two halves of a twin graph. The problem with this principle is that it may require us to ignore information gathered from our observation $\boldsymbol{z}$. For example, if $\boldsymbol{Z} = \boldsymbol{z}$ implies that $\Pi_D = \pi_D$ in the actual world and $\mathcal{I}$ is *causally downstream* of Π_D (and hence can have no effect on the value of Π_D), then this means we also know that $\Pi_D = \pi_D$ in the counterfactual world, i.e., $\Pi_D \in \boldsymbol{\Pi}(\mathcal{I})$.

To solve this problem, we can instead invoke the *closest possible world principle*, where we retain as much information as possible from our knowledge of the rational policies $\mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z})$. While the values of some decision rules may still need to be recomputed, by keeping $\boldsymbol{\Pi}(\mathcal{I})$ as large as possible, we avoid the need for re-solving the entire game, and can provide a more accurate answer to counterfactual queries. The process of computing $\boldsymbol{\Pi}(\mathcal{I})$ under this principle is slightly more complicated, however, as it involves propagating the effects of an intervention through models that contain both cycles and non-determinism. In the remainder of the paper

we therefore employ the simplicity principle above, which is also more in keeping with prior work.[17] An algorithm for computing $\boldsymbol{\Pi}(\mathcal{I})$ according to the closest possible world principle is, however, provided in Appendix B.1 for reference.

Defining Counterfactuals in Games

Given a set of invariant decision rule variables $\boldsymbol{\Pi}(\mathcal{I})$, the answer produced by the three-step process given above can be written as follows.

Definition 23. Given a mechanised SCG $\mathsf{m}\mathcal{M}$ with rationality relations $\mathcal{R}$, the **answer to a counterfactual query** of the probability of $\boldsymbol{x}$ given observation $\boldsymbol{z}$ and intervention $\mathcal{I}$ on variables $\boldsymbol{Y}$ is given by the set:

$$\mathrm{Pr}^{\mathcal{R}}(\boldsymbol{x}_{\mathcal{I}} \mid \boldsymbol{z}) \coloneqq \left\{ \int_{dom(\mathbf{E}')} \mathrm{Pr}^{\boldsymbol{\pi}'}(\boldsymbol{x}_{\mathcal{I}} \mid \mathbf{e}, \mathbf{e}^*) \, \mathrm{Pr}(\mathbf{e}^*) \, \mathrm{Pr}^{\boldsymbol{\pi}}(\mathbf{e} \mid \boldsymbol{z}) \, d\mathbf{e}' \right\}_{(\boldsymbol{\pi}, \boldsymbol{\pi}') \in \mathcal{R}(\mathsf{m}\mathcal{M}_{\mathcal{I}} \mid \boldsymbol{z})},$$

where $\mathbf{E}^* = \mathbf{E}' \setminus \mathbf{E}$ are any newly added exogenous variables as a result of a soft intervention,

$$\mathcal{R}(\mathsf{m}\mathcal{M}_{\mathcal{I}} \mid \boldsymbol{z}) \coloneqq \left\{ (\boldsymbol{\pi}, \boldsymbol{\pi}') \in \mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z}) \times \mathcal{R}(\mathsf{m}\mathcal{M}_{\mathcal{I}}) : \boldsymbol{\pi}(\mathcal{I}) = \boldsymbol{\pi}'(\mathcal{I}) \right\},$$

is the set of **actual-counterfactual rational outcomes**, and $\boldsymbol{\Pi}(\mathcal{I})$ is the set of **invariant decision rule variables**. Note that if $\mathcal{I}$ is fully post-policy, then the rational outcomes remain the same, i.e., $\boldsymbol{\Pi}(\mathcal{I}) = \boldsymbol{\Pi}$ and $\mathcal{R}(\mathsf{m}\mathcal{M}_{\mathcal{I}} \mid \boldsymbol{z}) = \left\{ (\boldsymbol{\pi}, \boldsymbol{\pi}) : \boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z}) \right\}$ when $\boldsymbol{Y} \subseteq \boldsymbol{V}$.

In the definition of $\mathrm{Pr}^{\mathcal{R}}(\boldsymbol{x}_{\mathcal{I}} \mid \boldsymbol{z})$, for every actual policy $\boldsymbol{\pi}$ and counterfactual policy $\boldsymbol{\pi}'$, we compute the product of three quantities. $\mathrm{Pr}^{\boldsymbol{\pi}}(\mathbf{e} \mid \boldsymbol{z})$ is the updated distribution over the exogenous variables under $\boldsymbol{\pi}$, and corresponds to the first step. If $\mathcal{I}$ is a soft intervention, then we add further variables $\mathbf{E}^* = \mathbf{E}' \setminus \mathbf{E}$ to capture the stochasticity of $\mathcal{I}$, which leads to the term $\mathrm{Pr}(\mathbf{e}^*)$.[18] We then condition on both $\mathbf{e}$ and $\mathbf{e}^*$ and compute the value of $\boldsymbol{x}_{\mathcal{I}}$ under $\boldsymbol{\pi}'$, hence the term $\mathrm{Pr}^{\boldsymbol{\pi}'}(\boldsymbol{x}_{\mathcal{I}} \mid \mathbf{e}, \mathbf{e}^*)$. Finally, in the third step, we marginalise over all exogenous variables $\mathbf{E}'$. The set $\mathcal{R}(\mathsf{m}\mathcal{M}_{\mathcal{I}} \mid \boldsymbol{z})$ defines the pairs of policies that we must consider. Namely, actual rational outcomes $\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z})$ and counterfactual rational outcomes $\boldsymbol{\pi}' \in \mathcal{R}(\mathsf{m}\mathcal{M}_{\mathcal{I}})$ such that the decision rules invariant to $\mathcal{I}$, denoted $\boldsymbol{\Pi}(\mathcal{I})$, remain the same: $\boldsymbol{\pi}(\mathcal{I}) = \boldsymbol{\pi}'(\mathcal{I})$.

We briefly demonstrate the three step process above by returning to queries 3a and 3b. To answer 3a we must compute $\mathrm{Pr}^{\boldsymbol{\pi}}(u_g^1 \mid \neg g)$. First, we note that as this involves a post-policy intervention then we only need to consider the *actual* rational outcomes, as $\boldsymbol{\Pi}(\mathcal{I}) = \boldsymbol{\Pi}$. Thus, for each $\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M} \mid \neg g)$ we begin by updating $\mathrm{Pr}(\mathbf{e})$ to $\mathrm{Pr}^{\boldsymbol{\pi}}(\mathbf{e} \mid \neg g)$. In this case, such an update amounts to changing $\mathrm{Pr}^{\boldsymbol{\pi}}(\mathsf{e}_T, \mathsf{e}_{D^1}, \mathsf{e}_{D^2})$ to $\mathrm{Pr}^{\boldsymbol{\pi}}(\mathsf{e}_T, \mathsf{e}_{D^1}, \mathsf{e}_{D^2} \mid \neg g)$, where note that we independently update each $\mathsf{E}_D^{\pi_D, \overline{\mathbf{pa}}_D}$ for every such $\boldsymbol{\pi}$ and endogenous decision context $\overline{\mathbf{pa}}_D$. Following this, we apply the intervention $\mathrm{do}(D^1 = g)$. The final answer to the query is therefore rather simple in this case, and is given by $\mathrm{Pr}^{\mathcal{R}}(u_g^1 \mid \neg g) = \left\{ \mathrm{Pr}^{\boldsymbol{\pi}}(u_g^1 \mid \neg g) \right\}_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M} \mid \neg g)}$. We thus have $\left\{ \mathbb{E}_{\boldsymbol{\pi}}[U_g^1 \mid \neg g] \right\}_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M} \mid \neg g)} = \left\{ \frac{7}{2} \right\}$.

Query 3b involves a hard pre-policy intervention $\mathcal{I}$ that sets Π_{D^1} to $\overline{\pi}_{D^1}$. As $\boldsymbol{\Pi}(\mathcal{I}) = \varnothing$, the answers to the query are given under the interventional outcomes, i.e., $\mathcal{R}(\mathsf{m}\mathcal{M}_{\mathcal{I}} \mid \tilde{\pi}_{D^1}) = \left\{ (\boldsymbol{\pi}', \boldsymbol{\pi}') : \boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M}_{\mathcal{I}}) \right\}$. In this particular game, $\mathcal{R}(\mathsf{m}\mathcal{M}_{\mathcal{I}}) = \mathcal{R}(\mathsf{m}\mathcal{M} \mid \overline{\pi}_{D^1})$ and so the answer is the same as the answer to query 1b.

Remark 3. *The reasons for choosing between a CG or an SCG to model a given problem are analogous to the respective reasons for choosing a CBN or an SCM. Using the latter, one can reason about counterfactuals, path-specific effects, and questions of identifiability [5]. The*

---

[17]In practice, for queries 3a and 3b, it happens to be the case that both principles lead to the same answer.

[18]Note that the distribution over these 'fresh' exogenous variables does not depend on the policy, actual or counterfactual.

former, however, is a simpler formalism, and requires less knowledge about the precise functions holding between the variables, making CGs a more attractive choice when one does not require any of the features listed above.

5 Solution Concepts and Subgames

In Section 3, we explained how a set of rationality relations can be used to capture the process by which agents choose their decision rules, and thus which mechanisms agents need to consider when doing so. In this section, we build on these ideas in three subsections. Firstly, we detail the distinction between mixed and behavioural policies and their relation to NEs in MAIDs. Secondly, we introduce the concept of *subgames* within MAIDs, which, analogously to their EFG counterparts, allow us to analyse and solve parts of the game independently. Finally, we introduce several *equilibrium refinements* for MAIDs, which are discussed in relation to their EFG counterparts in Section 6.2. With these contributions, we aim to place causal games on a more equal footing with EFGs as a tool for game-theoretic analysis. Note that as (S)CGs are refinements of MAIDs, the results in this section also apply to these models. Concepts in this section will be explained with the help of the following example, shown in Figure 8a.

Example 2 (Warehouse Robots). *Two robots are working together in a warehouse. Robot one is responsible for fulfilling orders; it can decide to move quickly or slowly. It will not break anything if it moves slowly, but might break something if it moves quickly. Robot two is responsible for keeping the warehouse tidy and so must decide whether to patrol or not, but it can only observe what robot one does. If it patrols, robot two can repair broken items, but by doing so it might obstruct robot one and prevent it from completing its order. Robot one is rewarded for fulfilling orders, the quicker the better, and penalised for breaking things. Robot two is rewarded for everything in the warehouse ending up in a state of repair, but incurs a small energy cost for patrolling.*

To parameterise this game, let us suppose that robot one breaks something if it moves quickly ($D^1 = q$) with probability $\frac{1}{3}$ but will not break anything otherwise, and that robot two obstructs robot one with probability $\frac{1}{2}$ if it patrols ($D^2 = p$) and probability zero otherwise. Finally, we define the utility functions such that robot one receives a reward of 5 or 2 for completing an order quickly or slowly respectively, but it also incurs a penalty of -3 for breakages. Robot two receives a reward of 6 for everything being in a state of repair, but incurs a penalty of -1 for patrolling. Given this parameterisation, we can easily calculate the expected payoff of each agent for the four possible action combinations. For reference, we show these in Figure 8c using an EFG that only bifurcates on the two robots' decisions; the chance variable B has been marginalised out to create a reduced EFG (as detailed in Appendix 6.1).

5.1 Nash Equilibria

A solution concept aims to identify a subset of the possible outcomes of a game that may occur if agents act rationally. In non-cooperative games, the most fundamental solution concept is a Nash equilibrium (NE) [68], a policy profile such that no agent may benefit by unilaterally deviating. In Example 2, for instance, the policy profile $\boldsymbol{\pi}^{\text{NE}}$ in which robot one chooses $D_1 = q$ and robot two chooses $D_2 = p$ whatever the value of D_1 is an NE. Previous work introduced this concept to MAIDs, as recalled in Definition 11 [51], but did not fully characterise when an NE is guaranteed to exist in a MAID. This existence depends on which class of policies agents are permitted to choose from.

So far in this paper, we have viewed agents as employing *behavioural policies*, where each agent selects decision rules for each of their decisions independently. In contrast, a *mixed policy* allows an agent to coordinate their choice of decision rules at different decisions; it is a distribution
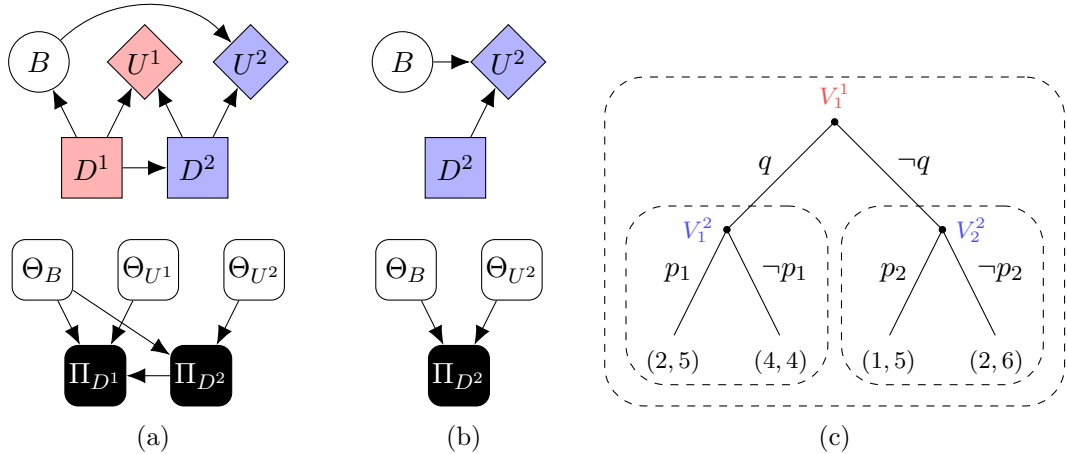
Figure 8: (a) A MAID $\mathcal{M} = (\mathcal{G}, \theta)$ representing Example 2 and its s-relevance graph. (b) The smallest proper s-subdiagram $\mathcal{G}'$ of $\mathcal{G}$ and its s-relevance graph. (c) A reduced EFG where each of the proper subgames corresponds to the two possible s-subgames for $\mathcal{G}'$, where $D^1 = q$ or $D^1 = \neg q$ respectively.

over pure policies. In what follows, we use a dot $\dot{}$ to denote the determinism of pure policies and μ to denote mixed policies.

Definition 24. Let $dom(\dot{\Pi}_D)$ be the set of all possible pure decision rules for D, and recall that we write $dom(\boldsymbol{V}) = \times_{V \in \boldsymbol{V}} dom(V)$. A **mixed policy** for agent i is some $\mu^i \in \Delta(dom(\dot{\Pi}_{\boldsymbol{D}^i}))$, a **behavioural policy** is some $\boldsymbol{\pi}^i \in dom(\boldsymbol{\Pi}_{\boldsymbol{D}^i})$, and a **pure policy** is some $\dot{\boldsymbol{\pi}}^i \in dom(\dot{\boldsymbol{\Pi}}_{\boldsymbol{D}^i})$.

Proposition 5. *A (behavioural policy) NE is not guaranteed to exist in a MAID.*

Nash's theorem establishes that any finite game is guaranteed to have an NE, as long as all agents are allowed to choose mixed policies [68]. However, in general, there is no such guarantee when agents are only permitted to use behavioural policies (for an example, see Appendix B.2). Despite this negative result, behavioural policies are often regarded as more 'natural' due to their representation of agents randomising at each decision point instead of once at the beginning of the game [72]. Moreover, behavioural policies respect the conditional independencies of the MAID's graph. As such, an interesting question is when an NE in behavioural policies is guaranteed to exist. As a first step, it is relatively straightforward to prove an analogue of Kuhn's theorem: if all agents in the game have perfect recall (i.e., agents never forget their past moves nor any of the information they knew previously, as introduced in Definition 13), then an NE in behavioural policies is guaranteed to exist [54].

Lemma 2. *Let $\boldsymbol{\pi}^{-i} \in \Delta(dom(\boldsymbol{D}^{-i}) \mid dom(\boldsymbol{Pa}_{\boldsymbol{D}^{-i}}))$ be a partial (behavioural or mixed) policy profile for agents $N \setminus \{i\}$ in a MAID $\mathcal{M}$. If agent i has perfect recall in $\mathcal{M}$, then for every mixed policy μ^i there exists a behavioural policy $\boldsymbol{\pi}^i$ such that $\Pr^{(\mu^i, \boldsymbol{\pi}^{-i})}(\boldsymbol{v}) = \Pr^{(\boldsymbol{\pi}^i, \boldsymbol{\pi}^{-i})}(\boldsymbol{v})$.*

Proposition 6. *In any MAID $\mathcal{M}$ with perfect recall, there exists a (behavioural) policy profile $\boldsymbol{\pi}$ that is an NE.*

Going further, because MAIDs reveal conditional independencies between variables, only a weaker criterion of *sufficient recall* is sufficient for the existence of an NE in behavioural policies. K&M implicitly prove this result – included as Proposition 2 – when proving that their Algorithm for finding an NE always converges under certain conditions (as these correspond with sufficient recall) [51]. Appendix B.2 provides an example of a MAID in which all agents have sufficient but imperfect recall. In this example, there exists a mixed policy profile with no behavioural policy resulting in the same distribution over outcomes. There is still, however, an

NE in behavioural policies. On the other hand, a MAID without sufficient recall may *not* have an NE in behavioural policies (an example is again given in Appendix B.2).

Proposition 7. *If an agent i in a MAID $\mathcal{M}$ has perfect recall, then they also have sufficient recall. However, if an agent has sufficient recall, then they do not always have perfect recall.*

5.2 Subgames

Subgames in EFGs represent parts of the game that can be solved independently from the rest; we now introduce the analogous notion of $\mathcal{R}$-*subgames* in MAIDs. These subgames have three uses: they allow us to introduce further equilibrium refinements (in Section 5.3); they can reduce the cost of computing equilibria [42]; and they allow us to analyse agents' decision-making in isolation from other parts of the game, which may be useful for other forms of analysis (such as those discussed in Section 7). In Appendix C.2, we demonstrate the second of these facts by empirically showing how subgames in MAIDs can be used to compute NEs much more efficiently than in EFGs.

However, there are two key differences between subgames in EFGs and those in MAIDs. Firstly, because MAIDs explicitly represent conditional independencies between variables, we can often find more subgames in a MAID than in its corresponding EFG. Secondly, because the ability to solve part of a game independently varies with the solution concept, the $\mathcal{R}$-subgames vary with $\mathcal{R}$. Given a set of graphical criteria, $\mathcal{R}$-reachability, such as those in Section 3.2, one can identify the structure of any $\mathcal{R}$-subgames – which we refer to as an $\mathcal{R}$-subdiagram – using only the underlying graph.

Definition 25. Given a mechanised MAID $\mathsf{m}\mathcal{M} = (\mathsf{m}\mathcal{G}, \boldsymbol{\theta}, \mathcal{R})$ and a set of sound and complete graphical criteria for $\mathcal{R}$-relevance – ie., $\mathcal{R}$-reachability – we refer to the subgraph $(\boldsymbol{V}', \mathscr{E}')$ of $\mathcal{G}$, along with the set of agents $N' \subseteq N$ possessing decision variables in that subgraph, as an $\mathcal{R}$-**subdiagram** $\mathcal{G}' = (N', \boldsymbol{V}', \mathscr{E}')$ if:

- $\boldsymbol{V}'$ contains every variable Z such that M_Z is $\mathcal{R}$-reachable from some Π_D with $D \in \boldsymbol{V}'$;

- $\boldsymbol{V}'$ contains, for all $X, Y \in \boldsymbol{V}'$, every variable that lies on a directed path $X \dashrightarrow Y$ in $\mathcal{G}$.

As before, we drop $\mathcal{R}$ from our notation where unimportant or unambiguous.

The first condition on $\boldsymbol{V}'$ ensures that for any decision variable D in the subdiagram, any variable whose mechanism may impact the rational response for D is also included in the graph. This means that the rational responses for the decision rules in this part of the game are independent of what happens elsewhere. The second condition says that additional variables may also be included in the subdiagram as long as mediators are included too. This ensures that the CPDs for all the variables in the subgame remain consistent, and allows us to define a correspondence between subgames in MAIDs and subgames in EFGs (in Section 6.2). To find the subgames for each subdiagram, we must update the parameterisation of the remaining variables to be consistent with the original game and the structure of the graph.

Definition 26. Given a mechanised MAID $\mathsf{m}\mathcal{M} = (\mathsf{m}\mathcal{G}, \boldsymbol{\theta}, \mathcal{R})$, an $\mathcal{R}$-**subgame** of $\mathcal{M}$ is a new MAID $\mathcal{M}' = (\mathcal{G}', \boldsymbol{\theta}')$ where $\mathcal{G}'$ is an $\mathcal{R}$-subdiagram of $\mathcal{G}$ and $\boldsymbol{\theta}'$ is defined by $\Pr'(\boldsymbol{v}'; \boldsymbol{\theta}') \coloneqq \Pr(\boldsymbol{v}' \mid \boldsymbol{z}; \boldsymbol{\theta})$, where $\boldsymbol{z}$ is some instantiation of the variables $\boldsymbol{Z} = \boldsymbol{V} \setminus \boldsymbol{V}'$.[19] An $\mathcal{R}$-subgame is **feasible** if there exists a policy profile $\boldsymbol{\pi}$ where $\Pr^{\boldsymbol{\pi}}(\boldsymbol{z}) > 0$.

$\mathcal{R}$-subgames can be found for any $\mathcal{R}$ using only $\mathcal{R}$-reachability. In particular, s-reachability produces s-subgames, which in many ways are the most natural form of subgame in MAIDs (because of their connection to subgames in EFGs, as shown formally in Section 6.2). In the

---

[19]In fact, it can easily be appreciated that only the setting $\boldsymbol{z}$ of the variables that have a child in $\boldsymbol{V}'$ will matter.
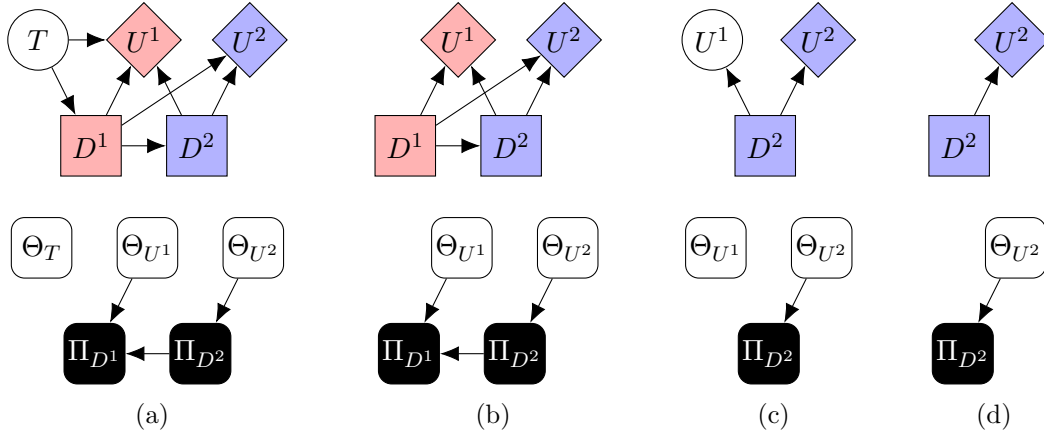
Figure 9: (a) A MAID $\mathcal{M} = (\mathcal{G}, \theta)$ for the modified version of Example 1 – in which the firm's profits are also function of the worker's decision but not of their temperament – and resulting s-relevance graph. The graph $\mathcal{G}$ is also an (improper) s-subdiagram and the full game an (improper) s-subgame. Figures (b), (c), and (d) illustrate the remaining (proper) s-subdiagrams of $\mathcal{G}$ and their s-relevance graphs.

remainder of the paper, unless otherwise specified, we therefore focus our attention on this case. Note also that any MAID is an $\mathcal{R}$-subgame of itself, and so an $\mathcal{R}$-subgame on a strictly smaller set of variables is called a *proper* $\mathcal{R}$-subgame. For example, the MAID for Example 1 (shown in Figure 3b) has no proper $\mathcal{R}^{\mathrm{BR}}$-subgames because Π_{D^1} and Π_{D^2} are both $\mathcal{R}^{\mathrm{BR}}$-reachable from one another.

Identifying More Subgames in MAIDs

Before continuing, we note that the conditional dependencies captured in MAIDs allow for a richer and stronger notion of subgames than in EFGs. Not only can different notions of $\mathcal{R}$-subgame be introduced for different rationality relations, but it is often possible to identify more subgames (and hence rule out more non-credible threats) in a MAID than in the corresponding EFG. This can be seen by considering a minor variation on Example 1. Suppose that the firm has a new vacancy in which what is important is not whether the worker is hard-working or lazy, but whether they have studied at university. In other words, U^2 is a function of D^1 and D^2 but not T.

The game graph and s-relevance graph for this example are shown in Figure 9a. Note that the *structure* of the EFG (shown in Figure 3a) does not change, only the payoffs for the firm. As such, there are no proper subgames in the EFG because when the hiring system is making its decision (D^2), it cannot observe the value of T and so no proper subtree is closed under both descendants and information sets. In contrast, we can recognise three proper s-subdiagrams (shown in Figures 9b, 9c, and 9d) of the equivalent MAID. Each of these s-subdiagrams has two s-subgames associated with it owing to the two values that T can take (for the s-subdiagram in Figure 9b) and the two values that D^1 can take (for the s-subdiagrams in Figures 9c and 9d).

5.3 Equilibrium Refinements

When more than one NE exists, it is useful to specify additional criteria to rule out less plausible outcomes. This corresponds to making additional assumptions about the rationality of agents, which can be encoded as stricter rationality relations. Below, we provide definitions of two of the most important equilibrium refinements – subgame perfect equilibria [85] and trembling hand perfect equilibria [84] – within MAIDs. Later, in Section 6.2, we provide proofs regarding

various equivalences between these definitions and those in EFGs.

Subgame Perfect Equilibrium

The concept of a subgame perfect equilibrium (SPE) was introduced into EFGs in order to eliminate NEs containing *non-credible threats* – choices made by an agent in a sequential game that would not be in their best interest to carry out if given the opportunity [84, 85]. In MAIDs, we can rule out non-credible threats by ensuring that each agent plays a best response in every feasible s-subgame. In games with sufficient recall, an SPE (in behavioural policies) can always be constructed by performing backwards induction over the s-subgames and finding an NE in each; it is this technique in Appendix C.2 that allows an NE to be computed much more efficiently than in a corresponding EFG. However, if even one agent in the game has insufficient recall, an SPE may not exist even when allowing for mixed policies (see Appendix B.2 for an example).

Definition 27. A policy profile $\boldsymbol{\pi}$ in a MAID $\mathcal{M}$ is a **subgame perfect equilibrium (SPE)** if $\boldsymbol{\pi}$ is an NE in every feasible s-subgame of $\mathcal{M}$, when restricted to that subgame.[20]

Proposition 8. *Any MAID $\mathcal{M}$ with sufficient recall has at least one SPE in behavioural policies.*

Recall $\boldsymbol{\pi}^{\mathrm{NE}}$ in Example 2, introduced in Section 5.1, in which robot one chooses $D_1 = q$ and robot two chooses $D_2 = p$ whatever the value of D_1. We can immediately see that in the feasible s-subgame where $D_1 = \neg q$ – the s-subdiagram for the smallest such s-subgame is shown in Figure 8b – then choosing $D_2 = p$ is a non-credible threat, resulting in expected utility 5 instead of 6 for robot two. Instead, an example of an SPE is the policy profile $\boldsymbol{\pi}^{\mathrm{SPE}}$ in which robot one chooses $D_1 = \neg q$, and robot two chooses $D_2 = p$ if and only if $D_1 = q$. Under such a policy profile, robot two achieves its optimal expected utility in any of the feasible s-subgames, and given that robot two is following this policy, robot two receives expected utility 2 regardless of whether they move quickly or not.

Before continuing, we note that rationality relations allow us to capture arbitrary sets of policy profiles as rational outcomes, including the equilibrium refinements in this section. For example, (when each agent has at most one decision) the rational outcomes $\mathcal{R}^{\mathrm{BR}}(\mathsf{m}\mathcal{M})$ are simply the NEs of $\mathcal{M}$, as can easily be seen via inspection of Equation (1) and Definition 11. Similarly, for a MAID $\mathcal{M}$ let us denote by $\mathcal{M}(D)$ the set of all feasible s-subgames containing D, and define:

$$\pi_D \in r_D^{\mathrm{SP}}(\mathbf{pa}_{\Pi_D}) \quad \Leftrightarrow \quad \pi_D \in \underset{\hat{\pi}_D \in dom(\Pi_D)}{\arg\max} \sum_{U \in \boldsymbol{U}^i \cap \boldsymbol{V}'} \mathbb{E}_{(\hat{\pi}_D, \boldsymbol{\pi}'_{-D})}[U] \; \forall \; \mathcal{M}' \in \mathcal{M}(D),$$

expressing that each agent plays a best response for their decision rule in every s-subgame containing that decision. In other words, $\boldsymbol{\pi} \in \mathcal{R}^{\mathrm{SP}}(\mathsf{m}\mathcal{M})$ if $\boldsymbol{\pi}' \in \mathcal{R}^{\mathrm{BR}}(\mathsf{m}\mathcal{M}')$ for each s-subgame $\mathcal{M}'$ of $\mathcal{M}$, where $\boldsymbol{\pi}'$ is $\boldsymbol{\pi}$ restricted to the decision variables in $\mathcal{M}'$. If $\mathcal{M}$ has sufficient recall, $\mathcal{R}^{\mathrm{SP}}(\mathsf{m}\mathcal{M})$ are the SPEs of the game, as stated formally below. While such representations may sometimes be slightly more cumbersome, encoding equilibria via rationality relations facilitates the use of $\mathcal{R}$-relevance and hence $\mathcal{R}$-reachability. This offers a principled way to identify independencies that can be useful both for causal and game-theoretic reasoning.

Proposition 9. *Suppose that a MAID $\mathcal{M}$ has sufficient recall. Then, the set of SPEs of $\mathcal{M}$ is equal to the set of rational outcomes $\mathcal{R}^{SP}(\mathsf{m}\mathcal{M})$.*

Trembling Hand Perfect Equilibrium

In an SPE, agents make decisions on the assumption that an SPE will be played in all proper subgames. As a result, however, the optimality of their strategies may not be robust to events

---

[20] Note that this notion of subgame perfectness can be generalised to other choices of rationality relations.

in which other agents make mistakes, or 'tremble', with some small probability. To solve this problem, we can stipulate that each agent must play a best response (leading to an NE) in each *perturbed game* [84]. Let ζ_k be a perturbation vector containing, for every $D \in \boldsymbol{D}$, $d \in dom(D)$, and decision context $\mathbf{pa}_D$, a value $\epsilon_d^{\mathbf{pa}_D} \in (0,1)$ such that $\sum_{d \in dom(D)} \epsilon_d^{\mathbf{pa}_D} \leq 1$. Then, given a game $\mathcal{M}$, the perturbed game $\mathcal{M}(\zeta_k)$ is defined such that each decision rule π_D is forced to have $\pi_D(d \mid \mathbf{pa}_D) \geq \epsilon_d^{\mathbf{pa}_D}$.

Definition 28. A policy profile $\boldsymbol{\pi}$ is a **trembling hand perfect equilibrium (THPE)** in a MAID $\mathcal{M}$ if there is a sequence of perturbation vectors $\{\zeta_k\}_{k \in \mathbb{N}}$ such that $\lim_{k \to \infty} \|\zeta_k\|_\infty = 0$ and for each perturbed MAID $\mathcal{M}(\zeta_k)$ there is an NE $\boldsymbol{\pi}_k$ such that $\lim_{k \to \infty} \boldsymbol{\pi}_k = \boldsymbol{\pi}$.

For example, the policy profile $\boldsymbol{\pi}^{\text{SPE}}$ is *not* a THPE. To see this, suppose that robot two trembles with probability $\epsilon > 0$ when $D_1 = q$ and $\epsilon' > 0$ when $D_1 = \neg q$. Then, robot one's expected utility when playing q is $2 + 2\epsilon$ and when playing $\neg q$ is $2 - \epsilon'$. Therefore in any NE of the perturbed game robot one will play q with probability one. One THPE is given by the policy profile $\boldsymbol{\pi}^{\text{THPE}}$ in which robot one chooses $D_1 = q$ and robot two chooses $D_2 = p$ if and only if $D_1 = q$. Note that in any two-agent game, THPEs rule out all weakly dominated policies. In this example, robot one's policy of always choosing $D^1 = \neg q$ is weakly dominated by the policy of always choosing $D^1 = q$ (which itself is not weakly dominated by any other policy).

6 Connections to EFGs

EFGs are perhaps the most widely studied model of dynamic strategic decision-making. Despite their intuitive appeal, these tree-based models can be less concise and reveal less of the underlying structure of a game than the DAG-based models we use in this paper. With that said, a natural question is whether the game-theoretic definitions for MAIDs in Section 5 successfully capture the familiar concepts of their EFG counterparts, or whether we might lose something by working with MAIDs – and hence (S)CGs – instead.

In this section, we show that such concerns are unwarranted: these game-theoretic concepts are preserved when we convert between representations. We begin by briefly describing two algorithms, `maid2efg` and `efg2maid`, that implement these conversions. Using these procedures, we then provide equivalence results for the definitions from Section 5. Following this, we briefly discuss how tree-based models can also be used for causal reasoning in certain cases.

6.1 Transformations

We briefly summarise two procedures for converting between MAIDs and EFGs. A more formal treatment of both can be found in Appendix A.1.

From MAID to EFG

There are many ways to convert a MAID $\mathcal{M} = (\mathcal{G}, \boldsymbol{\theta})$ into an EFG $\mathcal{E}$, but these differ in their computational costs [51, 74]. The basic idea, as employed by K&M, is to use a topological ordering $\prec$ over the variables of $\mathcal{G}$ to construct $\mathcal{E}$'s game tree by splitting on each of the variables in order. This splitting is required due to the fact that a variable in $\mathcal{G}$ defines what happens given any context (a setting of the parent variables), whereas a node in $\mathcal{E}$ defines what happens only in one context (the path taken to reach that node). As such, we may end up with exponentially more nodes in $\mathcal{E}$ than variables in $\mathcal{G}$.

By querying the CPDs of each variable, branches from chance nodes are labelled with probabilities based on the path taken from the root of the tree to that node, and similarly for the utilities assigned to each leaf; branches from decision nodes are labelled with the possible actions available. Given two nodes corresponding to the same decision variable D in $\mathcal{G}$, we assign them

to the same information set if and only if the values taken by their ancestors along the paths from the root of the tree to each node agree on all those variables in $\mathbf{Pa}_D$. Note that because there can be more than one topological ordering $\prec$, we regard the output of `maid2efg` as a *set* of EFGs – one for each possible ordering.

Remark 4. *This procedure can be made more efficient by marginalising out all variables not in $\mathbf{U} \cup \mathbf{Fa_D}$, such as in the reduced EFG for Example 2 (shown in Figure 8c), which only has 2^2 leaves, as opposed to the 2^3 leaves that would have resulted had we retained all the variables in the original MAID (in Figure 8a). Importantly, the information in this reduced EFG is sufficient for computing its equilibria, and can thus offer significant efficiency gains (since the cost of solving an EFG depends on its size, which is exponential in the length of $\prec$). In our codebase [27], we therefore implement this more efficient transformation (originally described by K&M), which can be used in conjunction with Gambit, a popular tool for solving EFGs [63].*

From EFG to MAID

By encoding the CPDs for each variable in the MAID using trees as opposed to tables, MAIDs can represent any game using at most the same (but often exponentially less) space than an EFG [51]. In general, there are many MAIDs that can represent a given EFG. For instance, upon converting the EFG representation (in Figure 3a) of Example 1 to a MAID, we could naïvely create a decision variable for each information set. Alternatively, we could recognise, for example, that V_1^1 and V_2^1 correspond to the same real-world variable – the worker's choice to go to university or not – and thus combine them (as shown in Figure 3b, in which the two information sets for the second agent have also been combined).

Whether two nodes represent the same variable in an EFG is a matter of domain knowledge external to the EFG, but this knowledge typically exists when creating a model of a game and is then lost by viewing the interaction as an EFG. By formalising this notion as a question of whether nodes belong to the same *intervention set*, we provide a procedure `efg2maid` (described fully in Appendix A.1), which maps an EFG to a *unique*, canonical MAID.

Definition 29. An **intervention set** J in an EFG $\mathcal{E}$ is either a set of chance nodes, information sets belonging to the same agent, or leaves, such that:

- Every node $V \in J$ or $V \in I_J^i \in J$ has the same number of children;
- No path from the root of $\mathcal{E}$ to one of its leaves passes through J more than once.

Moreover, for a valid partition of information sets and chance variables into intervention sets, we require that any path to a node $V \in J$ or $V \in I_J^i \in J$ passes through the same intervention sets before reaching J.

In `efg2maid`, we assume that intervention sets are given; in the absence of such knowledge, one can simply choose each intervention set to be a singleton. The resulting MAID will remain equivalent to the EFG (in the sense described in the following subsection), but may not be as compact as it otherwise could be. The basic idea behind the conversion is to first view the game tree as a MAID, then to add missing edges to each node from its ancestors whilst merging nodes that are in the same intervention set into single variables. Incoming edges from any variable V whose value is not observed by a decision D are then removed from $\mathbf{Pa}_D$, along with any duplicate edges produced by merging nodes. Each leaf intervention set is split into one utility variable for each agent. The distributions over each variable V are then formed for each context $\mathbf{pa}_V$ by summing the relevant probabilities of sets of merged edges.

6.2 Equivalences

Using these transformations, we derive equivalence results between EFGs and MAIDs to demonstrate that the fundamental game-theoretic notions of subgames and various equilibrium refinements in Section 5 are *preserved* when converting between representations. We begin by defining what we mean for two game representations to be 'equivalent'; the underlying idea is that the (behavioural) policy and strategy spaces in each game should be the same and each corresponding (behavioural) policy and strategy profile should lead to the same expected utility for each agent in both games. An immediate consequence of this is that NEs are also preserved.

Definition 30. We say that a MAID $\mathcal{M}$ is **equivalent** to an EFG $\mathcal{E}$ (and vice versa) if there is a bijection for each agent $f^i : \Sigma^i \to dom(\mathbf{\Pi}^i)/\sim$ between their strategies in $\mathcal{E}$ and a partition of their policies in $\mathcal{M}$ (the quotient set of $dom(\mathbf{\Pi}^i)$ by an equivalence relation $\sim$ where $\boldsymbol{\pi}^i \sim \hat{\boldsymbol{\pi}}^i$ if and only if $\boldsymbol{\pi}^i$ and $\hat{\boldsymbol{\pi}}^i$ differ only on infeasible decision contexts) such that for every $\boldsymbol{\pi} \in f(\sigma)$ and every agent i, we have $\mathbb{E}_\sigma\big[U(\rho[\boldsymbol{L}])[i]\big] = \sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{\boldsymbol{\pi}}[U]$, where $f(\sigma) := \bigtimes_{i \in N} f^i(\sigma^i)$. We refer to any such f as a **natural mapping** between $\mathcal{E}$ and $\mathcal{M}$.

Lemma 3. *Let f be a natural mapping between $\mathcal{E}$ and $\mathcal{M}$. Then σ is an NE in $\mathcal{E}$ if and only if every $\boldsymbol{\pi} \in f(\sigma)$ is an NE in $\mathcal{M}$.*

The reason we use an equivalence relation on the space of policies is that `efg2maid` can introduce additional infeasible decision contexts: those $\mathbf{pa}_D$ such that $\Pr^{\boldsymbol{\pi}}(\mathbf{pa}_D) = 0$ for all $\boldsymbol{\pi}$, corresponding to paths in the EFG that do not exist. An agent's choice in such decision contexts has no bearing on the outcome of the game, meaning we may safely abstract away from them. Any natural mapping is therefore sufficient for preserving the essential game-theoretic features of each representation, as we show below, though it may not be unique. We begin with a supporting lemma that justifies the correctness of the procedures `maid2efg` and `efg2maid`, and forms the basis of our other results.

Lemma 4. *If $\mathcal{E} \in$ `maid2efg`$(\mathcal{M})$ or $\mathcal{M} =$ `efg2maid`$(\mathcal{E})$, then $\mathcal{E}$ and $\mathcal{M}$ are equivalent.*

This lemma follows directly from the construction of a natural mapping f using the two procedures, `maid2efg` and `efg2maid` respectively. The intuition is that the information sets in an EFG correspond to the feasible decision contexts in a MAID, and thus a behavioural strategy profile σ in the EFG corresponds to a behavioural policy profile $\boldsymbol{\pi}$ in the MAID, and vice versa. The following result is a direct corollary of Lemma 3 and Lemma 4.

Corollary 1. *If $\mathcal{E} \in$ `maid2efg`$(\mathcal{M})$ or $\mathcal{M} =$ `efg2maid`$(\mathcal{E})$, then there is a natural mapping f between $\mathcal{E}$ and $\mathcal{M}$ such that σ is an NE in $\mathcal{E}$ if and only if every $\boldsymbol{\pi} \in f(\sigma)$ is an NE in $\mathcal{M}$.*

For a subgame $\mathcal{E}'$ in an EFG, it can be shown that the variables outside $\mathcal{E}'$ are not s-relevant to those in the corresponding s-subgame $\mathcal{M}'$. This means that subgames in EFGs have equivalent counterparts in their equivalent MAID, as established by the following proposition. A minor caveat here is that some utility variables may not occur in $\mathcal{M}'$, and so we must add their value (under the setting of the variables outside $\mathcal{G}'$ that defines $\mathcal{M}'$) to the payoff for the agents in order to equate their expected utilities with that of $\mathcal{E}'$. For any subgame $\mathcal{M}'$, however, this change in value is constant for each agent under any policy profile, and so has no effect on their decisions.

Proposition 10. *If $\mathcal{E} \in$ `maid2efg`$(\mathcal{M})$ or $\mathcal{M} =$ `efg2maid`$(\mathcal{E})$, then there is a natural mapping f between $\mathcal{E}$ and $\mathcal{M}$ such that, for every subgame $\mathcal{E}'$ in $\mathcal{E}$ there is an s-subgame $\mathcal{M}'$ in $\mathcal{M}$ that is equivalent (modulo a constant difference between the utilities for each agent under any policy in $\mathcal{M}'$) to $\mathcal{E}'$ under the natural mapping f restricted to the strategies of $\mathcal{E}'$.*

This restriction of f to the strategies in $\mathcal{E}'$ can be made precise by considering only those feasible decision contexts that correspond to the information sets contained in $\mathcal{E}'$. By first

applying Proposition 10 and then Lemma 3 to each of the resulting subgames, we see that any SPE in $\mathcal{M}$ is carried over to $\mathcal{E}$. We note, however, that as there may be more s-subgames in a MAID than in its equivalent EFG, the criterion of subgame perfectness may be slightly stronger in the MAID. In other words, not all SPEs in an EFG may be SPEs in the equivalent MAID. This additional strength can be useful in ruling out non-credible threats even when they do not fall under a particular subgame in the EFG.

Corollary 2. *If $\mathcal{E} \in$ maid2efg$(\mathcal{M})$ or $\mathcal{M} =$ efg2maid$(\mathcal{E})$, then there is a natural mapping f between $\mathcal{E}$ and $\mathcal{M}$ such that if every $\boldsymbol{\pi} \in f(\sigma)$ is an SPE in $\mathcal{M}$, then σ is an SPE in $\mathcal{E}$.*

Finally, we derive an equivalence between the THPEs in EFGs and those in MAIDs. In order to do so, it suffices to prove an equivalence between perturbed versions of the corresponding games $\mathcal{E}(\zeta_k)$ and $\mathcal{M}(\zeta_k)$, which can easily be done by construction using maid2efg and efg2maid, and then by applying Lemma 4 and Lemma 3 to each of these perturbed games.

Proposition 11. *If $\mathcal{E} \in$ maid2efg$(\mathcal{M})$ or $\mathcal{M} =$ efg2maid$(\mathcal{E})$, then there is a natural mapping f between $\mathcal{E}$ and $\mathcal{M}$ such that σ is a THPE in $\mathcal{E}$ if and only if every $\boldsymbol{\pi} \in f(\sigma)$ is a THPE in $\mathcal{M}$.*

This series of equivalence results serves to justify MAIDs as an appropriate choice of game representation. Not only do they provide computational advantages over EFGs, they preserve some of the most fundamental game-theoretic concepts commonly employed in EFGs.

6.3 Causality in EFGs

One alternative to causal games would be to define causal concepts directly via EFGs. In this section, we explain how this can be done for a limited set of causal queries by extending methods designed for probability trees [32], and explain some shortcomings of this approach relative to the models we focus on in this paper.

Interventions

Intuitively, interventions in EFGs correspond to replacing the probability distribution(s) governing some node(s) in the game tree. More formally, consider an EFG $\mathcal{E}$ with intervention sets (if each intervention set is a singleton, we recover the case where each EFG variable is treated separately). As in CGs, we can apply both pre- and post-strategy interventions, and both operations have the same form. We make an intervention $\mathcal{I}$ on an intervention set J by replacing P_j or σ_j^i with an arbitrary probability distribution $P_j^*(V_j \mid \cap_{V_j \in J} \mathbf{Anc}_{V_j})$ or $\sigma_j^{i*}(V_j \mid \cap_{V_j \in J} \mathbf{Anc}_{V_j})$, for each $V_j \in J$. Hard interventions represent the special case when this distribution is given by $\delta(V_j, v_j)$, where v_j corresponds to an outgoing edge from each node $V_j \in J$. Pre-strategy interventions result from performing an intervention on the game tree *before* agents choose their strategies – in essence, forming a new game $\mathcal{E}_{\mathcal{I}}$ – whereas post-strategy interventions result from intervening on the probability tree that results *after* agents have chosen their strategies – in other words, forming a new distribution $P_{\mathcal{I}}^\sigma$ over paths in the tree.

While the graph of an EFG is typically viewed as encoding informational constraints upon the decisions of agents, as opposed to temporal or causal structure, it can also be given a causal interpretation in much the same way that a MAID or BN can, by restricting the EFG to be consistent with the set of all possible (hard) interventional distributions. This restriction produces a level two model in which the interventional queries described in the paragraph above are semantically meaningful. Due to the total ordering imposed by EFGs over the variables in a game, however, only a restricted subset of interventions are available (those whose distributions $\mathcal{I}$ condition only on *ancestors* of the node in question). Similarly, some queries (be they probabilistic or causal) over EFGs are not well-defined, due to the fact that, in general, not all variables are assigned a value on a path ρ from the root to a leaf.

Counterfactuals

Answering counterfactual queries is slightly more complicated. We briefly describe an existing procedure for probability trees, and refer the interested reader to this work for further details [32]. Suppose that we wish to calculate $P^\sigma(\boldsymbol{x}_\mathcal{I} \mid z)$ – the value that $\boldsymbol{X} = \boldsymbol{V} \setminus \{Y, Z\}$ would have taken if Y were distributed according to $\mathcal{I}$, given that in fact z is true. The first step is to condition on z by setting $P_Z(z) = 1$ or $\sigma_Z^i(z) = 1$, which is equivalent to first performing a hard intervention $\mathrm{do}(Z = z)$ and then renormalising the probability distributions over the branches *upstream* of this intervention. The second step is to perform the soft intervention $\mathcal{I}$ on Y, without any renormalising. Thirdly and finally, we restore the probability distributions over the branches *downstream* of this second intervention to the original settings given by P^σ. The resulting probability tree describes the *post-strategy* counterfactual distribution $P^\sigma(\boldsymbol{x}_\mathcal{I} \mid z)$. By modifying this procedure such that the normalisation upstream of Z and the restoring of distributions downstream of Y takes place according to the actual strategy profile σ, then allowing agents to play a new counterfactual strategy profile σ', we produce a *pre-strategy* counterfactual distribution (though if conditioning is performed downstream of differences between σ and σ', the resulting distribution will not, in general, be correct).

Counterfactual queries in EFGs are *not* invariant to the choice of tree representation, even if the distributions P^σ are the same for any strategy profile σ. This difficulty arises because standard EFGs do not reside at level three of the causal hierarchy, but may be overcome (as in CBNs and CGs) by relegating all stochasticity to a set of exogenous chance nodes that appear at the top of the game tree, and allowing agents only to select deterministic strategies (over endogenous decision nodes) as a function of some exogenous noise. This also resolves the difficulties of pre-policy counterfactuals in EFGs, as updates made due to conditioning on z will now be *upstream* of differences between σ and σ'. Existing work on probability trees [32], on the other hand, treats stochasticity as endogenous, and only possible to learn about (in the counterfactual world) for those variables that are *upstream* of the intervention $\mathcal{I}$. The ordering of variables in the tree therefore implicitly represents a substantive modelling assumption that resolves the subtle complexities surrounding counterfactuals in models with intrinsic stochasticity [17, 73].

7 Applications

Having described and analysed the theoretical features of causal games, we now consider potential domains of application. In the first subsection, we provide a case study based on the UK home insurance market. In the second, we discuss how several existing concepts for the analysis of (artificial) agents can be embedded and further developed within our framework. We provide worked examples of these embeddings in Appendix B.3. More substantive results, however, are beyond the scope of this paper, whose fundamental aim is to lay the theoretical foundations for causal games, rather than to explore specific applications.

7.1 Case Study: Insurance Pricing

While our primary motivation is to use causal games to analyse AI systems, another natural domain of application is in economics, where both causal and game-theoretic modelling are common. For example, when a regulator in the UK proposes a policy intervention, they are required to conduct a cost-benefit analysis to justify it. This entails making a comparison between the consequences of the intervention and the consequences of no intervention and/or alternative interventions. Some regulators, such as the Financial Conduct Authority (FCA), include a 'causal chain' in their analyses, though these are only heuristic diagrams rather than formal causal models. In this case study, we demonstrate the usefulness of causal games for economic analysis with an example based on the UK home insurance market, worth £6.31 billion

in 2021 [34], which was the focus of a recent FCA report [25].

The Model

In this setting: some customers are inert and do not change their insurance provider from year to year; firms charge a low price to attract customers in the first year, then increase their prices upon renewal; inert customers, therefore, end up paying more, while savvy customers spend time negotiating or switching provider each year. This allows firms to increase profits at customers' expense, either from increased prices (if inert), or switching costs (if savvy). These practices are especially troubling because it is often more vulnerable customers (such as older or less educated individuals) who are likely to be exploited [61]. While this is not the focus of the aforementioned FCA report, it is conceivable that the use of sophisticated AI algorithms to set insurance prices based on customer details might potentially lead to an even greater degree of exploitation. A simplified model of this setting is as follows.

Example 3 (Insurance Market). *A customer in a duopolistic insurance market must choose an insurance provider, and aims to minimise their costs. They already have a contract with one of the two providers, and so must decide whether to renew or switch. The customer is either savvy (in which case they are modelled as having a low switching cost) or inert (in which case they are modelled as having a high switching cost). The two firms use internal pricing algorithms to offer their quotes simultaneously, with the knowledge of which firm the customer has an existing contract with, but without the knowledge of the customer's type (savvy or inert), in order to maximise their profits.*

We represent this setting as a causal game $\mathcal{M}$ with three agents (in Figure 10a) and parameterise the model using real-world data on pricing strategies and consumer tendencies in the UK home insurance market within the period 2013-18 [25, 26, 61]. The two firms (agents one and two) use their pricing algorithms to decide prices $d^1, d^2 \in \mathbb{N}$ for their insurance premiums based on which firm the customer (agent three) currently has a contract with, $c \in \{1, 2\}$.[21] The customer uses this information, along with their type T – savvy (s) or inert ($\neg s$) – to choose a firm $d^3 \in \{1, 2\}$ to purchase insurance from, or to exit the market ($\perp$).

Firm i's utility (their profit) is assumed to be given by their policy price d^i minus the marginal cost of providing the policy, if selected by the customer, and zero otherwise. The customer's utility is defined as their valuation of the policy, minus the price they pay and switching costs, with an extra $-\infty$ term if the customer is inert (i.e., to model the searching and switching cost being prohibitively high). The customer's utility is zero if they exit the market. The chance variables C (the customer's current firm) and T (the customer's type) are parameterised following the values in Table 10b.

The Proposed Intervention

In the aforementioned report, the FCA proposes to ban insurance companies from setting different prices for new and existing customers [25]. Because firms are made aware of the intervention before deploying their pricing algorithms, we represent this as a pre-policy intervention on Π_{D^1} and Π_{D^2}. More concretely, this corresponds to changing the rationality relations r_{D^1} and r_{D^2} by restricting their codomains – $dom(\Pi_1)$ and $dom(\Pi_2)$ – to be such that $\pi_{D^1}(D^1 \mid C) = \pi_{D^1}(D^1)$ and $\pi_{D^2}(D^2 \mid C) = \pi_{D^2}(D^2)$, effectively removing the edges $C \to D^1$ and $C \to D^2$ (shown as

---

[21]In practice, the firms' prices will be bounded between the marginal cost of supplying the policy, and the point at which the customer exits the market due to prices being too high.

[21]This quantity is estimated as the average amount received per policy by firms (£231) minus the average profit per policy (£25) [25]. In practice, there are other small fees that affect a firm's revenue.

[22]Taken as those who self-report as 'actively' searching for new policies [25].
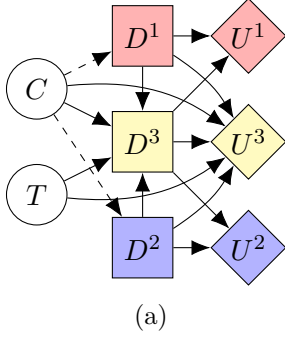
Average Quantity	Amount
Marginal cost of supplying policy[22]	£206
Customer valuation of policy	£296
Switching cost for customer	£38
Savvy customers[23]	72%
Customers currently with first firm	p%

(a) (b)

Figure 10: (a) A causal game representing Example 3. (b) Data from the UK home insurance market within the period 2013-18 [25, 26, 61], used to parameterise the game.

dashed in Figure 10a). In what follows we denote this intervention by $\mathcal{I}$. The FCA offers a number of hypotheses about the likely effects of this intervention, namely that:

i) The percentage of customers switching will reduce;

ii) Switching customers will end up paying higher prices, and renewing customers will end up paying lower prices;

iii) Overall, customers will end up better off on average.

Representing the intervention within a causal game $\mathcal{M}$ allows us to express these hypotheses formally by measuring and quantifying the causal effects of $\mathcal{I}$, which cannot, in general, be done without a formal causal model. For the purpose of this analysis, we set $p = 0.6$ and assume that agents play a pure THPE, which, while not truly realistic, illustrates the real-world applications of causal games (which can also be used to capture boundedly rational agents). In what follows, we denote the rationality relations describing this assumption as $\mathcal{R}$.

Before addressing the three hypotheses, it is a simple exercise to see that for each $\boldsymbol{\pi} \in \mathcal{R}(\mathcal{M})$, firms do indeed end up setting minimal prices (i.e., equal to their marginal cost) for new customers and higher prices for renewing customers, up to the point of driving away sales due to switching costs from customers. In other words, we have that $\pi_{D^i}(D^i \mid C = 3 - i) = \delta(D^i, 206)$ and $\pi_{D^i}(D^i \mid C = i) = \delta(D^i, 244)$ for each $i \in \{1, 2\}$. For their part, customers always seek the lowest price, taking into account switching costs, and are indifferent when the price at an alternative firm is the same as the price at their current firm plus the switching cost. After the intervention $\mathcal{I}$, the customer's policy stays the same, but we have $\pi_{D^1}(D^1 \mid C) = \delta(D^1, 288)$ and $\pi_{D^2}(D^2 \mid C) = \delta(D^2, 250)$ for the two pricing algorithms, respectively, for any $\boldsymbol{\pi} \in \mathcal{R}(\mathcal{M}_{\mathcal{I}})$.

The rational outcomes of the original game correspond to whether savvy customers (at each of the two firms) switch to the other firm or not. Depending on whether customers with neither firm, the first firm, the second firm, or both firms switch, then we have 0%, 43.2%, 28.8%, or 72% of customers switching, respectively. After intervention $\mathcal{I}$, customers already with the second firm have no incentive to switch, and so the percentage of customers switching is either 0% or 43.2%. If we were to assume a uniform prior over rational outcomes, for instance, then we would see that the causal effect of $\mathcal{I}$ on the percentage of customers switching insurance provider is given by:

$$\frac{1}{|\mathcal{R}(\mathcal{M}_{\mathcal{I}})|} \sum_{\boldsymbol{\pi} \in \mathcal{R}(\mathcal{M}_{\mathcal{I}})} \mathrm{switch}(\boldsymbol{\pi}) - \frac{1}{|\mathcal{R}(\mathcal{M})|} \sum_{\boldsymbol{\pi} \in \mathcal{R}(\mathcal{M})} \mathrm{switch}(\boldsymbol{\pi}) = 21.6\% - 36\% = -14.4\%,$$

where $\mathrm{switch}(\boldsymbol{\pi}) \coloneqq \sum_i \Pr^{\boldsymbol{\pi}}(D^3 = i \mid C = 3 - i)$ is the probability that a customer switches provider. Thus, hypothesis i) is confirmed.

39

For the first part of hypothesis ii), we wish to measure the causal effect of $\mathcal{I}$ on the quantity $\mathbb{E}_{\boldsymbol{\pi}}[D^i \mid C = i, D^3 = 3 - i]$ for each $i \in \{1, 2\}$, in settings such that $\Pr^{\boldsymbol{\pi}}(C = i, D^3 = 3 - i) > 0$. For $\boldsymbol{\pi} \in \mathcal{R}(\mathcal{M})$, this quantity is equal to 206, which may occur when either $i = 1$ or $i = 2$. For $\boldsymbol{\pi} \in \mathcal{R}(\mathcal{M}_{\mathcal{I}})$ this quantity is equal to 250, which occurs when $i = 2$; the customer never switches to firm one. This confirms the first half of the hypothesis. For the second part, a similar analysis shows that renewing customers pay £244 before the intervention either £250 or £288 after the intervention, which is *contrary* to the hypothesis. We discuss the reasons for this in the following paragraph, as well as how another proposal can be employed to achieve the desired result.

Hypothesis iii) concerns simply the causal effect of $\mathcal{I}$ on $\mathbb{E}_{\boldsymbol{\pi}}[U^3]$. The customer's expected utility is the same under any of the rational outcomes and interventional rational outcomes respectively, so we can write simply:

$$\mathbb{E}_{\boldsymbol{\pi}}[U_{\mathcal{I}}^3] - \mathbb{E}_{\boldsymbol{\pi}}[U^3] = 23.12 - 52 = -28.88.$$

As in the case for the second half of hypothesis ii), this *disconfirms* the hypothesis (in our model). The reason for both of these results is that the relatively high switching costs for customers (£38) and the relatively high percentage of inert customers (28%) mean that the firms' pricing algorithms are better off sticking to higher prices to exploit customers who will not switch, rather than lower prices in order to attract customers away from other firms.

To achieve the desired effects, we can instead consider another intervention proposed by the FCA, which is to reduce switching costs for customers (for example, by allowing them to stop their insurance policy from auto-renewing more easily) [25]. If we therefore consider a second intervention $\mathcal{I}'$ that makes the same change as $\mathcal{I}$ but also modifies Θ_{U^3} to set the switching costs of customers to £0, for example, then we have that:

$$\mathbb{E}_{\boldsymbol{\pi}}[U_{\mathcal{I}'}^3] - \mathbb{E}_{\boldsymbol{\pi}}[U^3] = 72 - 52 = 20,$$

as $\pi_{D^i}(D^i \mid C) = \delta(D^i, 224)$ for $i \in \{1, 2\}$ in any $\boldsymbol{\pi} \in \mathcal{R}(\mathcal{M}_{\mathcal{I}'})$. Returning to hypothesis ii), this implies that all customers now pay £224 when renewing as opposed to £244, thus completing the set of all desired outcomes.

Extensions

As shown above, not only can our analysis be used to investigate the qualitative hypotheses put forwards by the FCA, it provides quantitative results too. Nevertheless, we emphasise that the model we present is highly simplified, and considers only a very simple intervention. More complex models and queries are beyond the scope of the present paper, but represent a natural avenue for further applied work. Before continuing, we briefly remark on possible extensions.

First, we could extend the model to make it more realistic. This might include: increasing the number of firms; different marginal costs for different firms; different customer valuations of insurance policies from different firms; small costs to firms from customers switching; including intermediary firms; more explicit modelling of the firms' pricing algorithms; not making savviness a binary variable; and modelling the market over time instead of as a one-shot game. With these added complexities to the model come greater opportunities to exploit the richness of causal games.

In particular, in the simple model above, the lack of more intricate causal dependencies between the agents' actions means that applying a pre-policy intervention essentially reduces to re-solving a slightly different game (where the dashed edges in Figure 10a are removed). In larger games, it is both easier and more important to avoid this cost, which cannot (in general) be avoided in models such as EFGs that do not explicitly represent the causal structure of a game. We could

also use the model for more advanced analyses, such as answering counterfactuals, identifying subgames, and computing mixtures of pre- and post-policy queries.

7.2 Blame, Intent, Incentives, and Fairness

Many of the most important concepts for reasoning about the safe and ethical use of AI systems are implicitly causal in nature. Moreover, in recent years, substantial progress has been made on formalising these concepts using causal models [22, 39, 55, 65, 80]. Because causal games are the first causal framework to explicitly capture multi-agent and game-theoretic reasoning, they open up the possibility of further work in these directions. In what follows, we explain how causal games might be applied to existing concepts in order to arrive at richer and more general results. Concrete examples of how these concepts can be modelled using causal games are provided in Appendix B.3.

Blame and Intent

SCMs have been fruitfully employed in order to formally define the notion of *actual causation* (i.e., what it means for some event c to, in fact, cause some event e) [37, 40]. Such ideas are not only of philosophical interest, but have been argued as being crucial for building AI systems that can automatically generate *explanations*, either of their own workings or that of other agents [41, 65]. In recent years, this underlying theory has been used to formalise the concepts of blameworthiness and intention [28, 39]. As these works have highlighted, issues of blame and intent become particularly acute in multi-agent settings.

While these concepts can be accommodated within (structural) causal games (see Appendix B.3), they also allow for an even richer formalism of blame and intent in multi-agent settings. More specifically, there are at least four generalisations we might make. Firstly, the use of sets of utility variables that define *multiple* utility functions, one for each agent, would mean that costs (and hence both blame and intent) can be viewed from the perspective of multiple agents. Secondly, we could consider blame and intent with respect to *strategies* rather than single actions, which corresponds to considering soft (post-policy) interventions as opposed to hard (post-policy) interventions. Thirdly, we might wish to consider *pre-policy* interventions by using mechanised games. For example, if a pre-policy change to one agent's decision means that the only rational response for the second agent is something that causes a bad outcome, we might think that the first agent is at least somewhat to blame for this turn of events (although these concepts become more nuanced when there are cyclic dependencies between decision rule variables). Finally, all of these extensions can also be viewed in terms of *coalitions* of agents and their strategies, which has already begun to be explored in prior work [28].

Incentives

In order to build intelligent agents that do not behave in undesirable ways, it is useful to be able to reason about their *incentives* [19, 22, 57] and *reasoning patterns* [3, 76]. For example, given a (possibly multi-agent) decision-making scenario, we might be interested in asking whether an agent has an incentive to base its policy on a protected attribute, or to influence a variable that we would prefer to be left alone. Similarly, notions of fairness are often naturally expressed in terms of whether a different outcome would result in some counterfactual situation.

Recent work has sought to formalise incentives by modelling such scenarios as single-agent causal games (i.e., causal influence diagrams). This work has developed a number of sound and complete graphical criteria for identifying Value of Information (VoI) [44] and Value of Control (VoC) [87]. Sound and complete graphical criteria have also been introduced for two novel concepts: *response incentives*, which indicate when a decision-maker can benefit from

responding to a variable, and *instrumental control incentives*, which establish whether an agent can benefit from manipulating a variable [22].

Existing work, however, has been limited to the single-agent, single-decision setting; it is our hope that causal games will provide the basis for generalising this work to the multi-agent, multi-decision setting and that the use of mechanised games may lead to additional insights. For instance, for some incentive concept C we may wish to ask questions such as 'is there any $\mathcal{R}$-rational outcome in which agent $i \in N$ has a C incentive with respect to a particular variable X in the game?'. Given the fact that many proposals for safe AI systems are multi-agent [23], this generalisation to the multi-agent setting marks an important next step in analysing agent incentives. We provide an example of such a proposal – cooperative inverse reinforcement learning [36] – modelled as a causal game, in Appendix B.3.

Fairness

Another important and useful application of incentives is for reasoning about fairness, a number of popular and influential definitions of which are based explicitly on causal frameworks [4, 49, 55, 67, 102]. Indeed, it can be shown that all optimal policies π^* in a single-decision SCIM are *counterfactually unfair* [55] with respect to a protected attribute A (meaning that a change to the protected attribute would change the decision made) if and only if there is an RI on A [22].

The question of fairness arguably becomes even more important and interesting in the multi-agent setting, in which one not only has to ask whether or not a process is fair, but fair to *whom*, and whether we might be forced to trade off fairness with respect to different agents. Interestingly, despite the large number of existing works investigating fairness in games (see, e.g., the seminal papers of Rabin [79] or Fehr and Schmidt [24]) and the recent insights gained from using causal definitions of fairness, to the best of our knowledge there has been no application of these definitions to the multi-agent setting. One possible (partial) explanation for this is that researchers have, until now, been lacking precisely the kind of models that we introduce.

8 Discussion

In this paper, we have introduced a framework that we argue provides a unifying formalism for reasoning about causality in games. As mentioned in Section 1, combinations of causal and game-theoretic reasoning have long been considered by various research communities, and so we conclude with a brief summary of the relative advantages and disadvantages of causal games compared to other models, before offering some final thoughts about future directions.

8.1 Advantages and Disadvantages of Causal Games

The primary benefit of causal games compared to prior work is that no previous framework captured both game-theoretic and causal features in a general, principled way. Alongside our discussion of related work in Section 1.1, we expand briefly upon this claim by considering other causal and game-theoretic models in turn.

Causal Models

Standard causal models do not take into account the presence of rational, self-interested agents. Doing so requires more than a simple re-labelling of some of the variables as decisions and utilities, as the presence of strategic, decision-making agents violates the standard assumption of independent causal mechanisms [75], represented by the edges between mechanism variables in mechanised games. Alternative causal models that do include agents, such as CIDs [18, 22, 45], typically only consider a single agent.

To the best of our knowledge, the only exception to this rule is settable systems [95, 96], though these models do not capture the multiplicity of equilibria that arise in game-theoretic analyses, and thus do not naturally support the analysis we seek to do in this paper. The focus in these works is instead on capturing lower-level algorithmic details, which may make them more appropriate when reasoning about the data and attributes of, say, an optimisation or machine learning process.

Our use of *non-deterministic* mechanisms (arising whenever agents are indifferent between alternatives) also serves to generalise existing work on cyclic causal models [8, 38] to the relational setting [1, 33, 62]. However, the cyclic dependencies in causal games are of a specific form due to the fact that they represent game-theoretic equilibria, which means that causal games are not necessarily an appropriate model for analysing the more arbitrary dynamical systems and sets of simultaneous equations studied in these works.

The fact that we build on top of MAIDs means that causal games inherit some of the existing game-theoretic concepts introduced in these models, such as NEs, while being consistent with the more standard probabilistic graphical models upon which MAIDs are based. This further allows us to derive notions of subgames and other equilibrium refinements, which are critical for game-theoretic reasoning and yet are not supported by other causal models. Such concepts have been historically underexplored in the context of graphical games when compared to, say, EFGs or strategic-form games.

Game-Theoretic Models

The most natural game-theoretic model with which to compare causal games is EFGs, which are tree-based models as opposed to DAG-based. For our purposes, the most important benefit of DAG-based models is that they can be more readily used to natively support a wide range of causal queries. This is bolstered by the use of $\mathcal{R}$-relevance graphs, which form an explicit representation of the causal dependencies between agents' decision rules; there is no such representation for these dependencies in EFGs.[24] Mechanised games can be used to define a wide range of both pre- and post-policy queries in games. In contrast, as explained in Section 6.3, there are many causal queries that *cannot* be natively answered in EFGs.

DAG-based models also more compactly and explicitly represent dependencies between variables, which can often only be understood in an EFG through inspecting the parameterisation of the game. Moreover, a DAG-based representation of a game need never be bigger than the corresponding EFG and can be exponentially smaller. On the other hand, if a game is highly asymmetric in its play paths (such as when if play proceeds down one path the game stops immediately, and on the other path it continues for several more moves), then this structure is not immediately observable from a causal game, and may effectively make many valuations of the variables irrelevant due to their inconsistency with the shorter game path.

In addition to their transparency, MAIDs allow us to further exploit the conditional independencies between variables using d-separation. Using $\mathcal{R}$-reachability, we may construct the $\mathcal{R}$-relevance graph for any game and find more subgames in a MAID than in its equivalent EFG. This can significantly reduce the computational complexity of solving a game (as shown by the example in Appendix C.2), offers analytical benefits, and provides a way to define a stronger subgame perfectness condition. On the other hand, in the case of *context-specific independencies* – such as when A sometimes depends on B, and B sometimes depends on A – it is well-known that DAG-based models are a less natural choice than tree-based models [10].[25]

---

[24]This further implies that we cannot take advantage of the structure of these dependencies in order to answer pre-strategy queries more efficiently.

[25]Though it is possible to support said independencies in DAGs via tree-based representations of the CPDs, which can graphically capture different independencies on different branches.

Finally, though we have significantly extended the number of standard game-theoretic concepts for MAIDs (and hence causal games) and proved their equivalence to their EFG counterparts, EFGs remain a more well-investigated representation. Thus, if one is interested in more exotic equilibrium refinements, for example, EFGs are likely to be a more suitable model. It is our hope that further research on MAIDs and causal games will reduce this last difference.

8.2 Future Work

Our priority is to use causal games to further analyse incentives in multi-agent systems, which has important applications in ensuring that we build AI systems that are safe and fair. As mentioned in Appendix 7.2, existing work has already characterised some of these incentives using CIDs. Therefore, a natural next step is to extend these definitions to multi-agent (and multi-decision) scenarios, though this is by no means trivial. For example, in single agent-settings VoI is always non-negative, whereas in multi-agent settings this need not be the case (if other agents are aware of the information gain) [78]. Further, we might wish to rule certain incentives in or out based on whether or not they occur under all or some policy profiles satisfying a particular equilibrium refinement, or more generally, falling within the set of rational outcomes.

Other specific applications for which causal games may prove fruitful are, for example: designing mechanisms for auctions and other multi-agent systems, or analysing possible interventions on those mechanisms; generalising counterfactual fairness to multi-agent settings; providing artificial agents with the means to more easily provide explanations and reason about qualitative concepts (such as blame and intent or reasoning patterns) that can be defined using causal models of games; and deriving new definitions for similar concepts. We might also hope to extend the framework presented here with: model variations that can more easily capture dynamic settings, fine-grained subjective beliefs, or optimisation; definitions capturing other classic equilibrium refinements such as perfect Bayesian equilibrium [29] or sequential equilibrium [53]; and methods of causal discovery for games.

Given that we propose this paper and our accompanying codebase as a robust foundation for reasoning about causality in games, we believe our work presents many other interesting avenues for further research. We hope that the advantages causal games confer based on their generality, explainability, and succinctness (not to mention their compatibility with existing mainstream models) make them an attractive choice for researchers and practitioners alike who are interested in the intersection of causality and game theory.

Acknowledgements

References

[1] Ahsan, R., Arbour, D., and Zheleva, E. "Relational Causal Models with Cycles: Representation and Reasoning". *First Conference on Causal Learning and Reasoning, Proceedings of Machine Learning Research* 177 (2022), pp. 1–18.

[2] Ankan, A. and Panda, A. "Pgmpy: Probabilistic Graphical Models Using Python". *Proceedings of the 14th Python in Science Conference*. 2015, pp. 6–11.

[3] Antos, D. and Pfeffer, A. "Identifying Reasoning Patterns in Games". *Proceedings of the Twenty-fourth Conference on Uncertainty in Artificial Intelligence*. 2008, pp. 9–18.

[4] Ashurst, C., Carey, R., Chiappa, S., and Everitt, T. "Why Fair Labels Can Yield Unfair Predictions: Graphical Conditions for Introduced Unfairness". *Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence*. 2022, pp. 9494–9503.

[5] Avin, C., Shpitser, I., and Pearl, J. "Identifiability of Path-specific Effects". *Proceedings of the 19th International Joint Conference on Artificial Intelligence*. 2005, pp. 357–363.

[6] Balke, A. and Pearl, J. "Probabilistic Evaluation of Counterfactual Queries". *Proceedings of the Twelfth AAAI National Conference on Artificial Intelligence*. 1994, pp. 230–237.

[7] Bauer, H. *Probability Theory*. Vol. 23. Translated by: Robert B. Burckel. De Gruyter, 1996.

[8] Bongers, S., Forré, P., Peters, J., and Mooij, J. M. "Foundations of Structural Causal Models with Cycles and Latent Variables". *The Annals of Statistics* 49.5 (2021), pp. 2885–2915.

[9] Bostrom, N. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, 2014.

[10] Boutilier, C., Friedman, N., Goldszmidt, M., and Koller, D. "Context-specific Independence in Bayesian Networks". *Proceedings of the Twelfth International Conference on Uncertainty in Artificial Intelligence*. 1996, pp. 115–123.

[11] Bradfield, J., Gutierrez, J., and Wooldridge, M. "Partial-order Boolean Games: Informational Independence in a Logic-based Model of Strategic Interaction". *Synthese* 193.3 (2015), pp. 781–811.

[12] Chatterjee, K., Henzinger, T. A., and Piterman, N. "Strategy Logic". *Information and Computation* 208.6 (2010). Special Issue: 18th International Conference on Concurrency Theory, pp. 677–693.

[13] Correa, J. and Bareinboim, E. "A Calculus for Stochastic Interventions: Causal Effect Identification and Surrogate Experiments". *Proceedings of the Thirty-fourth AAAI Conference on Artificial Intelligence*. 2020, pp. 10093–10100.

[14] Cozman, F. G. "Credal Networks". *Artificial Intelligence* 120.2 (2000), pp. 199–233.

[15] Daskalakis, C., Goldberg, P. W., and Papadimitriou, C. H. "The Complexity of Computing a Nash Equilibrium". *Siam Journal on Computing* 39.1 (2009), pp. 195–259.

[16] Daskalakis, C., Mehta, A., and Papadimitriou, C. "A Note on Approximate Nash Equilibria". *International Workshop on Internet and Network Economics*. Springer. 2006, pp. 297–306.

[17] Dawid, A. P. "Causal Inference without Counterfactuals". *Journal of the American Statistical Association* 95.450 (2000), pp. 407–424.

[18] Dawid, A. P. "Influence Diagrams for Causal Modelling and Inference". *International Statistical Review* 70.2 (2002), pp. 161–189.

[19] Déletang, G., Grau-Moya, J., Martic, M., Genewein, T., McGrath, T., Mikulik, V., Kunesch, M., Legg, S., and Ortega, P. A. "Causal Analysis of Agent Behavior for AI Safety". *arXiv:2103.03938* (2021).

[20] Doshi, P., Zeng, Y., and Chen, Q. "Graphical Models for Interactive POMDPs: Representations and Solutions". *Autonomous Agents and Multi-agent Systems* 18.3 (2008), pp. 376–416.

[21] Eberhardt, F. and Scheines, R. "Interventions and Causal Inference". *Philosophy of Science* 74.5 (2007), pp. 981–995.

[22] Everitt, T., Carey, R., Langlois, E. D., Ortega, P. A., and Legg, S. "Agent Incentives: A Causal Perspective". *Proceedings of the Thirty-fifth AAAI Conference on Artificial Intelligence*. 2021, pp. 11487–11495.

[23] Everitt, T., Kumar, R., Krakovna, V., and Legg, S. "Modeling AGI Safety Frameworks with Causal Influence Diagrams". *IJCAI AI Safety Workshop*. 2019.

[24] Fehr, E. and Schmidt, K. M. "A Theory of Fairness, Competition, and Cooperation". *The Quarterly Journal of Economics* 114.3 (1999), pp. 817–868.

[25] Financial Conduct Authority. *General Insurance Pricing Practices Market Study: Consultation on Handbook changes*. CP20/19. 2020.

[26] Financial Conduct Authority. *General insurance Pricing Practices: Interim Report*. MS18/1.2. 2019.

[27] Fox, J., Everitt, T., Carey, R., Langlois, E., Abate, A., and Wooldridge, M. "PyCID: A Python Library for Causal Influence Diagrams". *Proceedings of the 20th Python in Science Conference*. 2021, pp. 43–51.

[28] Friedenberg, M. and Halpern, J. Y. "Blameworthiness in Multi-agent Settings". *Proceedings of the Thirty-third AAAI Conference on Artificial Intelligence*. 2019, pp. 525–532.

[29] Fudenberg, D. and Tirole, J. "Perfect Bayesian Equilibrium and Sequential Equilibrium". *Journal of Economic Theory* 53.2 (1991), pp. 236–260.

[30] Gal, Y. and Pfeffer, A. "Networks of Influence Diagrams: A Formalism for Representing Agents' Beliefs and Decision-making Processes". *Journal of Artificial Intelligence Research* 33 (2008), pp. 109–147.

[31] Geiger, D., Verma, T., and Pearl, J. "Identifying Independence in Bayesian Networks". *Networks* 20.5 (1990), pp. 507–534.

[32] Genewein, T., McGrath, T., Déletang, G., Mikulik, V., Martic, M., Legg, S., and Ortega, P. A. "Algorithms for Causal Reasoning in Probability Trees". *arXiv:2010.12237* (2020).

[33] Getoor, L., Friedman, N., Koller, D., Pfeffer, A., and Taskar, B. "Probabilistic Relational Models". *Introduction to Statistical Relational Learning*. Ed. by Getoor, L. and Taskar, B. The MIT Press, 2007.

[34] GlobalData. *United Kingdom (UK) Household Insurance Market Size, Trends, Competitor Dynamics and Opportunities*. Report Code: GDFS0406IA. 2022.

[35] Gonzalez-Soto, M., Sucar, L. E., and Escalante, H. J. "Causal Games and Causal Nash Equilibrium". *arXiv:1910.06729* (2019).

[36] Hadfield-Menell, D., Dragan, A., Abbeel, P., and Russell, S. "Cooperative Inverse Reinforcement Learning". *Proceedings of the 30th International Conference on Neural Information Processing Systems*. 2016, pp. 3916–3924.

[37] Halpern, J. Y. "A Modification of the Halpern-Pearl Definition of Causality". *Proceedings of the 24th International Conference on Artificial Intelligence*. 2015, pp. 3022–3033.

[38] Halpern, J. Y. "Axiomatizing Causal Reasoning". *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*. 1998, pp. 202–210.

[39] Halpern, J. Y. and Kleiman-Weiner, M. "Towards Formal Definitions of Blameworthiness, Intention, and Moral Responsibility". *Proceedings of the Thirty-second AAAI Conference on Artificial Intelligence*. 2018, pp. 1853–1860.

[40] Halpern, J. Y. and Pearl, J. "Causes and Explanations: A Structural-model Approach. Part I: Causes". *The British Journal for the Philosophy of Science* 56.4 (2005), pp. 843–887.

[41] Halpern, J. Y. and Pearl, J. "Causes and Explanations: A Structural-model Approach. Part II: Explanations". *The British Journal for the Philosophy of Science* 56.4 (2005), pp. 889–911.

[42] Hammond, L., Fox, J., Everitt, T., Abate, A., and Wooldridge, M. "Equilibrium Refinements for Multi-agent Influence Diagrams: Theory and Practice". *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems*. 2021, pp. 574–582.

[43] Heckerman, D. and Shachter, R. "A Decision-based View of Causality". *Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence*. 1994, pp. 302–310.

[44] Howard, R. "Information Value Theory". *IEEE Transactions on Systems Science and Cybernetics* 2.1 (1966), pp. 22–26.

[45] Howard, R. A. and Matheson, J. E. "Influence Diagrams". *Decision Analysis* 2.3 (2005), pp. 127–143.

[46] Jaques, N., Lazaridou, A., Hughes, E., Gülçehre, Ç., Ortega, P. A., Strouse, D., Leibo, J. Z., and Freitas, N. de. "Social Influence As Intrinsic Motivation for Multi-agent Deep Reinforcement Learning". *Proceedings of the 36th International Conference on Machine Learning*. 2019, pp. 3040–3049.

[47] Jiang, A. X., Leyton-Brown, K., and Pfeffer, A. "Temporal Action-graph Games: A New Representation for Dynamic Games". *Proceedings of the Twenty-fifth Conference on Uncertainty in Artificial Intelligence*. 2009, pp. 268–276.

[48] Kearns, M. "Graphical Games". *Algorithmic Game Theory*. Ed. by Nisan, N., Roughgarden, T., Tardos, E., and Vazirani, V. V. Cambridge University Press, 2007. Chap. 7, pp. 159–180.

[49] Kilbertus, N., Rojas-Carulla, M., Parascandolo, G., Hardt, M., Janzing, D., and Schölkopf, B. "Avoiding Discrimination through Causal Reasoning". *Proceedings of the 31st International Conference on Neural Information Processing Systems*. 2017, pp. 656–666.

[50] Koller, D. and Milch, B. "Multi-agent Influence Diagrams for Representing and Solving Games". *Proceedings of the 17th International Joint Conference on Artificial Intelligence*. 2001, pp. 1027–1034.

[51] Koller, D. and Milch, B. "Multi-agent Influence Diagrams for Representing and Solving Games". *Games and Economic Behavior* 45.1 (2003), pp. 181–221.

[52] Korb, K. B., Hope, L. R., Nicholson, A. E., and Axnick, K. "Varieties of Causal Intervention". *PRICAI 2004: Trends in Artificial Intelligence*. Ed. by Zhang, C., W. Guesgen, H., and Yeap, W.-K. 2004, pp. 322–331.

[53] Kreps, D. M. and Wilson, R. "Sequential Equilibria". *Econometrica* 50.4 (1982), p. 863.

[54] Kuhn, H. W. "Extensive Games and the Problem of Information". *Contributions to the Theory of Games (AM-28)*. Vol. 2. Princeton University Press, 1953, pp. 193–216.

[55] Kusner, M., Loftus, J., Russell, C., and Silva, R. "Counterfactual Fairness". *Proceedings of the 31st International Conference on Neural Information Processing Systems*. 2017, pp. 4069–4079.

[56] Kwiatkowska, M., Norman, G., Parker, D., and Santos, G. "Multi-player Equilibria Verification for Concurrent Stochastic Games". *Proceedings of the 17th International Conference on Quantitative Evaluation of SysTems*. 2020.

[57] Langlois, E. D. and Everitt, T. "How RL Agents Behave When Their Actions Are Modified". *Proceedings of the Thirty-fifth AAAI Conference on Artificial Intelligence*. 2021, pp. 11586–11594.

[58] Lauritzen, S. L. and Nilsson, D. "Representing and Solving Decision Problems with Limited Information". *Management Science* 47.9 (2001), pp. 1235–1251.

[59] Lauritzen, S. L. and Richardson, T. S. "Chain Graph Models and Their Causal Interpretations". *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64.3 (2002), pp. 321–348.

[60] Levi, I. *The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability, and Chance*. MIT Press, 1980.

[61] London Economics, YouGov, and Kudos Research. *General Insurance Pricing Practices*. 2019.

[62] Maier, M., Marazopoulou, K., and Jensen, D. "Reasoning about Independence in Probabilistic Models of Relational Data". *arXiv:1302.4381* (2013).

[63] McKelvey, R. D., McLennan, A. M., and Turocy, T. L. *Gambit: Software Tools for Game Theory*. Version 0.2006. 01.20. 2016.

[64] Milch, B. and Koller, D. *Ignorable Information in Multi-agent Scenarios*. Tech. rep. MIT-CSAIL-TR-2008-029. Computer Science and Artificial Intelligence Laboratory, MIT, 2008.

[65] Miller, T. "Contrastive Explanation: A Structural-model Approach". *The Knowledge Engineering Review* 36 (2021).

[66] Miller III, A. C., Merkhofer, M. W., Howard, R. A., Matheson, J. E., and Rice, T. R. *Development of Automated Aids for Decision Analysis*. Tech. rep. 1976.

[67] Nabi, R. and Shpitser, I. "Fair Inference on Outcomes". *Proceedings of the Thirty-second AAAI Conference on Artificial Intelligence*. 2018, pp. 1931–1940.

[68] Nash, J. F. "Equilibrium Points in N-person Games". *Proceedings of the National Academy of Sciences* 36.1 (1950), pp. 48–49.

[69] Nielsen, M., Plotkin, G., and Winskel, G. "Petri Nets, Event Structures and Domains". *Semantics of Concurrent Computation*. Springer, 1979, pp. 266–284.

[70] Nielsen, T. D. and Jensen, F. V. "Welldefined Decision Scenarios". *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence*. 1999, pp. 502–511.

[71] Nisan, N., Roughgarden, T., Tardos, E., and Vazirani, V. V., eds. *Algorithmic Game Theory*. Cambridge University Press, 2007.

[72] Osborne, M. J. and Rubinstein, A. *A Course in Game Theory*. MIT Press, 1994.

[73] Pearl, J. *Causality*. Cambridge University Press, 2009.

[74] Pearl, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers Inc., 1988.

[75] Peters, J., Janzing, D., and Schölkopf, B. *Elements of Causal Inference: Foundations and Learning Algorithms*. The MIT Press, 2017.

[76] Pfeffer, A. and Gal, Y. "On the Reasoning Patterns of Agents in Games". *Proceedings of the 22nd National Conference on Artificial Intelligence*. 2007, pp. 102–109.

[77] Polich, K. and Gmytrasiewicz, P. "Interactive Dynamic Influence Diagrams". *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems*. 2007, pp. 1–3.

[78] Ponssard, J.-P. "On the Concept of the Value of Information in Competitive Situations". *Management Science* 22.7 (1976), pp. 739–747.

[79] Rabin, M. "Incorporating Fairness into Game Theory and Economics". *The American Economic Review* 83.5 (1993), pp. 1281–1302.

[80] Richens, J. G., Beard, R., and Thompson, D. H. "Counterfactual Harm". *arXiv:2204.12993* (2022).

[81] Rideau, S. and Winskel, G. "Concurrent Strategies". *2011 IEEE 26th Annual Symposium on Logic in Computer Science*. 2011.

[82] Rubin, D. B. "Causal Inference Using Potential Outcomes". *Journal of the American Statistical Association* 100.469 (2005), pp. 322–331.

[83] Russell, S. *Human Compatible*. Penguin LCC US, 2019.

[84] Selten, R. "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games". *International Journal of Game Theory* 4.1 (1975), pp. 25–55.

[85] Selten, R. "Spieltheoretische Behandlung Eines Oligopolmodells Mit Nachfrageträgheit: Teil I: Bestimmung Des Dynamischen Preisgleichgewichts". *Journal of Institutional and Theoretical Economics* H. 2 (1965), pp. 301–324.

[86] Shachter, R. D. "Bayes-ball: Rational Pastime (for Determining Irrelevance and Requisite Information in Belief Networks and Influence Diagrams)". *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*. 1998, pp. 480–487.

[87] Shachter, R. D. "Evaluating Influence Diagrams". *Operations Research* 34.6 (1986), pp. 871–882.

[88] Shoham, Y. and Leyton-Brown, K. *Multiagent Systems: Algorithmic, Game-theoretic, and Logical Foundations*. Cambridge University Press, 2008.

[89] Simon, H. A. *Models of Man. Social and Rational. Mathematical Essays on Rational Human Behavior in a Social Setting*. John Wiley & Sons, 1957.

[90] Spence, M. "Job Market Signaling". *The Quarterly Journal of Economics* 87.3 (1973), p. 355.

[91] Spirtes, P., Glymour, C., and Scheines, R. *Causation, Prediction, and Search*. Springer, 1993.

[92] Toulis, P. and Parkes, D. C. "Long-term Causal Effects Via Behavioral Game Theory". *Advances in Neural Information Processing Systems 29*. 2016, pp. 2604–2612.

[93] Verma, T. and Pearl, J. "Causal Networks: Semantics and Expressiveness". *Proceedings of the Fourth Annual Conference on Uncertainty in Artificial Intelligence*. 1990, pp. 69–78.

[94] von Neumann, J. "Zur Theorie Der Gesellschaftsspiele". *Mathematische Annalen* 100.1 (1928), pp. 295–320.

[95] White, H. and Chalak, K. "Settable Systems: An Extension of Pearl's Causal Model with Optimization, Equilibrium, and Learning". *Journal of Machine Learning Research* 10 (2009), pp. 1759–1799.

[96] White, H., Xu, H., and Chalak, K. "Causal Discourse in a Game of Incomplete Information". *Journal of Econometrics* 182.1 (2014), pp. 45–58.

[97] Wichardt, P. C. "Existence of Nash Equilibria in Finite Extensive Form Games with Imperfect Recall: A Counterexample". *Games and Economic Behavior* 63.1 (2008), pp. 366–369.

[98]   Winskel, G. "Distributed Probabilistic and Quantum Strategies". *Electronic Notes in Theoretical Computer Science* 298 (2013), pp. 403–425.

[99]   Wooldridge, M. *An Introduction to Multiagent Systems.* John Wiley & Sons, 2009.

[100] Wooldridge, M., Gutierrez, J., Harrenstein, P., Marchioni, E., Perelli, G., and Toumi, A. "Rational Verification: From Model Checking to Equilibrium Checking". *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence.* 2016, pp. 4184–4190.

[101] Zahoransky, V., Gutierrez, J., Harrenstein, P., and Wooldridge, M. "Partial Order Games". *Games* 13.1 (2021), p. 2.

[102] Zhang, J. and Bareinboim, E. "Fairness in Decision-making - the Causal Explanation Formula". *Proceedings of the Thirty-second AAAI Conference on Artificial Intelligence.* 2018, pp. 2037–2045.

A   Proofs

A.1   Transformations between Game Representations

MAID to EFG

In this section, we provide an encoding transforming a MAID, $\mathcal{M} = (\mathcal{G}, \boldsymbol{\theta})$, into an EFG, $\mathcal{E} = (N, T, P, A, \lambda, I, U)$. One can decide on whether one wants to keep record of the structure of the original MAID, or if one wants to optimise the encoding by splitting only only some of $\mathcal{G}$'s variables. If one wants to be able to preserve the full structure of $\mathcal{M}$ then the set of splitting variables in the resulting tree is $\boldsymbol{S} = \boldsymbol{X} \cup \boldsymbol{D}$.[26] If instead, however, one wishes to minimise the complexity of computing equilibria, then one only needs to split on the MAID's decision variables and their parents, $\boldsymbol{S} = \mathbf{Fa}_{\boldsymbol{D}}$ [74]. This is because the EFG's tree size will be exponential in $|\boldsymbol{S}|$. The following procedure, which we refer to as `maid2efg`, is based on that of K&M [51]. Given a MAID $\mathcal{M} = (\mathcal{G}, \boldsymbol{\theta})$ we define an equivalent EFG $\mathcal{E} = (N, T, P, A, \lambda, I, U)$ as follows:

- Choose a topological ordering $S_1 \prec \cdots \prec S_n$ over all variables in $\boldsymbol{S}$ such that if S_k is a descendent of S_j, then $S_j \prec S_k$.[27] Further, define $\boldsymbol{S}_k^{\prec} = \{S' \in \boldsymbol{S} : S' \prec S_k\}$.

- The set of agents, N, in $\mathcal{E}$ is the same as in $\mathcal{M}$.

- The tree T is a symmetric tree with each path containing splits over all the variables in $\boldsymbol{S}$ in the order defined by $\prec$.

- For all variables $S \in \boldsymbol{S}$, if S is a chance variable $S \in \boldsymbol{X}$, add it to $\mathcal{E}$'s set of chance nodes $\boldsymbol{V}^0$. Else, if S is a decision variable $S \in \boldsymbol{D}^i$, add S to agent i's set of nodes in $\mathcal{E}$, $\boldsymbol{V}^i$.

- Label each node V in T with an instantiation $\mu(V)$ corresponding to the values taken by each EFG node (i.e., the branch followed from this node) on the path from the tree's root node R to V.

- For every node V in T such that its corresponding variable $S_V \in \boldsymbol{S}$ is a chance variable in $\mathcal{M}$, we determine the probability distribution $P_V : \mathbf{Ch}_V \to [0, 1]$ over its children $V' \in \mathbf{Ch}_V$ (each child corresponds to a value $s_v \in dom(S_V)$) by querying that variable's CPD with the instantiation label at V. In other words, $P_V(V') \coloneqq \Pr\left(s_v \mid \mu(V)\right)$. This equation can be simplified because in a BN the CPD for each variable S_V only depends on the values of its parents, $\mathbf{pa}_{S_V}$. Therefore, we have:

$$P_V(V') \coloneqq \Pr\left(s_v \mid \mu(V)\right) = \Pr\left(s_v \mid \mu(V)[\mathbf{Pa}_{S_V}]\right),$$

---

[26]We assume that $\mathbf{desc}_U = \varnothing$.

[27]This topological ordering will, in general, be non-unique.

where $\mu(V)[\mathbf{Pa}_{S_V}])$ is the restriction of $\mu(V)$ to the values of $\mathbf{Pa}_{S_V}$. If $P_V(V') = 0$ for some child V', we remove that child from the EFG.

- The set of available actions A_j^i for agent i at node V_j^i in the EFG is given by the domain of the corresponding decision variable $D = S_{V_j^i}$ in the MAID, i.e. $A_j^i := dom(D)$.

- Define $\lambda : \mathbf{V} \times \mathbf{V} \to A$ for each decision node $V \in \mathbf{V}^1 \cup \cdots \cup \mathbf{V}^n$ and $V' \in \mathbf{Ch}_V$ such that $\lambda(V, V') = \mu(V')[S_V]$ to label the outcome of the decision.

- Define an equivalence relation $\sim$ over $\mathbf{V}^1 \cup \cdots \cup \mathbf{V}^n$ such that $V \sim V'$ if and only if $S_V = S_{V'} = S$ and $\mu(V)[\mathbf{Pa}_{S_V}] = \mu(V')[\mathbf{Pa}_{S_{V'}}]$. Then the set of information sets I^i for each agent $i \in N$ is simply the quotient set $\mathbf{V}^i/\sim$ – the set of $\sim$ equivalence classes partitioning $\mathbf{V}^i$. In other words, two nodes $V, V' \in \mathbf{V}$ which correspond to the same decision variable S in $\mathcal{M}$ are in the same information set I_j^i for agent i if and only if their instantiation labels, restricted to their parents in $\mathcal{M}$, are the same.

- The payoff $U : \mathbf{L} \to \mathbb{R}^n$ for each leaf $L \in \mathbf{L}$ of the EFG's tree is $U(L) = (u^1, \dots, u^n)$, where $u^i := \sum_{U \in \mathbf{U}^i} \mathbb{E}[U \mid \mu(L)]$ is the expected utility of agent i in $\mathcal{M}$, given the instantiation $\mu(L)$ of the variables $\mathbf{S}$.

EFG to MAID

We now detail the construction `efg2maid`. Unlike in the case for `maid2efg`, we show how every EFG can be converted to a *unique*, canonical equivalent MAID. Given an EFG $\mathcal{E} = (N, T, P, A, \lambda, I, U)$ (including intervention sets) we define an equivalent MAID $\mathcal{M} = (\mathcal{G}, \boldsymbol{\theta})$ as follows:

- The set of agents N remains the same in $\mathcal{M}$ as in $\mathcal{E}$.

- Initialise the MAID's graph $(\mathbf{V}, \mathscr{E})$ as T, where edges are directed from parents to children.

- For each of $\mathcal{E}$'s chance nodes $V \in \mathbf{V}^0$, label each outgoing edge from V with a value v. Every other node with an outgoing edge is labelled, by λ, with the decision corresponding to that edge. Thus, let $dom(V)$ in $\mathcal{M}$ contain the labels of the outgoing edges from V.

- For each variable $V \in \mathbf{V}$ let ρ_V be the unique path formed by the sequence of labels from the root R of $\mathcal{E}$ to the corresponding EFG node for V and let $\rho_V[V']$ denote the label of the outgoing edge from node V' on the path ρ_V.

- For each information set I_j^i in $\mathcal{E}$, let $\rho^{\prec}(I_j^i)$ be the set of paths from R into the nodes of I_j^i. Next, define a function $\mu : \bigcup_{i \in N} I^i \to 2^{\mathbf{V}}$ that maps each information set I_j^i to the set of variables in $\mathbf{Anc}_{I_j^i}$ whose outgoing label is the same in every path in $\rho^{\prec}(I_j^i)$. Note that by Definition 29, if information sets I and I' are in the same intervention set, then the nodes whose values are in $\mu(I)$ are the same as the nodes whose values are in $\mu(I')$.

- We then consider $\mathcal{E}$'s chance, decision, and leaf nodes in turn:

  - For every outgoing edge with some label v_j from a chance node $V \in \mathbf{V}^0$, add the label $(v_j \mid \rho_{V_j}) : p$ where $p = P_j(v_j)$.

  - For every information set I_j^i and each variable $V \in I_j^i$, add the label $(v \mid \rho_V[\mu(I_j^i)]) : \_$ where $\rho_V[\mu(I_j^i)]$ denotes the labels of ρ_D restricted to those in $\mu(I_j^i)$ and $\_$ is a placeholder to be parameterised by a decision rule.

  - For each leaf variable $L \in \mathbf{L}$ with payoff vector $U(L) = (u^1, \dots, u^n)$, split L into n utility variables $U^1, \dots, U^n$ (duplicating incoming edges) with labels $(u^i \mid \rho_L) : 1$ respectively.

- Given these labellings we proceed by merging variables according to the intervention sets:

  - Merge each information set I_j^i into a single variable $D_j \in \boldsymbol{D}^i$, collecting the labels $(v \mid \rho_V[\mu(I_j^i)]) : \_$ for each $V \in I_j^i$ and retaining all incoming and outgoing edges.

  - Begin by adding a directed edge from every $V' \in \mathbf{Anc}_V$ to V for every variable V. Then for every variable D_j corresponding to an information set I_j^i, remove all the incoming edges from variables that are not in $\mu(I_j^i)$.

  - Merge each group of variables, collecting their incoming and outgoing labelled edges, that belong to the same intervention set.

  - Merge any utility variables U_j^i and U_k^i belonging to the same agent i that have the same sets of incoming edges, and collect their labels.

- We then let $\boldsymbol{V} := \boldsymbol{X} \cup \boldsymbol{D} \cup \boldsymbol{U}$ where $\boldsymbol{X} = \boldsymbol{V}^0$, $\boldsymbol{D}$ is the union of variable sets $\boldsymbol{D}^i$ defined above, and $\boldsymbol{U}$ is the collection of utility variables, also defined above.

- $\mathscr{E}$ is the set of edges in the graph defined above.

- We conclude by defining the CPDs $\Pr(V \mid \mathbf{Pa}_V)$, for every $V \in \boldsymbol{X} \cup \boldsymbol{U}$.

  - For certain instantiations, $\mathbf{pa}_V$, the CPD over V is undefined because $\mathbf{pa}_V$ does not represent a path through the original game tree T (as mentioned in Section 6.2). For non-decision variables, this is dealt with by simply adding a null value $\perp$ for every variable in $\boldsymbol{X}$ and a value of 0 for every variable in $\boldsymbol{U}$.

  - Recall the labels of the form $(v \mid \rho_V) : p$ and let $l(V)$ denote the set of V's labels.

    * For each variable $V \in \boldsymbol{V} \setminus \boldsymbol{D}$, we define $\Pr(v \mid \mathbf{pa}_V) := \sum_{(v\mid\mathbf{pa}_V):p \,\in\, l(V)} p$.

    * For any $\mathbf{pa}_V$ such that for all $v \in dom(V)$, $\Pr(v \mid \mathbf{pa}_V) = 0$, we set $\Pr(\perp \mid \mathbf{pa}_V) = 1$ if $V \in \boldsymbol{X}$ and $\Pr(0 \mid \mathbf{pa}_V) = 1$ if $V \in \boldsymbol{U}$.

- By construction, $\prod_{V \in \boldsymbol{V} \setminus \boldsymbol{D}} \Pr(v \mid \mathbf{pa}_V)$ therefore forms a partial distribution over the non-decision variables in $\mathcal{M}$.

- For decision variables, we now see that the only changes in parameterisations of $\pi_D(D \mid \mathbf{Pa}_D)$ that can have any effect on any of the other variables are those that occur under settings $\mathbf{pa}_D$ such that there exists a strategy σ and path ρ in $\mathcal{E}$ capturing all values in $\mathbf{pa}_D$ with $\Pr^\sigma(\rho) > 0$. In other words, those values of $\pi_D(d \mid \mathbf{pa}_D)$ corresponding to a parametrisation of $(d \mid \rho_D[\mu(I_j^i)]) : \_$.

A.2 Theoretical Results

Section 3

Proposition 3. M_V *is* $\mathcal{R}^{BR}$*-relevant to* Π_D *if and only if* $\mathsf{M}_V \not\perp_{\mathsf{m}_\perp \mathcal{G}} \boldsymbol{U}^i \cap \boldsymbol{Desc}_D \mid D, \boldsymbol{Pa}_D$ *or* $\mathsf{M}_V \not\perp_{\mathsf{m}_\perp \mathcal{G}} \boldsymbol{Pa}_D$*, where if* $D \in \boldsymbol{D}^i$*, then* $\boldsymbol{U}^i \cap \boldsymbol{Desc}_D \neq \varnothing$*.*

Proof. First, recall from Definition 17, that $\mathsf{M}_V \in \mathbf{Pa}_{\Pi_D}$ is $\mathcal{R}^{\mathrm{BR}}$-**relevant** to Π_D if there exists $\mathbf{pa}_{\Pi_D} \neq \mathbf{pa}'_{\Pi_D}$ such that $r_D^{\mathrm{BR}}(\mathbf{pa}_{\Pi_D}) \neq r_D^{\mathrm{BR}}(\mathbf{pa}'_{\Pi_D})$, where $\mathbf{pa}_{\Pi_D}$ and $\mathbf{pa}'_{\Pi_D}$ differ only on M_V. Moreover, recall that for each $D \in \boldsymbol{D}^i$, $\pi_D \in r_D^{\mathrm{BR}}(\mathbf{pa}_{\Pi_D})$ if and only if:

$$\pi_D \in \operatorname*{arg\,max}_{\hat{\pi}_D \in dom(\Pi_D)} \sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{(\hat{\pi}_D, \boldsymbol{\pi}_{-D})}[U].$$

Given some setting $\mathbf{m}_{-D}$ of all mechanism variables apart from Π_D, then for each $U \in \boldsymbol{U}^i$ and any $\hat{\pi}_D \in dom(\Pi_D)$ we have that:

$$\Pr^{(\hat{\pi}_D, \boldsymbol{\pi}_{-D})}(u) = \sum_{d \in dom(D)} \sum_{\mathbf{pa}_D \in dom(\mathbf{Pa}_D)} \Pr^{(\hat{\pi}_D, \boldsymbol{\pi}_{-D})}(u, d, \mathbf{pa}_D)$$

$$= \sum_{d \in dom(D)} \sum_{\mathbf{pa}_D \in dom(\mathbf{Pa}_D)} \Pr^{\boldsymbol{\pi}_{-D}}(u \mid d, \mathbf{pa}_D) \hat{\pi}_D(d \mid \mathbf{pa}_D) \Pr^{\boldsymbol{\pi}_{-D}}(\mathbf{pa}_D)$$

$$= \sum_{d \in dom(D)} \sum_{\mathbf{pa}_D \in dom(\mathbf{Pa}_D)} \Pr^{\boldsymbol{\pi}}(u \mid d, \mathbf{pa}_D) \hat{\pi}_D(d \mid \mathbf{pa}_D) \Pr^{\boldsymbol{\pi}}(\mathbf{pa}_D)$$

The first equality simply rewrites the marginal distribution as a sum over a joint distribution; the second factorises this joint distribution; and the third results from the observations that the distribution over $\mathbf{Pa}_D$ cannot depend on the CPD $\pi_D(D \mid \mathbf{Pa}_D)$, and nor can the distribution over U *given* some $d, \mathbf{pa}_D$. Next, note that in the definition of r^{BR} we quantify over $\hat{\pi}_D$ via the arg max operation, and that the choice of $\hat{\pi}_D$ (given $\boldsymbol{\pi}_{-D}$) can only impact the expected value of those utility variables $U \in \boldsymbol{U}^i$ that are also in $\mathbf{Desc}_D$. Therefore, any decision rule that is in $r_D^{\mathrm{BR}}(\mathbf{pa}_{\Pi_D})$ satisfies:

$$\arg\max_{\hat{\pi}_D \in dom(\Pi_D)} \left[\sum_{U \in \boldsymbol{U}^i \cap \mathbf{Desc}_D} \sum_{d \in dom(D)} \sum_{\mathbf{pa}_D \in dom(\mathbf{Pa}_D)} \Pr^{\boldsymbol{\pi}}(u \mid d, \mathbf{pa}_D) \hat{\pi}_D(d \mid \mathbf{pa}_D) \Pr^{\boldsymbol{\pi}}(\mathbf{pa}_D) \cdot u \right]. \quad (4)$$

Therefore, if the expression (4) is independent of the value m_V, then M_V is not $\mathcal{R}^{\mathrm{BR}}$-**relevant** to Π_D. As explained in Section 3.2, this is equivalent to asking whether V is a requisite probability node for any distribution in $\mathcal{Q}_D^{\mathrm{BR}} = \{\Pr^{\boldsymbol{\pi}}(\boldsymbol{u}^i \cap \mathbf{desc}_D \mid d, \mathbf{pa}_D), \Pr^{\boldsymbol{\pi}}(\mathbf{pa}_D)\}$. The graphical criterion for checking whether something is a requisite probability node is given in Lemma 1, and results in the criteria (which we refer to as $\mathcal{R}^{\mathrm{BR}}$-reachability) $\mathsf{M}_V \not\perp_{\mathsf{m}_\perp \mathcal{G}} \boldsymbol{U}^i \cap \mathbf{Desc}_D \mid D, \mathbf{Pa}_D$ or $\mathsf{M}_V \not\perp_{\mathsf{m}_\perp \mathcal{G}} \mathbf{Pa}_D$, where if $D \in \boldsymbol{D}^i$, then $\boldsymbol{U}^i \cap \mathbf{Desc}_D \neq \varnothing$.

We begin by proving soundness, i.e., that if M_V is not $\mathcal{R}^{\mathrm{BR}}$-reachable from Π_D, then M_V is not $\mathcal{R}^{\mathrm{BR}}$-relevant to Π_D. Let $\mathbf{m} \neq \mathbf{m}'$ differ only on the value of $\mathsf{M}_V \in \mathbf{Pa}_{\Pi_D}$, and let $\mathbf{pa}_{\Pi_D} \neq \mathbf{pa}'_{\Pi_D}$ be the respective settings of $\mathbf{Pa}_{\Pi_D}$. Additionally, let us write $\mathbf{m} = (\boldsymbol{\pi}, \boldsymbol{\theta})$ and $\mathbf{m}' = (\boldsymbol{\pi}', \boldsymbol{\theta}')$. If M_V is not $\mathcal{R}^{\mathrm{BR}}$-reachable from Π_D then we have that both $\mathsf{M}_V \perp_{\mathsf{m}_\perp \mathcal{G}} \boldsymbol{U}^i \cap \mathbf{Desc}_D \mid D, \mathbf{Pa}_D$ and $\mathsf{M}_V \perp_{\mathsf{m}_\perp \mathcal{G}} \mathbf{Pa}_D$, implying that both $\Pr^{\boldsymbol{\pi}}(u \mid d, \mathbf{pa}_D; \boldsymbol{\theta}) = \Pr^{\boldsymbol{\pi}'}(u \mid d, \mathbf{pa}_D; \boldsymbol{\theta}')$ and $\Pr^{\boldsymbol{\pi}}(\mathbf{pa}_D; \boldsymbol{\theta}) = \Pr^{\boldsymbol{\pi}'}(\mathbf{pa}_D; \boldsymbol{\theta}')$. Thus, the value of expression (4) is independent of the value of M_V and so $r_D^{\mathrm{BR}}(\mathbf{pa}_{\Pi_D}) = r_D^{\mathrm{BR}}(\mathbf{pa}'_{\Pi_D})$, meaning that M_V is not $\mathcal{R}^{\mathrm{BR}}$-relevant to Π_D, as required.

We now turn to completeness, i.e., that if M_V is $\mathcal{R}^{\mathrm{BR}}$-reachable from Π_D in some mechanised graph $\mathsf{m}\mathcal{G}$, where $D \in \boldsymbol{D}^i$ and $\boldsymbol{U}^i \cap \mathbf{Desc}_D \neq \varnothing$, then for some mechanised MAID $\mathsf{m}\mathcal{M} = (\mathsf{m}\mathcal{G}, \boldsymbol{\theta}, \mathcal{R}^{\mathrm{BR}})$, M_V is $\mathcal{R}^{\mathrm{BR}}$-relevant to Π_D. Our goal is therefore to find some parameterisation of the mechanised graph $\mathsf{m}\mathcal{G}$ such that if the parameterisation of Π_D's parents $\mathbf{pa}_{\Pi_D}$ is changed to $\mathbf{pa}'_{\Pi_D}$, where these only differ on M_V, then $r_D(\mathbf{pa}_{\Pi_D}) \neq r_D(\mathbf{pa}'_{\Pi_D})$.

If $\mathsf{M}_V \not\perp_{\mathsf{m}_\perp \mathcal{G}} \boldsymbol{U}^i \cap \mathbf{Desc}_D \mid D, \mathbf{Pa}_D$, then we can simply follow K&M's proof of completeness for s-reachability to construct a parameterisation of the mechanised graph such that M_V is s-relevant, and hence $\mathcal{R}^{\mathrm{BR}}$-relevant, to Π_D in the mechanised MAID. If, instead, we only have $\mathsf{M}_V \not\perp_{\mathsf{m}_\perp \mathcal{G}} \mathbf{Pa}_D$, then we proceed as follows. Below, we restrict our attention to those decision nodes that satisfy $\boldsymbol{U}^i \cap \mathbf{Desc}_D \neq \varnothing$ for $D \in \boldsymbol{D}^i$. This is because if $\boldsymbol{U}^i \cap \mathbf{Desc}_D = \varnothing$, then agent i will be indifferent between all of its decision rules for D and so trivially we will have $r_D^{\mathrm{BR}}(\mathbf{pa}_{\Pi_D}) = r_D^{\mathrm{BR}}(\mathbf{pa}'_{\Pi_D})$ for any $\mathbf{pa}_{\Pi_D}$ and $\mathbf{pa}'_{\Pi_D}$; in this case, M_V is never $\mathcal{R}^{\mathrm{BR}}$-relevant to Π_D.

Let us assume that $\mathsf{M}_V \not\perp_{\mathsf{m}_\perp \mathcal{G}} Z$ for some $Z \in \mathbf{Pa}_D$. Then, as M_V has no parents in $\mathsf{m}_\perp \mathcal{G}$ and all paths between M_V and Z that contain colliders are blocked, there must be a directed path

$\mathsf{M}_V \dashrightarrow Z$ in $\mathsf{m}_\perp \mathcal{G}$. Let us denote the variables along this directed path by $X_0 \dashrightarrow X_{n+1}$, where $X_0 = \mathsf{M}_V$ and $X_{n+1} = Z$. Next, following our remarks above, we know that $\boldsymbol{U}^i \cap \mathbf{Desc}_D \neq \varnothing$, and so there is a directed path $D \dashrightarrow U$ for some $U \in \boldsymbol{U}^i$. We denote the variables along this path by $Y_0 \dashrightarrow Y_{m+1}$, where $Y_0 = D$ and $Y_{m+1} = U$. Thus, in $\mathsf{m}_\perp \mathcal{G}$ we have a directed path $X_0 \dashrightarrow X_{n+1} \to Y_0 \dashrightarrow Y_{m+1}$ where the segments $X_1 \dashrightarrow X_n$ and $Y_1 \dashrightarrow Y_m$ might be empty.

We now give two parameterisations of $\mathsf{m}\mathcal{G}$ that are identical except for the two settings $\mathbf{pa}_{\Pi_D} \neq \mathbf{pa}'_{\Pi_D}$ that differ only on M_V. We begin with what the two parameterisations have in common. First, we let all non-utility variables be binary, although we will show how this assumption can easily be relaxed at the end. Next, we set the CPDs for all nodes along the paths $X_2 \dashrightarrow X_{n+1}$ and $Y_1 \dashrightarrow Y_m$ to copy the value of their parent in this path. That is, for $i \in \{2, n\}$, $x_i = x_{i-1}$ with probability 1, and for $i \in \{1, m\}$, $y_i = y_{i-1}$ with probability 1. Note that the value of $X_1 = V$ is determined according to the value of M_V. Finally, we set the distributions of the utility functions so that U copies the value of Y_m and any other $U' \in \boldsymbol{U}^i$ takes value 0 regardless of the value of $\mathbf{Pa}_{U'}$.

For $\mathbf{pa}_{\Pi_D}$ we let m_V set the distribution over V to $\delta(V, 1)$, implying that $X_{n+1} = Z$ takes value 1. Given some decision rule $\pi_D \in r_D^{\mathrm{BR}}(\mathbf{pa}_{\Pi_D})$, note that the decision rule $\bar{\pi}_D \in r_D^{\mathrm{BR}}(\mathbf{pa}_{\Pi_D})$ too, where $\bar{\pi}_D$ is identical to π_D apart from the fact that $\bar{\pi}_D(D = 0 \mid Z = 0, \mathbf{pa}'_D) = 1$ where $\mathbf{pa}'_D$ is any setting of the variables $\mathbf{Pa}'_D = \mathbf{Pa}_D \setminus \{Z\}$. When $D = 0$ then, by the construction above, $Y_{m+1} = U = 0$, but as $Z = 0$ with probability zero given m_V, then this outcome never occurs, and so $\bar{\pi}_D$ remains a rational response to $\mathbf{pa}_{\Pi_D}$.

For $\mathbf{pa}'_{\Pi_D}$, we let m'_V set the distribution over V to be such that $\mathrm{Pr}^{\boldsymbol{\pi}}(V = 0 \mid \mathbf{pa}_V) > 0$ for any $\mathbf{pa}_V$. In this case, $\bar{\pi}_D$ is *not* a rational response, as the context $Z = 0$ occurs with non-zero probability, and thus so to does $U = 0$, due to the construction above and the fact that $\bar{\pi}_D(D = 0 \mid Z = 0, \mathbf{pa}'_D) = 1$. Agent i's expected utility could instead be strictly improved by selecting the decision rule $\delta(D, 1)$. Thus, $\bar{\pi}_D \notin r_D^{\mathrm{BR}}(\mathbf{pa}'_{\Pi_D})$, and hence $r_D^{\mathrm{BR}}(\mathbf{pa}_{\Pi_D}) \neq r_D^{\mathrm{BR}}(\mathbf{pa}'_{\Pi_D})$, i.e., M_V is $\mathcal{R}^{\mathrm{BR}}$-relevant to Π_D, as required. To extend the above proof to the case where variables have arbitrary (as opposed to binary) domains, we simply label one value in the domain of each variable by '1' and another by '0', making sure that for the utility variables the value labelled as '1' is strictly greater than the value labelled as '0'. $\qquad \square$

Section 4

Proposition 4. *For distributions over E_D and D as governed by equations ([3](#)), there is a one-to-one correspondence between the set of stochastic decision rules $\Delta(dom(D) \mid dom(\overline{\boldsymbol{Pa}}_D))$ and the set of deterministic decision rules $dom(\dot{\Pi}_D) \subset \Delta(dom(D) \mid dom(\boldsymbol{Pa}_D))$. Moreover, given two such corresponding decision rules π_D and $\dot{\pi}_D$, then $\int_{dom(\mathsf{E}_D)} \dot{\pi}_D(d \mid \overline{\boldsymbol{pa}}_D, \mathsf{e}_D) \mathrm{Pr}(\mathsf{e}_D) \, d\mathsf{e}_D = \pi_D(d \mid \overline{\boldsymbol{pa}}_D)$.*

Proof. Recall from Section [4.3](#) that we define $\overline{\mathbf{Pa}}_D := \mathbf{Pa}_D \setminus \{\mathsf{E}_D\}$. Let us first examine the correspondence between the stochastic and non-stochastic decision rules. Let $f : \Delta(dom(D) \mid dom(\overline{\mathbf{Pa}}_D)) \to dom(\dot{\Pi}_D)$ be a function where for $\pi_D \in \Delta(dom(D) \mid dom(\overline{\mathbf{Pa}}_D))$ we define $f(\pi_D) = \dot{\pi}_D$ such that $\dot{\pi}_D(d \mid \overline{\mathbf{pa}}_D, \mathsf{e}_D) = \delta(d, \mathsf{e}_D^{\pi_D, \overline{\mathbf{pa}}_D})$. Given our definition of E_D then for any $\pi_D \neq \pi'_D$ then we have that $f(\pi_D) = f(\pi'_D)$ if and only if $\mathsf{E}_D^{\pi_D, \overline{\mathbf{pa}}_D} = \mathsf{E}_D^{\pi'_D, \overline{\mathbf{pa}}_D}$ for all $\overline{\mathbf{pa}}_D$. Clearly, however, for $\pi_D \neq \pi'_D$ then there exists some $\overline{\mathbf{pa}}_D$ and $d \in dom(\mathsf{E}_D^{\pi_D, \overline{\mathbf{pa}}_D}) = dom(\mathsf{E}_D^{\pi'_D, \overline{\mathbf{pa}}_D})$ such that:

$$\mathrm{Pr}(\mathsf{E}_D^{\pi_D, \overline{\mathbf{pa}}_D} = d) := \pi_D(d \mid \overline{\mathbf{pa}}_D) \neq \pi'_D(d \mid \overline{\mathbf{pa}}_D) =: \mathrm{Pr}(\mathsf{E}_D^{\pi'_D, \overline{\mathbf{pa}}_D} = d),$$

and so $f(\pi_D) \neq f(\pi'_D)$, hence f is an injection. Moreover, any $\dot{\pi}_D$ of this form can be identified with the set $\{\mathsf{E}_D^{\pi_D, \overline{\mathbf{pa}}_D}\}_{\overline{\mathbf{pa}}_D \in dom(\overline{\mathbf{Pa}}_D)} \subseteq \mathsf{E}_D$ and hence some $\pi_D \in \Delta(dom(D) \mid dom(\overline{\mathbf{Pa}}_D))$ such

that $f(\pi_D) = \dot{\pi}_D$, hence f is a surjection, and therefore a bijection, giving us our one-to-one correspondence.

Now take some $\pi_D \in \Delta(dom(D) \mid dom(\overline{\mathbf{Pa}}_D))$ and $\dot{\pi}_D = f(\pi_D)$. Let $\mathsf{E}_D^{-\pi_D, \overline{\mathbf{pa}}_D} := \mathsf{E}_D \setminus \{\mathsf{E}_D^{\pi_D, \overline{\mathbf{pa}}_D}\}$. Then, by our definition of E_D and its joint product of probability distributions [7], we have that:

$$\int_{dom(\mathsf{E}_D)} \dot{\pi}_D(d \mid \overline{\mathbf{pa}}_D, \mathsf{e}_D) \Pr(\mathsf{e}_D) \, d\mathsf{e}_D$$

$$= \int_{dom(\mathsf{E}_D)} \delta(d, \mathsf{e}_D^{\pi_D, \overline{\mathbf{pa}}_D}) \Pr(\mathsf{e}_D) \, d\mathsf{e}_D$$

$$= \sum_{\mathsf{e}_D^{\pi_D, \overline{\mathbf{pa}}_D} \in dom(\mathsf{E}_D^{\pi_D, \overline{\mathbf{pa}}_D})} \delta(d, \mathsf{e}_D^{\pi_D, \overline{\mathbf{pa}}_D}) \Pr(\mathsf{e}_D^{\pi_D, \overline{\mathbf{pa}}_D}) \cdot \int_{dom(\mathsf{E}_D^{-\pi_D, \overline{\mathbf{pa}}_D})} \Pr(\mathsf{e}_D^{-\pi_D, \overline{\mathbf{pa}}_D}) \, d\mathsf{e}_D^{-\pi_D, \overline{\mathbf{pa}}_D}$$

$$= \sum_{\mathsf{e}_D^{\pi_D, \overline{\mathbf{pa}}_D} \in dom(\mathsf{E}_D^{\pi_D, \overline{\mathbf{pa}}_D})} \delta(d, \mathsf{e}_D^{\pi_D, \overline{\mathbf{pa}}_D}) \pi_D(\mathsf{e}_D^{\pi_D, \overline{\mathbf{pa}}_D} \mid \overline{\mathbf{pa}}_D) \cdot 1$$

$$= \pi_D(d \mid \overline{\mathbf{pa}}_D).$$

$\square$

Section 5

Proposition 5. *A (behavioural policy) NE is not guaranteed to exist in a MAID.*

Proof. See Appendix B.2 for an example of a MAID that does not have a (behavioural policy) NE. $\square$

Lemma 2. *Let $\boldsymbol{\pi}^{-i} \in \Delta(dom(\boldsymbol{D}^{-i}) \mid dom(\boldsymbol{Pa}_{\boldsymbol{D}^{-i}}))$ be a partial (behavioural or mixed) policy profile for agents $N \setminus \{i\}$ in a MAID $\mathcal{M}$. If agent i has perfect recall in $\mathcal{M}$, then for every mixed policy μ^i there exists a behavioural policy $\boldsymbol{\pi}^i$ such that $\Pr^{(\mu^i, \boldsymbol{\pi}^{-i})}(\boldsymbol{v}) = \Pr^{(\boldsymbol{\pi}^i, \boldsymbol{\pi}^{-i})}(\boldsymbol{v})$.*

Proof. For the purposes of this proof, let us abuse notation by viewing each pure decision rule $\dot{\pi}_D$ as a function, so that $\dot{\pi}_D(\mathbf{pa}_D) = d$ just in case $\dot{\pi}_D(d \mid \mathbf{pa}_D) = 1$. Further, for some pure policy $\dot{\boldsymbol{\pi}}^i \in \bigtimes_{D \in \boldsymbol{D}^i} \dot{\Pi}_D$, let us write $\dot{\boldsymbol{\pi}}^i(\mathbf{pa}_{\boldsymbol{D}^i}) = \boldsymbol{d}^i$ just in case $\dot{\pi}_D(\mathbf{pa}_D) = d$ for every $D \in \boldsymbol{D}^i$. Then, given a fixed MAID $\mathcal{M}$ and a partial (behavioural or mixed) policy profile $\boldsymbol{\pi}^{-i}$, we have the partial distribution $\Pr^{\boldsymbol{\pi}^{-i}}(\boldsymbol{v}, \boldsymbol{d}^{-i} : \boldsymbol{d}^i) = \prod_{V \in \boldsymbol{V} \setminus \boldsymbol{D}^i} \Pr^{\boldsymbol{\pi}^{-i}}(v \mid \mathbf{pa}_V)$. It therefore suffices to show that for any $\mu^i \in \Delta(\bigtimes_{D \in \boldsymbol{D}^i} dom(\dot{\Pi}_D))$, there exists some $\boldsymbol{\pi}^i \in \bigtimes_{D \in \boldsymbol{D}^i} \Pi_D$ such that:

$$\mu^i(\boldsymbol{d}^i \mid \mathbf{pa}_{\boldsymbol{D}^i}) := \sum_{\dot{\boldsymbol{\pi}}^i \in \bigtimes_{D \in \boldsymbol{D}^i} \dot{\Pi}_D \, : \, \dot{\boldsymbol{\pi}}^i(\mathbf{pa}_{\boldsymbol{D}^i}) = \boldsymbol{d}^i} \mu^i(\dot{\boldsymbol{\pi}}^i) = \prod_{D \in \boldsymbol{D}^i} \pi_D(d \mid \mathbf{pa}_D) =: \boldsymbol{\pi}^i(\boldsymbol{d}^i \mid \mathbf{pa}_{\boldsymbol{D}^i}),$$

where again we abuse notation somewhat by using μ^i and $\boldsymbol{\pi}^i$ to denote both policies and the resulting partial distributions over variables in $\mathcal{M}$. In what follows, we also suppress the qualification that $\dot{\boldsymbol{\pi}}^i \in \bigtimes_{D \in \boldsymbol{D}^i} \dot{\Pi}_D$, which we take to be implicit when considering $\dot{\boldsymbol{\pi}}^i$. Suppose that $|\boldsymbol{D}^i| = m$ and let us assume, without loss of generality, that the indices $1, \ldots, m$ of these decision rules reflect the unique (in virtue of the fact that each agent has perfect recall) topological ordering of variables $\boldsymbol{D}^i$ in $\mathcal{M}$; i.e., $D_1 \prec \cdots \prec D_m$. In the remainder of the proof, we write $\boldsymbol{D}^i_{\leq j}$ to denote the set of variables $\{D_1, \ldots, D_j\} \subseteq \boldsymbol{D}^i$. Then for each $D_j \in \boldsymbol{D}^i$ we have that $\mathbf{Fa}_{\boldsymbol{D}_{<j}} \subseteq \mathbf{Pa}_{D_j}$. We now define each decision rule $\pi^i_{D_j}$ as follows:

$$\pi^i_{D_j}(d_j \mid \mathbf{pa}_{D_j}) := \begin{cases} \frac{1}{|dom(D_j)|} & \text{if } Z_j = 0 \\ \frac{1}{Z_j} \sum_{\dot{\boldsymbol{\pi}}^i_{\boldsymbol{D}_{\leq j}} \, : \, \dot{\boldsymbol{\pi}}^i_{\boldsymbol{D}_{\leq j}}(\mathbf{pa}_{\boldsymbol{D}^i_{\leq j}}) = \boldsymbol{d}^i_{\leq j}} \mu^i(\dot{\boldsymbol{\pi}}_{\boldsymbol{D}_{\leq j}}) & \text{otherwise,} \end{cases}$$

where:

$$Z_j := \begin{cases} 1 & \text{if } j = 1 \\ \sum_{\dot{\boldsymbol{\pi}}^i_{\boldsymbol{D}_{<j}} \;:\; \dot{\boldsymbol{\pi}}^i_{\boldsymbol{D}_{<j}}(\mathbf{pa}_{\boldsymbol{D}^i_{<j}})=\boldsymbol{d}^i_{<j}} \mu^i(\dot{\boldsymbol{\pi}}_{\boldsymbol{D}_{<j}}) & \text{otherwise.} \end{cases}$$

Note that this behavioural decision rule is well-formed. Concretely, $\pi^i_{D_j}(d_j \mid \mathbf{pa}_{D_j}) \geq 0$ as $\mu^i(\dot{\boldsymbol{\pi}}_{\boldsymbol{D}_{\leq j}}) \geq 0$ and thus $Z_j > 0$ whenever used to normalise $\pi^i_{D_j}(d_j \mid \mathbf{pa}_{D_j})$, and $\sum_{d_j} \pi^i_{D_j}(d_j \mid \mathbf{pa}_{D_j}) = 1$ as:

$$\sum_{d_j} \left(\sum_{\dot{\boldsymbol{\pi}}^i_{\boldsymbol{D}_{\leq j}} \;:\; \dot{\boldsymbol{\pi}}^i_{\boldsymbol{D}_{\leq j}}(\mathbf{pa}_{\boldsymbol{D}^i_{\leq j}})=\boldsymbol{d}^i_{\leq j}} \mu^i(\dot{\boldsymbol{\pi}}_{\boldsymbol{D}_{\leq j}}) \right) = \sum_{\dot{\boldsymbol{\pi}}^i_{\boldsymbol{D}_{<j}} \;:\; \dot{\boldsymbol{\pi}}^i_{\boldsymbol{D}_{<j}}(\mathbf{pa}_{\boldsymbol{D}^i_{<j}})=\boldsymbol{d}^i_{<j}} \mu^i(\dot{\boldsymbol{\pi}}_{\boldsymbol{D}_{<j}}).$$

Using this definition, we may now conclude the proof, proceeding by cases. First, suppose that $Z_j = 0$ for some $j \geq 2$. Then:

$$\sum_{\dot{\boldsymbol{\pi}}^i_{\boldsymbol{D}_{<j}} \;:\; \dot{\boldsymbol{\pi}}^i_{\boldsymbol{D}_{<j}}(\mathbf{pa}_{\boldsymbol{D}^i_{<j}})=\boldsymbol{d}^i_{<j}} \mu^i(\dot{\boldsymbol{\pi}}_{\boldsymbol{D}_{<j}}) = 0 \Rightarrow \pi_{D_{j-1}}(d_{j-1} \mid \mathbf{pa}_{D_{j-1}}) = 0 \Rightarrow \boldsymbol{\pi}^i(\boldsymbol{d}^i \mid \mathbf{pa}_{\boldsymbol{D}^i}) = 0.$$

But the first antecedent above also implies that $\mu^i(\boldsymbol{d}^i \mid \mathbf{pa}_{\boldsymbol{D}^i}) := \sum_{\dot{\boldsymbol{\pi}}^i \;:\; \dot{\boldsymbol{\pi}}^i(\mathbf{pa}_{\boldsymbol{D}^i})=\boldsymbol{d}^i} \mu^i(\dot{\boldsymbol{\pi}}^i) = 0$, and so we have $\boldsymbol{\pi}^i(\boldsymbol{d}^i \mid \mathbf{pa}_{\boldsymbol{D}^i}) = \mu^i(\boldsymbol{d}^i \mid \mathbf{pa}_{\boldsymbol{D}^i})$, as required. Now suppose that there is no j such that $Z_j = 0$. Then, by telescoping terms we have the following:

$$\begin{aligned} \boldsymbol{\pi}^i(\boldsymbol{d}^i \mid \mathbf{pa}_{\boldsymbol{D}^i}) &:= \prod_{j=1}^m \pi_{D_j}(d_j \mid \mathbf{pa}_{D_j}) \\ &= \prod_{j=1}^m \frac{1}{Z_j} \sum_{\dot{\boldsymbol{\pi}}^i_{\boldsymbol{D}_{\leq j}} \;:\; \dot{\boldsymbol{\pi}}^i_{\boldsymbol{D}_{\leq j}}(\mathbf{pa}_{\boldsymbol{D}^i_{\leq j}})=\boldsymbol{d}^i_{\leq j}} \mu^i(\dot{\boldsymbol{\pi}}_{\boldsymbol{D}_{\leq j}}) \\ &= \sum_{\dot{\boldsymbol{\pi}}^i_{\boldsymbol{D}_{\leq m}} \;:\; \dot{\boldsymbol{\pi}}^i_{\boldsymbol{D}_{\leq m}}(\mathbf{pa}_{\boldsymbol{D}^i_{\leq m}})=\boldsymbol{d}^i_{\leq m}} \mu^i(\dot{\boldsymbol{\pi}}_{\boldsymbol{D}_{\leq m}}) \\ &=: \mu^i(\boldsymbol{d}^i \mid \mathbf{pa}_{\boldsymbol{D}^i}). \end{aligned}$$

$\square$

Proposition 6. *In any MAID $\mathcal{M}$ with perfect recall, there exists a (behavioural) policy profile $\boldsymbol{\pi}$ that is an NE.*

Proof. First, note that an immediate consequence of Lemma 2 is that, for a given partial policy $\boldsymbol{\pi}^{-i}$ (behavioural or mixed) in a MAID $\mathcal{M}$ with perfect recall, then under any mixed policy μ^i and equivalent behavioural policy $\boldsymbol{\pi}^i$ we have:

$$\sum_{U \in \boldsymbol{U}^k} \mathbb{E}_{(\mu^i, \boldsymbol{\pi}^{-i})}[U] = \sum_{U \in \boldsymbol{U}^k} \mathbb{E}_{(\boldsymbol{\pi}^i, \boldsymbol{\pi}^{-i})}[U],$$

for all agents $k \in N$. Next, by a straightforward application of Nash's theorem [68], we know that every MAID is guaranteed to have an NE μ in mixed policies. Given $\mu = (\mu^1, \ldots, \mu^n)$, we can therefore apply Lemma 2 n times to result in a behavioural policy $\boldsymbol{\pi} = (\boldsymbol{\pi}^1, \ldots, \boldsymbol{\pi}^n)$ such that $\sum_{U \in \boldsymbol{U}^k} \mathbb{E}_\mu[U] = \sum_{U \in \boldsymbol{U}^k} \mathbb{E}_{\boldsymbol{\pi}}[U]$ for all agents $k \in N$. Finally, let us assume (for a contradiction) that there is some agent $i \in N$ that may deviate from $\boldsymbol{\pi}^i$ in order to increase their expected utility. Let this deviation be denoted by $\hat{\boldsymbol{\pi}}^i$. Then we can construct a mixed

policy $\hat{\mu}^i \in \Delta(\times_{D \in \boldsymbol{D}^i} dom(\dot{\Pi}_D))$ that results in the same joint distribution as $\hat{\boldsymbol{\pi}}^i$ in the following manner. We define $\hat{\mu}^i(\dot{\boldsymbol{\pi}}^i) := \prod_{j=1}^m \hat{\mu}^i(\dot{\pi}_{D_j}^i)$ where we set:

$$\hat{\mu}^i(\dot{\pi}_{D_j}^i) := \prod_{\mathbf{pa}'_{D_j} \in dom(\mathbf{Pa}_{D_j})} \hat{\pi}_j^i\big(\dot{\pi}_{D_j}^i(\mathbf{pa}'_{D_j}) \mid \mathbf{pa}'_{D_j}\big).$$

Again, note that this results in a well-formed mixed policy. Concretely, $\hat{\mu}^i(\dot{\boldsymbol{\pi}}^i) \geq 0$ for any $\dot{\boldsymbol{\pi}}^i$ as $\hat{\pi}_j^i\big(\dot{\pi}_{D_j}^i(\mathbf{pa}'_{D_j}) \mid \mathbf{pa}'_{D_j}\big) \geq 0$, and we have:

$$\sum_{\dot{\boldsymbol{\pi}}^i} \hat{\mu}^i(\dot{\boldsymbol{\pi}}^i) = \sum_{\dot{\pi}_{D_1}^i} \cdots \sum_{\dot{\pi}_{D_m}^i} \prod_{j=1}^m \hat{\mu}^i(\dot{\pi}_{D_j}^i)$$

$$= \prod_{j=1}^m \sum_{\dot{\pi}_{D_j}^i} \hat{\mu}^i(\dot{\pi}_{D_j}^i)$$

$$= \prod_{j=1}^m \sum_{\dot{\pi}_{D_j}^i} \left(\prod_{\mathbf{pa}'_{D_j} \in dom(\mathbf{Pa}_{D_j})} \hat{\pi}_j^i\big(\dot{\pi}_{D_j}^i(\mathbf{pa}'_{D_j}) \mid \mathbf{pa}'_{D_j}\big) \right)$$

$$= \prod_{j=1}^m \left(\prod_{\mathbf{pa}'_{D_j} \in dom(\mathbf{Pa}_{D_j})} \left(\sum_{d'_j \in dom(D_j)} \left(\sum_{\dot{\pi}_{D_j}^i \, : \, \dot{\pi}_{D_j}^i(\mathbf{pa}'_{D_j})=d'_j} \hat{\pi}_j^i\big(\dot{\pi}_{D_j}^i(\mathbf{pa}'_{D_j}) \mid \mathbf{pa}'_{D_j}\big) \right) \right) \right)$$

$$= \prod_{j=1}^m \left(\prod_{\mathbf{pa}'_{D_j} \in dom(\mathbf{Pa}_{D_j})} 1 \right)$$

$$= 1.$$

In order to show that policy profiles $(\hat{\mu}^i, \boldsymbol{\pi}^{-i})$ and $(\hat{\boldsymbol{\pi}}^i, \boldsymbol{\pi}^{-i})$ result in the same joint distribution over all variables (and hence the same expected payoffs for each agent), it suffices to show the following, which holds for any $j \in \{1, \ldots, m\}$:

$$\hat{\mu}_j^i(d_j \mid \mathbf{pa}_{D_j}) := \sum_{\dot{\pi}_{D_j}^i \, : \, \dot{\pi}_{D_j}^i(\mathbf{pa}_{D_j})=d_j} \hat{\mu}^i(\dot{\pi}_{D_j}^i)$$

$$= \sum_{\dot{\pi}_{D_j}^i \, : \, \dot{\pi}_{D_j}^i(\mathbf{pa}_{D_j})=d_j} \left(\prod_{\mathbf{pa}'_{D_j} \in dom(\mathbf{Pa}_{D_j})} \hat{\pi}_j^i\big(\dot{\pi}_{D_j}^i(\mathbf{pa}'_{D_j}) \mid \mathbf{pa}'_{D_j}\big) \right)$$

$$= \sum_{\dot{\pi}_{D_j}^i \, : \, \dot{\pi}_{D_j}^i(\mathbf{pa}_{D_j})=d_j} \hat{\pi}_j^i(d_j \mid \mathbf{pa}_{D_j}) \left(\prod_{\mathbf{pa}'_{D_j} \in dom(\mathbf{Pa}_{D_j}) \backslash \{\mathbf{pa}_{D_j}\}} \hat{\pi}_j^i\big(\dot{\pi}_{D_j}^i(\mathbf{pa}'_{D_j}) \mid \mathbf{pa}'_{D_j}\big) \right)$$

$$= \hat{\pi}_j^i(d_j \mid \mathbf{pa}_{D_j}) \cdot \prod_{\mathbf{pa}'_{D_j} \in dom(\mathbf{Pa}_{D_j}) \backslash \{\mathbf{pa}_{D_j}\}} \left(\sum_{\dot{\pi}_{D_j}^i \, : \, \dot{\pi}_{D_j}^i(\mathbf{pa}_{D_j})=d_j} \hat{\pi}_j^i\big(\dot{\pi}_{D_j}^i(\mathbf{pa}'_{D_j}) \mid \mathbf{pa}'_{D_j}\big) \right)$$

$$= \hat{\pi}_j^i(d_j \mid \mathbf{pa}_{D_j}) \cdot \prod_{\mathbf{pa}'_{D_j} \in dom(\mathbf{Pa}_{D_j}) \backslash \{\mathbf{pa}_{D_j}\}} 1$$

$$= \hat{\pi}_j^i(d_j \mid \mathbf{pa}_{D_j}).$$

Thus, if $\hat{\boldsymbol{\pi}}^i$ is a beneficial deviation against $\boldsymbol{\pi}^{-i}$, $\hat{\mu}^i$ is also a beneficial deviation against $\boldsymbol{\pi}^{-i}$. Further, by again appealing to Lemma 2 we can see that:

$$\sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{(\hat{\mu}^i, \boldsymbol{\pi}^{-i})}[U] = \sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{(\hat{\mu}^i, \mu^{-i})}[U],$$

and so $\hat{\mu}^i$ is also a beneficial deviation against μ^{-i}. In which case, μ cannot be an NE, giving us our contradiction. $\qquad\square$

Proposition 7. *If an agent i in a MAID $\mathcal{M}$ has perfect recall, then they also have sufficient recall. However, if an agent has sufficient recall, then they do not always have perfect recall.*

Proof. For the first direction, we recognise that if agent i has perfect recall in $\mathcal{M}$, then this means there exists a unique topological ordering of variables $\boldsymbol{D}^i$ in $\mathcal{M}$; i.e., $D_1 \prec \cdots \prec D_m$. Furthermore, for any $j < k$ we have that $\mathbf{Fa}_{D_j} \subseteq \mathbf{Pa}_{D_k}$. This implies that:

$$\Pi_{D_j} \perp_{\mathsf{m}\perp\mathcal{G}} \boldsymbol{U}^i \cap \mathbf{Desc}_{D_k} \mid D_k, \mathbf{Pa}_{D_k},$$

and so Π_{D_j} is not s-reachable from Π_{D_k}. Therefore, there is no edge $\Pi_{D_j} \to \Pi_{D_k}$ in the s-relevance graph $\mathsf{r}_s\mathcal{G}$. This implies that $\mathsf{r}_s\mathcal{G}$, when restricted to just agent i's decision rule variables, must be acyclic (i.e., there can *only* exist edges $\Pi_{D_j} \to \Pi_{D_k}$ when $j > k$), as required. For the second direction, we provide an example in Appendix B.2 of a MAID where an agent has sufficient but imperfect recall. $\qquad\square$

Proposition 8. *Any MAID $\mathcal{M}$ with sufficient recall has at least one SPE in behavioural policies.*

Proof. Let $\mathcal{G}_1 \prec \cdots \prec \mathcal{G}_m$ be an ordering of MAID $\mathcal{M}$'s s-subdiagrams such that $\mathcal{G}_j \prec \mathcal{G}_k$ implies that $\mathcal{G}_k$ is *not* an s-subdiagram of $\mathcal{G}_j$. If all agents have sufficient recall, then $\mathcal{G}_1$ contains at most one decision variable for each agent, and for each s-subdiagram $\mathcal{G}_j$ where $1 \le j < m$, $\mathcal{G}_{j+1}$ contains *at most* one additional decision variable for each agent. The newly added decision nodes in each subsequent s-subdiagram for such an ordering correspond exactly with a topological ordering of the strong connected components in $\mathcal{M}$'s s-relevance graph. This means that we can employ the same method as in a similar proof of Theorem 6.2 in K&M [51] to show that if we use a backwards induction procedure to optimise the decision rules of each subgame in this order, then the resulting policy profile is guaranteed to be an NE in every feasible s-subgame of $\mathcal{M}$, and hence an SPE. $\qquad\square$

Proposition 9. *Suppose that a MAID $\mathcal{M}$ has sufficient recall. Then, the set of SPEs of $\mathcal{M}$ is equal to the set of rational outcomes $\mathcal{R}^{SP}(\mathsf{m}\mathcal{M})$.*

Proof. As in the previous proof, let $\mathcal{G}_1 \prec \cdots \prec \mathcal{G}_m$ be an ordering of $\mathcal{M}$'s s-subdiagrams such that $\mathcal{G}_j \prec \mathcal{G}_k$ implies that $\mathcal{G}_k$ is *not* an s-subdiagram of $\mathcal{G}_j$. If all agents have sufficient recall, then $\mathcal{G}_1$ contains at most one decision variable for each agent, and for each s-subdiagram $\mathcal{G}_j$ where $1 \le j < m$, $\mathcal{G}_{j+1}$ contains *at most* one additional decision variable for each agent. The fact that $\mathcal{G}_1$ contains at most one decision variable for each agent means that $\boldsymbol{\pi}_1$ is an NE in any (feasible) s-subgame $\mathcal{M}_1$ over $\mathcal{G}_1$ just in case:

$$\pi_D \in \underset{\hat{\pi}_D \in dom(\Pi_D)}{\arg\max} \sum_{U \in \boldsymbol{U}^i \cap \boldsymbol{V}_1} \mathbb{E}_{(\hat{\pi}_D, \boldsymbol{\pi}_{1,-D})}[U],$$

for each $\mathcal{M}_1$, i.e., just in case $\pi_D \in r_D^{\mathrm{SP}}(\mathbf{pa}_{\Pi_D})$. This is akin to the result that the $\mathcal{R}^{\mathrm{BR}}$-rational outcomes of a game where each agent has only one decision variable are simply the NEs of the game.

Similarly, if we reason analogously to the backwards induction procedure over s-subgames as described in the proof of Proposition 8 and assume that $\boldsymbol{\pi}_j$ forms an NE in every feasible s-subgame of every $\mathcal{M}_j$, then by fixing $\boldsymbol{\pi}_j$ in each (feasible) s-subgame $\mathcal{M}_{j+1}$ we induce a new

MAID where each agent has at most one decision variable. A partial policy profile $\boldsymbol{\pi}'_{j+1}$ is an NE in this game if and only if:

$$\pi_D \in \underset{\hat{\pi}_D \in dom(\Pi_D)}{\arg\max} \sum_{U \in \boldsymbol{U}^i \cap \boldsymbol{V}_{j+1}} \mathbb{E}_{(\hat{\pi}_D, \boldsymbol{\pi}'_{j,-D})}[U],$$

for each $D \in \boldsymbol{D}_{j+1} \setminus \bigcup_{1 \leq k \leq j} \boldsymbol{D}_k$. Moreover, by our inductive hypothesis the resulting policy profile $\boldsymbol{\pi}_{j+1} = (\boldsymbol{\pi}'_{j+1}, \boldsymbol{\pi}_j)$ forms an NE (and in fact an SPE) in each (feasible) $\mathcal{M}_{j+1}$, and so we have that each such $\pi_D \in r_D^{\mathrm{SP}}(\mathbf{pa}_{\Pi_D})$. At the conclusion of this line of argument we see that any (full) policy profile $\boldsymbol{\pi} = \boldsymbol{\pi}_m$ is an NE in every feasible s-subgame of $\mathcal{M}$, when restricted to that subgame, just in case each $\pi_D \in r_D^{\mathrm{SP}}(\mathbf{pa}_{\Pi_D})$, i.e., just in case $\boldsymbol{\pi} \in \mathcal{R}^{\mathrm{SP}}(\mathsf{m}\mathcal{M})$. $\qquad \square$

Section 6

Lemma 3. *Let f be a natural mapping between $\mathcal{E}$ and $\mathcal{M}$. Then σ is an NE in $\mathcal{E}$ if and only if every $\boldsymbol{\pi} \in f(\sigma)$ is an NE in $\mathcal{M}$.*

Proof. This result follows straightforwardly from Definition 30. Suppose that σ is an NE. Then if f is a natural mapping, we have that $\sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{\boldsymbol{\pi}}[U] = \mathbb{E}_{\sigma}\big[U(\rho[\boldsymbol{L}])[i]\big]$ for every $\boldsymbol{\pi} \in f(\sigma)$ and every agent i. If $\hat{\boldsymbol{\pi}}^i$ is a profitable deviation for player i in $\mathcal{M}$ then there must exist some $f^i(\hat{\sigma}^i) \ni \boldsymbol{\pi}^i$ such that $\hat{\sigma}^i$ is a profitable deviation for player i in $\mathcal{E}$. I.e., we must have:

$$\sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{\boldsymbol{\pi}}[U] < \sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{(\boldsymbol{\pi}^{-i}, \hat{\boldsymbol{\pi}}^i)}[U] = \mathbb{E}_{(\sigma^{-i}, \hat{\sigma}^i)}\big[U(\rho[\boldsymbol{L}])[i]\big] > \mathbb{E}_{\sigma}\big[U(\rho[\boldsymbol{L}])[i]\big],$$

which contradicts our assumption that σ is an NE. The same argument also applies in the opposite direction, thus concluding the proof. $\qquad \square$

Lemma 4. *If $\mathcal{E} \in \mathtt{maid2efg}(\mathcal{M})$ or $\mathcal{M} = \mathtt{efg2maid}(\mathcal{E})$, then $\mathcal{E}$ and $\mathcal{M}$ are equivalent.*

Proof. In what follows we make the trivial assumption that any non-leaf node V in $\mathcal{E}$ has more than one child (otherwise we could simply remove such nodes), and that if V is a chance node then all of its children are reached with positive probability (otherwise we could delete the subtrees rooted at such children). The proof follows directly by construction using the procedures $\mathtt{maid2efg}$ and $\mathtt{efg2maid}$ respectively.

To see this, first suppose we have a MAID, $\mathcal{M}$, and $\mathcal{E}$ is an EFG resulting from $\mathtt{maid2efg}(\mathcal{M})$. A decision rule π_D defines a probability distribution over $dom(D)$ conditional on each decision context, $\mathbf{pa}_D$. Following the $\mathtt{maid2efg}$ procedure, each feasible decision context is an instantiation of a set of variables which corresponds to one information set I_j^i where $S_V = D$ for all $V \in I_j^i$, and for each $d \in dom(D)$ there exists precisely one node $V' \in \mathbf{Ch}_V$ such that $\lambda(V, V') = d$. In particular, if $\mathbf{pa}_D$ is infeasible then the probability of reaching I_j^i is zero and it would be removed from $\mathcal{E}$ (as per our assumption above).

For a policy profile $\boldsymbol{\pi}$, let us define σ such that $\sigma_j^i(d) := \pi_D(d \mid \mathbf{pa}_D)$ for each feasible decision context $\mathbf{pa}_D$. Note that for a given $\mathbf{pa}_D$ then this construction results in a one-to-one correspondence between $\sigma_j^i(D), \pi_D(D \mid \mathbf{pa}_D) \in \Delta(dom(D))$, and that for any two $\boldsymbol{\pi} \sim \boldsymbol{\pi}'$ (i.e., $\boldsymbol{\pi}$ and $\boldsymbol{\pi}'$ differ only on infeasible decision contexts) then the same strategy σ will be constructed. This construction therefore results in a bijection $f : \Sigma \to dom(\Pi)/\sim$. By reasoning analogously about the chance and utility variables, and observing that the difference between chosen actions in infeasible decision contexts has no bearing on the expected utility of each agent in $\mathcal{M}$ (because they all occur with probability zero) it follows that the expected utility for each agent i, $\mathbb{E}_{\sigma}\big[U(\rho[\boldsymbol{L}])[i]\big] = \sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{\boldsymbol{\pi}}[U]$ for any σ such that $\boldsymbol{\pi} \in f(\sigma)$, hence f is a natural mapping and $\mathcal{M}$ and $\mathcal{E}$ are equivalent.

For the second part of the proof, note first that the deterministic nature of the `efg2maid` procedure guarantees uniqueness of the resulting MAID, and so suppose we have some $\mathcal{M} = $ `efg2maid`$(\mathcal{E})$. In our construction above, the incoming edges for each decision variable $D \in \boldsymbol{D}^i$ are precisely those that originate from the variables whose values agent i can determine when in the corresponding information set(s) I^i_j. Hence the strategy σ^i_j, which assigns a probability distribution over the set of available decisions A^i_j at node V^i_j, is determined as a function of $\mathbf{pa}_D$, where $\mathbf{pa}_D$ corresponds to a particular information set I^i_j from which D was created.

Thus, given a strategy profile σ in $\mathcal{E}$ we fix the corresponding policy profiles $\boldsymbol{\pi}$ in $\mathcal{M}$ to have $\pi_D(d \mid \mathbf{pa}_D) \coloneqq \sigma^i_j(d)$ when $\mathbf{pa}_D$ is feasible, and let $\pi_D(d \mid \mathbf{pa}_D)$ vary otherwise. As before, we therefore have a bijection $f : \Sigma \to dom(\boldsymbol{\Pi})/\sim$. The same form of correspondence (i.e., ignoring the infeasible settings of the parents) can analogously be seen to hold between the distributions P of $\mathcal{E}$ and the CPDs for $\boldsymbol{X}$ and $\boldsymbol{U}$. Thus, for any $\boldsymbol{\pi} \in f(\sigma)$, we have that $\sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{\boldsymbol{\pi}}[U] = \mathbb{E}_\sigma\big[U(\rho[\boldsymbol{L}])[i]\big]$ for each agent $i \in N$, and so f is a natural mapping, as required. $\qquad\square$

Definition 31. A strategy profile σ is a **Nash equilibrium (NE)** in an EFG if, for every agent $i \in N$, $\sigma^i \in \arg\max_{\hat{\sigma}^i \in \Sigma^i} \mathbb{E}_{(\hat{\sigma}^i, \sigma^{-i})}\big[U(\rho[\boldsymbol{L}])[i]\big]$.

Corollary 1. *If $\mathcal{E} \in$ `maid2efg`$(\mathcal{M})$ or $\mathcal{M} = $ `efg2maid`$(\mathcal{E})$, then there is a natural mapping f between $\mathcal{E}$ and $\mathcal{M}$ such that σ is an NE in $\mathcal{E}$ if and only if every $\boldsymbol{\pi} \in f(\sigma)$ is an NE in $\mathcal{M}$.*

Proof. By Lemma 4 we have that $\mathcal{E}$ and $\mathcal{M}$ are equivalent. Thus, let $f : \Sigma \to dom(\boldsymbol{\Pi})/\sim$ be a natural mapping between $\mathcal{E}$ and $\mathcal{M}$ as defined in Definition 30. By applying Lemma 3, we have that σ is an NE in $\mathcal{E}$ if and only if every $\boldsymbol{\pi} \in f(\sigma)$ is an NE in $\mathcal{M}$. $\qquad\square$

Proposition 10. *If $\mathcal{E} \in$ `maid2efg`$(\mathcal{M})$ or $\mathcal{M} = $ `efg2maid`$(\mathcal{E})$, then there is a natural mapping f between $\mathcal{E}$ and $\mathcal{M}$ such that, for every subgame $\mathcal{E}'$ in $\mathcal{E}$ there is an s-subgame $\mathcal{M}'$ in $\mathcal{M}$ that is equivalent (modulo a constant difference between the utilities for each agent under any policy in $\mathcal{M}'$) to $\mathcal{E}'$ under the natural mapping f restricted to the strategies of $\mathcal{E}'$.*

Proof. We begin by proving existence. Recall that a subgame $\mathcal{E}'$ in an EFG $\mathcal{E}$ is a subtree that is closed under information sets and descendants. Let $\boldsymbol{V}' \subseteq \boldsymbol{V}$ be the set of variables in $\mathcal{M}$ corresponding to the intervention sets overlapping with or contained in $\mathcal{E}'$. We first show that $\boldsymbol{V}'$ forms an s-subdiagram, i.e., that $\boldsymbol{V}'$ contains every variable Z in $\mathcal{M}$ whose mechanism node M_Z is s-reachable from Π_D for some $D \in \boldsymbol{V}'$, and for every $X, Y \in \boldsymbol{V}'$, $\boldsymbol{V}'$ contains every variable that lies on a directed path $X \dashrightarrow Y$ in $\mathcal{G}$.

Beginning with the first condition, it suffices to show that there exists no variable $Z \in \boldsymbol{Z} = \boldsymbol{V} \setminus \boldsymbol{V}'$ such that M_Z is s-reachable from the decision rule Π_D of some variable $D \in \boldsymbol{D}^i \cap \boldsymbol{V}'$ for $i \in N'$. Recall that this means we would have $\mathsf{M}_Z \not\perp_{\mathsf{m}\perp\mathcal{G}} \boldsymbol{U}^i \cap \mathbf{Desc}_D \mid \mathbf{Fa}_D$, where $\mathsf{m}\perp\mathcal{G}$ is the mechanised graph with no edges between mechanism variables. Any path supporting such a dependency must have one of the following two forms:

- $\mathsf{M}_Z \to Z \to \cdots \boldsymbol{U}^i \cap \mathbf{Desc}_D$. In this case, as $Z \notin \boldsymbol{V}'$ then, by either construction `efg2maid` or `maid2efg`, any node in $\mathcal{E}$ corresponding to Z must lie outside $\mathcal{E}'$. Further, as $\mathcal{E}'$ is closed under information sets, then the value of Z must be observed by any decision node in $\mathcal{E}'$ corresponding to D, and thus there is an information link $Z \to D$ in $\mathcal{M}$ and so $Z \in \mathbf{Fa}_D$. Hence, by conditioning on $\mathbf{Fa}_D$ we block the path above, meaning there is no dependency.

- $\mathsf{M}_Z \to Z \leftarrow \cdots \boldsymbol{U}^i \cap \mathbf{Desc}_D$. In this case, let W be the first variable in the path (from left to right) that is a fork variable, i.e., we have $\cdots \leftarrow W \to \cdots$. Such a variable must exist because we assume that utility variables do not have children. As before, notice that

by either construction `efg2maid` or `maid2efg` we have that $Z \notin \boldsymbol{V'}$ must lie outside $\mathcal{E'}$, and as there is a directed path $W \dashrightarrow Z$ in $\mathcal{M}$ then W must also lie outside $\mathcal{E'}$. But then the fact that $\mathcal{E'}$ is closed under information sets means that W is observed by D and so conditioning on $\mathbf{Fa}_D$ blocks the path above, meaning there is no dependency.

We next consider the second condition. Suppose that $X, Y \in \boldsymbol{V'}$ and there is a directed path $X \dashrightarrow Z \dashrightarrow Y$ in $\mathcal{M}$. Then if $\mathcal{E} \in \mathtt{maid2efg}(\mathcal{M})$ or $\mathcal{M} = \mathtt{efg2maid}(\mathcal{E})$, any topological ordering $\prec$ over the variables in $\mathcal{M}$ must have $X \prec Z \prec Y$ and hence if there are nodes in $\mathcal{E'}$ corresponding to both X and Y then as $\mathcal{E'}$ is closed under descendants we must have a node corresponding to Z in $\mathcal{E'}$. This means that the intervention set containing Z overlaps with $\mathcal{E'}$ and so $Z \in \boldsymbol{V'}$ by the definition of $\boldsymbol{V'}$.

For the second part of the existence proof, we show that there is a setting $\boldsymbol{z}$ of $\boldsymbol{Z}$ which, when combined with $\boldsymbol{V'}$, leads to an s-subgame $\mathcal{M'}$ of $\mathcal{M}$ that is equivalent to $\mathcal{E'}$. First, note that any node passed through on the path $\rho_{R'}$ from the root R of $\mathcal{E}$ to the root R' of $\mathcal{E'}$ must correspond to a variable in $\boldsymbol{Z}$. Let $\boldsymbol{z}$ be any setting of $\boldsymbol{Z}$ that is consistent with the path from R to R', and let $\mathcal{M'}$ be the resulting s-subgame that is obtained by combining $\boldsymbol{V'}$ with $\boldsymbol{z}$. Observe that by the argument above, the decision variables in $\boldsymbol{D'} \subseteq \boldsymbol{V'}$ are precisely those corresponding to the information sets in $\mathcal{E'}$, and moreover that any decision context of any $D \in \boldsymbol{D'}$ that does not correspond to some information set in $\mathcal{E'}$ is, by definition, not feasible.

Thus, the natural mapping f between $\mathcal{E}$ and $\mathcal{M}$, when restricted to $\mathcal{E'}$ and $\mathcal{M'}$, leads to a bijection between the strategies in $\mathcal{E'}$ and the quotient set of behavioural policy profiles in $\mathcal{M'}$ by $\sim$ (where two policies are in the same equivalence class if and only if they differ only on those decision contexts that are infeasible), as per the arguments in the proof of Lemma 4. Furthermore, for any equivalent strategy $\boldsymbol{\pi} \in f(\sigma)$, it follows directly by construction using `maid2efg` or `efg2maid` that if $\mathbb{E}'_\sigma\big[U(\rho[\boldsymbol{L}])[i]\big]$ and $\sum_{U' \in \boldsymbol{U}^i \cap \boldsymbol{V'}} \mathbb{E}'_{\boldsymbol{\pi}}[U']$ denote the expected utility of agent i when σ and $\boldsymbol{\pi}$ are restricted to $\mathcal{E'}$ and $\mathcal{M'}$ respectively, then:

$$
\begin{aligned}
\mathbb{E}'_\sigma\big[U(\rho[\boldsymbol{L'}])[i]\big] &:= \sum_\rho P'^\sigma(\rho) U(\rho[\boldsymbol{L}])[i] \\
&= \sum_\rho P^\sigma(\rho \mid \rho_{R'}) U(\rho[\boldsymbol{L}])[i] \\
&= \sum_{U \in \boldsymbol{U}^i} \sum_{u \in dom(U)} u \cdot \Pr^{\boldsymbol{\pi}}(u \mid \boldsymbol{z}; \theta) \\
&= \sum_{U \in \boldsymbol{U}^i \cap \boldsymbol{Z}} u + \sum_{U' \in \boldsymbol{U}^i \cap \boldsymbol{V'}} \sum_{u' \in dom(U')} u' \cdot \Pr^{\boldsymbol{\pi}}(u' \mid \boldsymbol{z}; \theta) \\
&= \sum_{U \in \boldsymbol{U}^i \cap \boldsymbol{Z}} u + \sum_{U' \in \boldsymbol{U}^i \cap \boldsymbol{V'}} \sum_{u' \in dom(U')} u' \cdot \Pr^{\boldsymbol{\pi}}(u'; \theta') =: \sum_{U \in \boldsymbol{U}^i \cap \boldsymbol{Z}} u + \sum_{U' \in \boldsymbol{U}^i \cap \boldsymbol{V'}} \mathbb{E}'_{\boldsymbol{\pi}}[U']
\end{aligned}
$$

for every agent $i \in N'$, where $P^\sigma(\rho \mid \rho_{R'})$ is the distribution over paths ρ in $\mathcal{E}$ conditional on ρ containing $\rho_{R'}$ as a sub-path. Note that because some utility variables $U \in \boldsymbol{U}^i$ may not occur in $\boldsymbol{V'}$ then we must add their value u according to $\boldsymbol{z}$ to the payoff for each agent in order to equate the expected utilities for each agent in both games. Given $\boldsymbol{z}$, however, this difference between the utilities for each agent is constant under any policy in $\mathcal{M'}$, and so has no effect on the optimality of policies in $\mathcal{M'}$. Modulo this small difference, f forms a natural mapping between $\mathcal{E'}$ and $\mathcal{M'}$, as required. $\qquad\square$

Definition 32. A strategy profile σ in an EFG $\mathcal{E}$ is a **subgame perfect equilibrium (SPE)** if σ is an NE in every subgame of $\mathcal{E}$, when restricted to that subgame.

Corollary 2. *If $\mathcal{E} \in \mathtt{maid2efg}(\mathcal{M})$ or $\mathcal{M} = \mathtt{efg2maid}(\mathcal{E})$, then there is a natural mapping f between $\mathcal{E}$ and $\mathcal{M}$ such that if every $\boldsymbol{\pi} \in f(\sigma)$ is an SPE in $\mathcal{M}$, then σ is an SPE in $\mathcal{E}$.*

Proof. The corollary can be seen to follow immediately from combining the definitions of SPEs in EFGs and MAIDs with Proposition 10 and Lemma 3. Concretely, if $\mathcal{E} \in \mathtt{maid2efg}(\mathcal{M})$ or $\mathcal{M} = \mathtt{efg2maid}(\mathcal{E})$, and if any $\boldsymbol{\pi} \in f(\sigma)$ is an SPE in $\mathcal{M}$ (for some natural mapping f between $\mathcal{E}$ and $\mathcal{M}$ satisfying Proposition 10) then $\boldsymbol{\pi}$ is an NE in any s-subgame of $\mathcal{M}$, and as every subgame of $\mathcal{E}$ is equivalent to an s-subgame of $\mathcal{M}$ under f (modulo a constant difference between the utilities for each agent under any policy in $\mathcal{M}'$, which does not affect whether a given policy profile in $\mathcal{M}'$ is an NE or not), then we must have that σ is an NE in each subgame of $\mathcal{E}$ (by Lemma 3). $\qquad\square$

Definition 33. A perturbation vector η_k contains perturbations $\epsilon_a^{i,j} \in (0,1)$ with $\sum_{a \in A_j^i} \epsilon_a^{i,j} \leq 1$ for every information set I_j^i such that in a perturbed game $\mathcal{E}(\eta_k)$, each strategy is forced to have $\sigma_j^i(a) \geq \epsilon_a^{i,j}$. A strategy profile σ is a **trembling hand perfect equilibrium (THPE)** in an EFG $\mathcal{E}$ if there is a sequence of perturbation vectors $\{\eta_k\}_{k \in \mathbb{N}}$ such that $\lim_{k \to \infty} \|\eta_k\|_\infty = 0$ and for each perturbed EFG $\mathcal{E}(\eta_k)$ there is an NE σ_k such that $\lim_{k \to \infty} \sigma_k = \sigma$.

Proposition 11. *If $\mathcal{E} \in \mathtt{maid2efg}(\mathcal{M})$ or $\mathcal{M} = \mathtt{efg2maid}(\mathcal{E})$, then there is a natural mapping f between $\mathcal{E}$ and $\mathcal{M}$ such that σ is a THPE in $\mathcal{E}$ if and only if every $\boldsymbol{\pi} \in f(\sigma)$ is a THPE in $\mathcal{M}$.*

Proof. Given Lemma 4 and Lemma 3, it suffices to show that the perturbed MAID resulting from $\mathtt{efg2maid}(\mathcal{E}(\eta_k))$ is equivalent to $\mathcal{E}(\eta_k)$, and similarly that any perturbed EFG resulting from $\mathtt{maid2efg}(\mathcal{M}(\zeta_k))$ is equivalent to $\mathcal{M}(\zeta_k)$. For the first part, given η_k we define the entries of ζ_k for each decision d and decision context $\mathbf{pa}_D$ as:

$$
\epsilon_d^{\mathbf{pa}_D} := \begin{cases} \epsilon_a^{i,j} & \text{if } d = a \text{ and } \mathbf{pa}_D \text{ is feasible} \\ \min_{i,j,a} \epsilon_a^{i,j} & \text{otherwise,} \end{cases}
$$

where $\mathbf{pa}_D$ corresponds to the information set I_j^i as per our $\mathtt{efg2maid}$ construction. Clearly for any $\boldsymbol{\pi} \in f(\sigma)$ where f is a natural mapping between $\mathcal{E}$ and $\mathcal{M}$ then $\sigma_j^i(a) \geq \epsilon_a^{i,j}$ if and only if $\pi_D(d \mid \mathbf{pa}_D) \geq \epsilon_d^{\mathbf{pa}_D}$. Based on this, it can easily be seen that any such f is also a natural mapping between the perturbed games $\mathcal{E}(\eta_k)$ and $\mathcal{M}(\zeta_k)$, and hence they are equivalent. For the second part, given ζ_k we construct η_k such that $\epsilon_a^{i,j} := \epsilon_d^{\mathbf{pa}_D}$ if $a = d$ and $\mathbf{pa}_D$ is feasible. A similar argument to the above shows that any natural mapping f between $\mathcal{M}$ and $\mathcal{E}$ induces a natural mapping between the perturbed games $\mathcal{M}(\zeta_k)$ and $\mathcal{E}(\eta_k)$, as required. $\qquad\square$

B Further Examples

B.1 Counterfactuals Using the Closest Possible World Principle

When computing counterfactuals in SCGs under the 'closest possible world' principle, our basic desideratum is to consider those counterfactual rational outcomes $\boldsymbol{\pi}' \in \mathcal{R}(\mathsf{m}\mathcal{M}_\mathcal{I})$ that are consistent with the decision rules used in some actual rational outcome $\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z})$ whenever those decision rules are *not* affected by $\mathcal{I}$. In other words, we wish to keep the set of invariant decision rules $\boldsymbol{\Pi}(\mathcal{I})$ as large as possible while propagating changes due to $\mathcal{I}$.

As a first attempt, we might set $\boldsymbol{\Pi}(\mathcal{I})$ to $\boldsymbol{\Pi} \setminus (\boldsymbol{Y} \cup \mathbf{Desc_Y})$. It can easily be seen, however, that this choice is flawed. Consider an $\mathcal{R}$-relevance graph with two variables $\Pi_D, \Pi_{D'}$, with $\mathbf{Pa}_{\Pi_{D'}} = \{\Pi_D\}$, to which we apply an intervention $\mathrm{do}(\Pi_D = \pi_D)$. Then it is perfectly possible that $r_{D'}(\pi_D) = \bigcup_{\hat{\pi}_D \in r_D()} r_{D'}(\hat{\pi}_D)$ – i.e., the set of rational responses for decision rule $\Pi_{D'}'$ does not change – and in which case there is no sense in which $\Pi_{D'}'$ is affected by $\mathcal{I}$, even though $\Pi_{D'}' \in \{\Pi_D\} \cup \mathbf{Desc}_{\Pi_D}$.

Instead, we can compute $\mathbf{\Pi}(\mathcal{I})$ by propagating the effects of $\mathcal{I}$ through the *maximal strongly connected components* (MSCCs) of the $\mathcal{R}$-relevance graph when restricted to decision rule variables.[28] Let $\mathsf{c}_\mathcal{R}\mathcal{G}$ denote the condensation of the $\mathcal{R}$-relevance graph $\mathsf{r}_\mathcal{R}\mathcal{G}$ when restricted to the decision rule variables – called the 'component graph' by K&M [51] – with topological ordering $C_1 \prec \cdots \prec C_m$ over the vertices in $\mathsf{c}_\mathcal{R}\mathcal{G}$ (where each $C_j \subseteq \mathbf{\Pi}$ is an MSCC of $\mathsf{r}_\mathcal{R}\mathcal{G}$ restricted to the decision rule variables), and let us write:

$$r_{C_j}(\mathbf{pa}_{C_j}) := \{c_j : \pi_D \in r_D(\mathbf{pa}_{\Pi_D}) \ \forall \ \Pi_D \in C_j\},$$

to denote the rational responses $c_j = (\pi_D)_{\Pi_D \in C_j}$. Then we may compute $\mathbf{\Pi}(\mathcal{I})$ using Algorithm 1.

Algorithm 1

    Input: $\mathsf{m}\mathcal{M}, \boldsymbol{z}, \mathcal{I}$
    Output: $\mathbf{\Pi}(\mathcal{I})$
1: $\Pr(\boldsymbol{\theta}) \leftarrow \Pr(\boldsymbol{\theta}_\mathcal{I})$
2: $\mathbf{\Pi}(\mathcal{I}) \leftarrow \mathbf{\Pi} \setminus (\boldsymbol{Y} \cup \mathbf{Desc_Y})$
3: form $\mathsf{c}_\mathcal{R}\mathcal{G}$ from $\mathsf{r}_\mathcal{R}\mathcal{G}$ with topological ordering $C_1 \prec \cdots \prec C_m$
4: **for** $j = 1, \ldots, m$ **do**
5:     **if** $C_j \subseteq \mathbf{\Pi}(\mathcal{I})$ **then** continue
6:     **else**
7:         $\texttt{actual} \leftarrow \mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z})$
8:         $\texttt{counterfactual} \leftarrow \bigcup_{\boldsymbol{\pi} \in \texttt{actual}} \{\boldsymbol{\pi}' \in \mathcal{R}(\mathsf{m}\mathcal{M}_\mathcal{I}) : \boldsymbol{\pi}(\mathcal{I}) = \boldsymbol{\pi}'(\mathcal{I})\}$
9:         $\boldsymbol{\Theta}_j \leftarrow \mathbf{Pa}_{C_j} \setminus \mathbf{\Pi}$
10:        $\mathbf{\Pi}_j \leftarrow \mathbf{Pa}_{C_j} \cap \mathbf{\Pi}$
11:        **if** $\bigcup_{\boldsymbol{\pi} \in \texttt{actual}} r_{C_j}(\boldsymbol{\theta}_j, \boldsymbol{\pi}_j) = \bigcup_{\boldsymbol{\pi}' \in \texttt{counterfactual}} r_{C_j}(\boldsymbol{\theta}_j, \boldsymbol{\pi}'_j)$ **then** $\mathbf{\Pi}(\mathcal{I}) \leftarrow \mathbf{\Pi}(\mathcal{I}) \cup C_j$
12: **return** $\mathbf{\Pi}(\mathcal{I})$

Algorithm 1 functions by incrementally expanding the set $\mathbf{\Pi}(\mathcal{I})$ – which is at first initialised to $\mathbf{\Pi} \setminus (\boldsymbol{Y} \cup \mathbf{Desc_Y})$ – by computing the rational responses for each MSCC to both the actual and counterfactual rational outcomes, adding the variables in the component to $\mathbf{\Pi}(\mathcal{I})$ if and only if the set of such responses remains the same. It can therefore be seen as generalising interventions to models containing both cycles and non-determinism, while maintaining a maximally justifiable similarity to an observed outcome. In other words, we maintain our updated beliefs (due to observation $\boldsymbol{z}$) about the values of a decision rule variable Π_D if and only if the set of all values π_D takes in any rational outcome $\boldsymbol{\pi}$ remains the same after the intervention $\mathcal{I}$ is performed, i.e., if and only if $\Pi_D \in \mathbf{\Pi}(\mathcal{I})$. We may then use this set $\mathbf{\Pi}(\mathcal{I})$ in Definition 23 to define counterfactual queries in games under this principle.

B.2   Non-Existence Results

No Nash Equilibrium in Behavioural Policies

Proposition 2 in Section 5.1 says that in a MAID with sufficient recall, there must exist an NE in behavioural policies. The example presented here, adapted from [97], demonstrates that if even one agent in a MAID does *not* have sufficient recall, then although there will be an NE in mixed policies, the same may not hold true for behavioural policies.

The example is shown as both a MAID $\mathcal{M}$ and an EFG $\mathcal{E}$ in Figure 11, in which agent 1 has two binary decision variables (A and B), agent 2 has one binary decision variable (D), and the

---

[28]Recall that a strongly connected component (SCC) is a subgraph containing a directed path between every pair of nodes and a maximal SCC is an SCC that is not a strict subset of any other SCC.
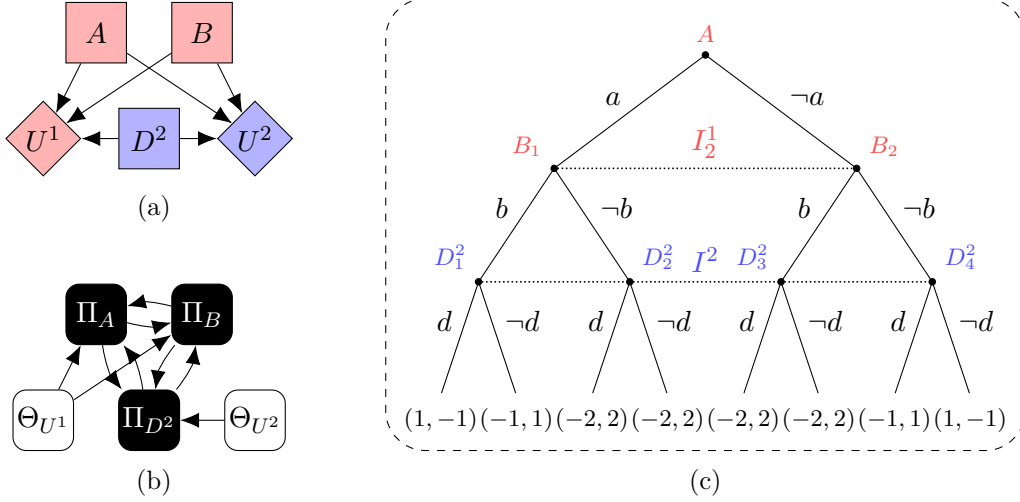
Figure 11: (a) A MAID representing the NE non-existence example and (b) its s-relevance graph. (c) An equivalent EFG representing the same game.

utilities are assigned according to the leaves in $\mathcal{E}$. For the sake of simplicity, we identify each pure policy with the decisions chosen, i.e., $dom(\dot{\mathbf{\Pi}}^1) = \{\neg a \neg b, \neg ab, a \neg b, ab\}$ and $dom(\dot{\mathbf{\Pi}}^2) = \{\neg d, d\}$. Note that agent 1 has insufficient recall as the s-relevance graph restricted to contain just Π_A and Π_B is not acyclic.

In order to show that $\mathcal{M}$ has no NE in behavioural policies, we first show that there exists no NE where agent 1 uses a pure policy. Suppose, by contradiction, that there exists some policy profile $\dot{\boldsymbol{\pi}} = (\dot{\boldsymbol{\pi}}^1, \dot{\boldsymbol{\pi}}^2)$ where $\dot{\boldsymbol{\pi}}^1$ is pure and $\dot{\boldsymbol{\pi}}$ is an NE. First, $\dot{\boldsymbol{\pi}}^1$ cannot be $\neg ab$ or $a\neg b$ because under any choice of $\dot{\boldsymbol{\pi}}^2$ agent 1 receives expected utility -2 and so could improve their expected utility by playing $\neg a \neg b$ or ab. Second, if $\dot{\boldsymbol{\pi}}^1 = \neg a \neg b$, then agent 2's best response is $\dot{\boldsymbol{\pi}}^2 = d$, and if $\dot{\boldsymbol{\pi}}^1 = ab$, then agent 2's best response is $\dot{\boldsymbol{\pi}}^2 = \neg d$. However, agent 1's best response to $\dot{\boldsymbol{\pi}}^2 = d$ is $\dot{\boldsymbol{\pi}}^1 = ab$ and agent 1's best response to $\dot{\boldsymbol{\pi}}^2 = \neg d$ is $\dot{\boldsymbol{\pi}}^1 = \neg a \neg b$. Therefore, there exists no choice of pure policy $\dot{\boldsymbol{\pi}}^1$ for any $\dot{\boldsymbol{\pi}}^2 \in \dot{\mathbf{\Pi}}^2$ such that both agents are simultaneously playing a best response. Hence, no NE in pure policies exists in $\mathcal{M}$.

We now consider the case where agent 1 is using a behavioural policy in which π_A^1 and/or π_B^1 is stochastic. Suppose, by contradiction, that there exists some policy profile $\boldsymbol{\pi} = (\boldsymbol{\pi}^1, \boldsymbol{\pi}^2)$ that is an NE in behavioural policies. We let agent 1's decision rules be parameterised by $p, q \in [0, 1]$ such that $\pi_A^1(a) = p$ and $\pi_B^1(b) = q$. First, consider the case where agent 2 plays d or $\neg d$ with probability 1, then agent 1's best response is to play the pure policy $\dot{\boldsymbol{\pi}}^1 = ab$ or $\dot{\boldsymbol{\pi}}^1 = \neg a \neg b$ with probability 1 respectively. Since we know that no NE in pure policies exists, agent 2 must instead select a stochastic decision rule $\pi_{D^2}^2$ and so agent 2 must be indifferent between d and $\neg d$. We thus obtain two constraints on p and q. On the one hand, agent 1's behavioural policy $\boldsymbol{\pi}^1$ must result in $\pi^1(\neg a, \neg b) = \pi^1(a, b)$, and hence we have:

$$(1-p)(1-q) = pq \Rightarrow p + q = 1.$$

On the other hand, agent 1 receives utility -2 if its policy $\boldsymbol{\pi}^1$ selects $\neg a$ and b or a and $\neg b$ (whatever the choice of $\boldsymbol{\pi}^2$). Therefore, we must have that $\pi^1(\neg a, b) + \pi^1(a, \neg b) < \pi^1(a, b) + \pi^1(\neg a, \neg b)$ and thus (by substituting in the result that $p + q = 1$):

$$(1-p)q + p(1-q) < pq + (1-p)(1-q) \Rightarrow (2p-1)^2 < 0.$$

This contradiction implies that $\mathcal{M}$ has no NE in behavioural policies. However, as expected by Nash's theorem [68], there does exist an NE of $\mathcal{M}$ in mixed policies (i.e., where both agents randomise over their pure policies). To find this mixed policy NE, we know from the principle

of indifference that the expected utility from all of an agent's pure policies in the support of their mixed policy must be the same. This means that we can immediately rule out the pure policies $\neg ab$ and $a \neg b$ being played with any positive probability. Assuming that agent 1 plays mixed policy μ_r^1, which we define as ab with probability r and $\neg a \neg b$ with probability $1 - r$, then $\mathbb{E}_{(\mu_r^1, d)}[U^2] = \mathbb{E}_{(\mu_r^1, \neg d)}[U^2]$ implies $r = \frac{1}{2}$. Furthermore, assuming that agent 2 plays mixed policy μ_s^2, which we define as d with probability s and $\neg d$ with probability $1 - s$, then $\mathbb{E}_{(ab, \mu_s^2)}[U^1] = \mathbb{E}_{(\neg a \neg b, \mu_s^2)}[U^1]$ implies $s = \frac{1}{2}$. Thus, the policy profile $\mu = (\mu_{\frac{1}{2}}^1, \mu_{\frac{1}{2}}^2)$ is an NE of $\mathcal{M}$ in mixed policies.

No Subgame Perfect Equilibrium

We now extend the MAID from the previous example (in Figure 11a) to construct a MAID that has no SPE. Proposition 8 in Section 5.3 says that in a MAID with sufficient recall, there must exist an SPE in behavioural policies. The example presented here demonstrates that if even one agent in a MAID does *not* have sufficient recall, then there may not exist an SPE, even if we allow mixed policies.

Figure 12a shows the graph $\mathcal{G}$ of the MAID $\mathcal{M}$ for our example and Figure 12d shows its s-relevance graph (where we can observe that agent 1 has insufficient recall). The proper s-subdiagrams – $\mathcal{G}_1$, $\mathcal{G}_2$, and $\mathcal{G}_3$ – of $\mathcal{G}$ are shown in Figures 12b, 12c, and 12e respectively. To parameterise $\mathcal{M}$, all non-utility variables are given a Boolean domain (which we interpret as integers, e.g., $a = 1$ and $\neg a = 0$) and U^1 and U^2 inherit the same parameterisation as the previous example (shown for reference via the EFG in Figure 11c). We let $U^3 = D^3$ and let $U^4 = 1$ if and only if $D^4 = B$, otherwise $U^4 = 0$. Finally, we set the CPD of X such that $X = A$ if $D^3 = 1$ and $X = 1 - A$ if $D^3 = 0$.
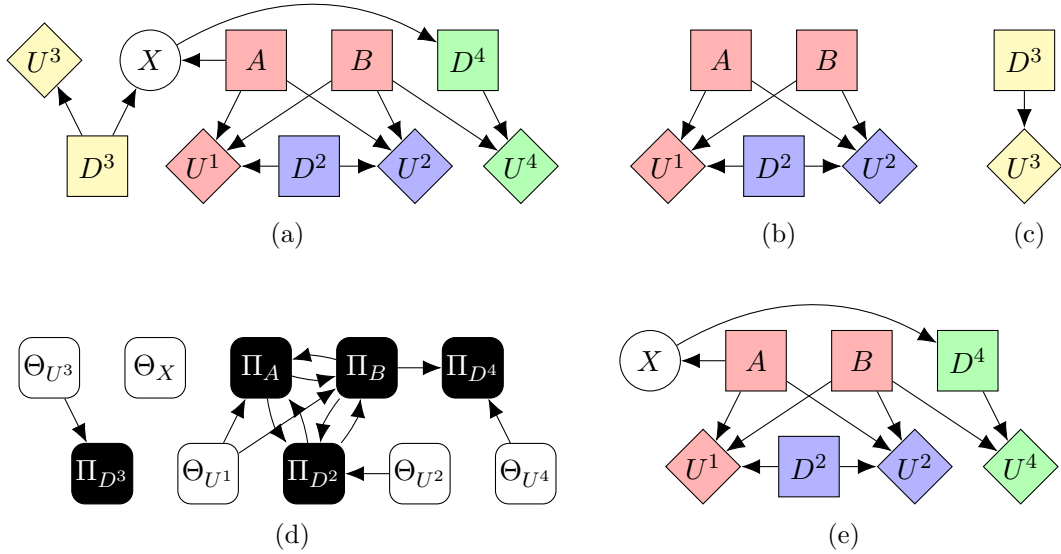

Figure 12: (a) A MAID $\mathcal{M}$ representing the SPE non-existence example. (b), (c), and (e) show the three proper s-subdiagrams of this MAID, $\mathcal{G}_1$, $\mathcal{G}_2$, and $\mathcal{G}_3$ respectively. (d) The s-relevance graph of $\mathcal{M}$.

From the previous example we know that any s-subgame defined over $\mathcal{G}_1$ has no behavioural NEs, and that the only mixed NE is given by $\mu_1 = (\mu_{\frac{1}{2}}^1, \mu_{\frac{1}{2}}^2)$ where $\mu_{\frac{1}{2}}^1$ plays the pure strategies ab and $\neg a \neg b$ each with probability $\frac{1}{2}$, and $\mu_{\frac{1}{2}}^2$ plays pure strategy d^2 with probability $\frac{1}{2}$. Thus, any mixed SPE $\mu = (\mu^1, \mu^2, \mu^3, \mu^4)$ in $\mathcal{M}$ must be such that $\mu^1 = \mu_{\frac{1}{2}}^1$ and $\mu^2 = \mu_{\frac{1}{2}}^2$.

With μ^1 and μ^2 fixed, we now show that there is no choice of μ^4 that leads to an NE in every

feasible subgame defined over $\mathcal{G}_3$. First, note that there are effectively only two such subgames, $\mathcal{M}_3$ and $\mathcal{M}'_3$ (defined by setting $D^3 = d^3$ and $D^3 = \neg d^3$ respectively), both of which are clearly feasible. In $\mathcal{M}_3$ the only policy μ^4 that forms an NE with μ^1 and μ^2 is one that assigns all probability mass to the pure policy that plays d^4 if and only if $X = 1$. However, in $\mathcal{M}'_3$, the only best response for agent 4 is to play $\neg d^4$ if and only if $X = 1$. As such, there is no policy profile that forms an NE in *every* s-subgame of $\mathcal{M}$, and hence no SPE in $\mathcal{M}$.

The reason for this result is that the use of a mixed policy by agent 1 means that variables A and B become correlated. This can be modelled graphically by the introduction of a shared parent C of A and B. Given this correlation, Π_{D^3} is now s-relevant to Π_{D^4} due to the new path:

$$\Pi_{D^3} \to D^3 \to X \leftarrow A \leftarrow C \to B \to U^4,$$

in the independent mechanised graph, which is active given $\mathbf{Fa}_{D^4} = \{X, D^4\}$.

No Equivalent Mixed and Behavioural Policies

Finally, we provide an example to demonstrate why sufficient recall is not a sufficient condition for every mixed policy to have an equivalent behavioural policy that results in the same probability distribution over game outcomes. It therefore serves as a proof for the second part of Proposition 7.
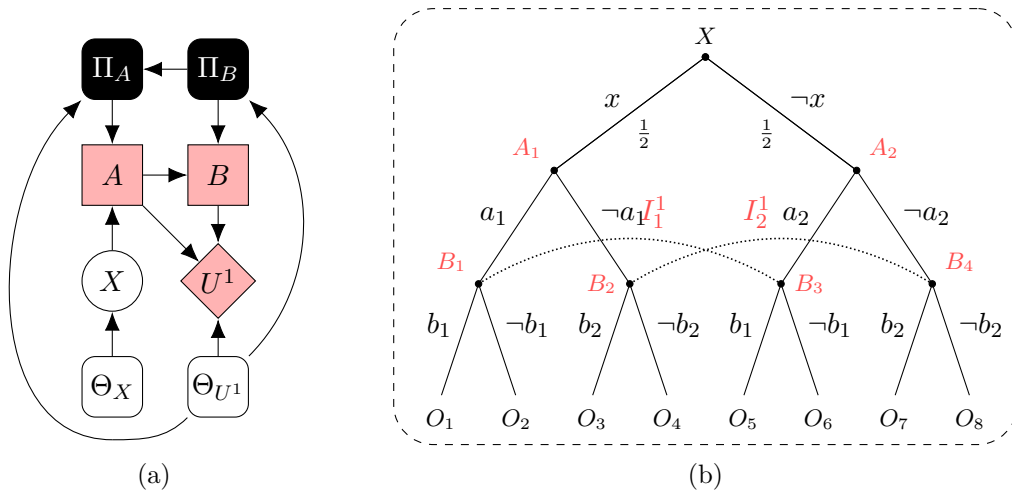

Figure 13: (a) An s-minimal mechanised MAID representing the equivalent behavioural and mixed policies non-existence example. (b) An equivalent EFG representing the same game.

Figure 13a shows the mechanised MAID for this example, $\mathcal{M}$, in which the single agent has sufficient, but imperfect recall. We assume that all variables are binary, and further that $\Pr(x) = \frac{1}{2}$. To aid our reasoning, Figure 13b shows the corresponding EFG for $\mathcal{M}$. The outcomes of the game (equivalent to a setting of all the variables in $\mathcal{M}$) are denoted by $\{O_1, \ldots, O_8\}$, e.g., O_3 represents $(x, \neg a, b, u_3^1)$. We use the same notation as the previous example to denote a pure policy. For example, $\dot{\pi}^1 = a_1 a_2 b_1 \neg b_2$ is the pure policy where the agent selects a whatever the value of X, and b if a, and $\neg b$ if $\neg a$. Because there are two decision contexts for A and B respectively (corresponding to the four information sets in the EFG), there are 2^4 pure policies. A mixed policy is then a distribution over these pure policies.

We will now show that there is no behavioural policy equivalent to the mixed policy $\mu^i = [\frac{1}{2}(a_1 \neg a_2 b_1 \neg b_2), \frac{1}{2}(a_1 a_2 \neg b_1 \neg b_2)]$. Let use begin by parameterising a general behavioural policy

as:

$$\pi_A(a \mid x) = p \qquad\qquad \pi_B(b \mid a) = r$$
$$\pi_A(a \mid \neg x) = q \qquad\qquad \pi_B(b \mid \neg a) = s$$

where $p, q, r, s \in [0, 1]$. Suppose, for a contradiction, that a behavioural policy equivalent to μ^i exists. This behavioural policy must then induce the same probability distribution over the outcomes of the game $\{O_1, \ldots O_8\}$ as μ^i, and so the following equalities must hold:

(O_1) $\frac{1}{2} \cdot p \cdot r = \frac{1}{4}$ $\qquad\qquad$ (O_5) $\frac{1}{2} \cdot q \cdot r = 0$

(O_2) $\frac{1}{2} \cdot p \cdot (1 - r) = \frac{1}{4}$ $\qquad\qquad$ (O_5) $\frac{1}{2} \cdot q \cdot (1 - r) = \frac{1}{4}$

(O_3) $\frac{1}{2} \cdot (1 - p) \cdot s = 0$ $\qquad\qquad$ (O_7) $\frac{1}{2} \cdot (1 - q) \cdot s = 0$

(O_4) $\frac{1}{2} \cdot (1 - p) \cdot (1 - s) = 0$ $\qquad\qquad$ (O_8) $\frac{1}{2} \cdot (1 - q) \cdot (1 - s) = \frac{1}{4}$

From (O_5), we must have $q = 0$ or $r = 0$, but $q = 0$ is ruled out by (O_6). Taking $r = 0$ implies that $q = \frac{1}{2}$ using (O_6), and so $s = 0$ using (O_7). If $s = 0$, then $p = 1$ using (O_4); however, this means that $r = \frac{1}{2}$ from (O_1) or (O_2) and yet we were forced to set $r = 0$. This contradiction implies that there is no behavioural policy equivalent to μ^i in $\mathcal{M}$.

B.3 Reasoning about Existing Concepts using Causal Games

Blame

The following definition is adapted from [39], and formalises the extent to which an agent is blameworthy for causing an event.

Definition 34 (39). Let $\mathcal{M}$ be an SCG, $\boldsymbol{\pi}$ a policy profile, $D \in \boldsymbol{D}^i$ a decision variable with $d, d' \in dom(D)$, and φ a Boolean combination of terms $V = v$. The **degree of blameworthiness** of d for φ relative to d' is denoted $db_S(d, d', \varphi)$, and is defined as:

$$db_S(d, d', \varphi) \coloneqq \delta_{d,d',\varphi} \cdot \frac{S - \max\left(c^i(d') - c^i(d), 0\right)}{S},$$

where:

- $\delta_{d,d',\varphi} \coloneqq \max\left(0, \Pr^{\boldsymbol{\pi}}(\varphi_d) - \Pr^{\boldsymbol{\pi}}(\varphi_{d'})\right)$ captures how much more likely it is that φ will result from decision d than from decision d';

- $c^i(d) \coloneqq \sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{\boldsymbol{\pi}}[U] - \sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{\boldsymbol{\pi}}[U_d]$ captures the cost to agent i of performing d;

- $S > \max_{d \in dom(D)} c^i(d)$ is a given measure of cost-sensitivity.

The **overall degree of blameworthiness** of d for φ is $db_S(d, \varphi) \coloneqq \max_{d' \in dom(D)} db_S(d, d', \varphi)$.

The factor $\delta_{d,d',\varphi}$ captures the extent to which the agent's decision causes the likelihood of φ to increase. The max operation ensures that no blame is possible if the likelihood *decreases*. The second factor of $db_S(d, d', \varphi)$ captures the costs of actions – the idea being that an agent is less liable to be blamed for φ if taking an alternative action would have been particularly costly. The extent to which these costs are taken into account is determined by S, where note that $\lim_{S \to \infty} db_S(d, d', \varphi) = \delta_{d,d',\varphi}$.

In Example 2, for instance, one might wish to compute $db_S(q, \neg q, B = b)$ – the degree of blame we should assign to robot one for moving quickly (as opposed to not moving quickly) with respect to breaking an item. Let us assume the policy profile in question is $\boldsymbol{\pi}^{\text{THPE}}$ as described

in Section 5.3 (in which robot one moves quickly and robot two patrols if and only if it observes robot one moving quickly). Then, we have the following:

$$
\begin{aligned}
db_S(q, \neg q, B = b) &= \delta_{q, \neg q, B = b} \cdot \frac{S - \max\left(c^1(\neg q) - c^1(q), 0\right)}{S} \\
&= \max\left(0, \Pr^{\boldsymbol{\pi}}(b_q) - \Pr^{\boldsymbol{\pi}}(b_{\neg q})\right) \cdot \frac{S - \max\left(0, \mathbb{E}_{\boldsymbol{\pi}}[U_q^1] - \mathbb{E}_{\boldsymbol{\pi}}[U_{\neg q}^1]\right)}{S} \\
&= \max\left(0, \frac{1}{3} - 0\right) \cdot \frac{S - \max(0, 2 - 2)}{S} \\
&= \frac{1}{3}.
\end{aligned}
$$

Before continuing, we note that definitions (and the definitions of intent below) were originally formalised in SCMs with a single 'action' variable A and a single utility function $\mathcal{U} : dom(\boldsymbol{V}) \to \mathbb{R}$, resulting in some small differences. Firstly, in an SCM, every variable must have a default value and actions are viewed as interventions on said variables, whereas, in SCGs, the values of the endogenous variables are typically undefined before actions are chosen. This is because SCMs model sequences of events and their causal connections in a broad sense, whereas SCGs represent a game-theoretic model to be analysed. These two framings can easily be reconciled if one posits a default value for every decision variable in an SCG, such as 'do nothing'. Relatedly, the use of utility variables in SCGs strictly generalises $\mathcal{U}$, and can be used to describe a more fine-grained structure. Finally, previous work views SCMs as representing a subjective, epistemic state possessed by the intervening agent.[29] In contrast, SCGs represent a more objective view of a sequential strategic interaction that is assumed to be common knowledge among all agents in the game. This perspective may be more appropriate when, say, assessing blame and intention from the point of view of an arbiter, system designer, or other third party.

Intent

In the same work as above, the authors also formalise the question of whether or not an agent intends to bring about an event, which we adapt below for use with SCGs.

Definition 35 (39)**.** Let $\mathcal{M}$ be an SCG, $\boldsymbol{\pi}$ a policy profile, $D \in \boldsymbol{D}^i$ a decision variable with $d \in dom(D)$ and $\text{alt}(d) \subseteq dom(D)$ be a set of alternative decisions under consideration. An agent **intends to bring about** $\boldsymbol{Y} = \boldsymbol{y}$ by performing d if and only if:

- There exists some $\boldsymbol{Z} \supseteq \boldsymbol{Y}$ such that $\sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{\boldsymbol{\pi}}[U_d] \leq \max_{d' \in \text{alt}(d)} \sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{\boldsymbol{\pi}}[U_{d', \boldsymbol{z}_d}]$, and $\boldsymbol{Z}$ is minimal with respect to this inequality;

- $\Pr^{\boldsymbol{\pi}}(\boldsymbol{y}_d) > 0$, where recall that $\Pr^{\boldsymbol{\pi}}(\boldsymbol{y}_d)$ denotes $\Pr^{\boldsymbol{\pi}}(\boldsymbol{Y}_d = \boldsymbol{y})$;

- For all $\boldsymbol{y}' \in dom(\boldsymbol{Y})$ such that $\Pr^{\boldsymbol{\pi}}(\boldsymbol{y}'_d) > 0$ we have $\sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{\boldsymbol{\pi}}[U_{\boldsymbol{y}'}] \leq \sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{\boldsymbol{\pi}}[U_{\boldsymbol{y}}]$.

Here $\boldsymbol{z}_d$ is used in a *nested counterfactual* to denote the value that $\boldsymbol{Z}$ would take under the intervention $\text{do}(D = d)$ (evaluated with respect to some marginalised setting $\mathbf{e}$ of the exogenous variables).

The notion of intent here addresses the problem of differentiating desirable, intended effects from undesirable, unintended effects. The first condition says that if the variables $\boldsymbol{Z}$ were set to the values that they would take under d, then another decision $d' \in \text{alt}(d)$ becomes at least as good, and further that $\boldsymbol{Z}$ is the minimal set of variables that the agent is intending to affect

---

[29]The philosophical motivation here is that arguably it only makes sense to judge an agent's, e.g. intent, relative to that agent's own beliefs. In our work we make a common prior assumption and assume that all agents are aware of the structure of the game, though relaxations of this assumption ought to support more subjective definitions of intent.

in this way. The second condition says that it is possible to bring about $\boldsymbol{y}$ by doing d (where it is assumed that the agent knows this). Finally, the third condition says that $\boldsymbol{y}$ is an optimal value of $\boldsymbol{Y}$ – among those possible under the intervention $\mathrm{do}(D = d)$ – for agent i.

For instance, we can see that robot two does not intend to obstruct robot one (meaning that it receives utility $U^1 = 0$) by patrolling. This is because the first property in Definition 35 fails to hold. Taking $\mathrm{alt}(p) = \{\neg p\}$, then the minimal set of variables $\boldsymbol{Z}$ such that:

$$c^2(p) \leq \max_{p' \in \mathrm{alt}(p)} \sum_{U \in \boldsymbol{U}^1} \mathbb{E}_{\boldsymbol{\pi}}[U_{p', \boldsymbol{z}_p}],$$

is $\{U^2\} \not\supseteq U^1$. In other words, given that robot one is moving quickly, $\{U^2\}$ is the minimal set of outcomes that robot two is trying to affect by performing p. If the values of U^2 were restricted to those that occur when $D^2 = p$, then there would be no incentive for robot two to patrol.

Incentives

In what follows, structural causal influence models (SCIMs) refer to SCGs with only one agent.

Definition 36 (22). A policy $\boldsymbol{\pi}$ in a single-decision SCIM $\mathcal{M}$ **responds** to a variable X if there exists $x \in dom(X)$ and $\mathbf{e} \in dom(\mathbf{E})$ such that $\mathrm{Pr}^{\boldsymbol{\pi}}(D_x \mid \mathbf{e}) \neq \mathrm{Pr}^{\boldsymbol{\pi}}(D \mid \mathbf{e})$. X has a **response incentive** (RI) if all optimal policies $\boldsymbol{\pi}^*$ respond to X, and a graph $\mathcal{G}$ admits a RI on X if there is a SCIM over $\mathcal{G}$ that has a RI on X.

Definition 37 (22). In a single-decision SCIM $\mathcal{M}$, there is an **instrumental control incentive** (ICI) on a variable X in decision context $\mathbf{pa}_D$ if for all optimal policies $\boldsymbol{\pi}^*$, there exists some $d \in dom(D)$ such that:

$$\sum_{U \in \boldsymbol{U}} \mathbb{E}_{\boldsymbol{\pi}^*}[U_{x_d} \mid \mathbf{pa}_D] \neq \sum_{U \in \boldsymbol{U}} \mathbb{E}_{\boldsymbol{\pi}^*}[U \mid \mathbf{pa}_D].$$

A graph $\mathcal{G}$ admits an ICI on X if there is a SCIM over $\mathcal{G}$ that has an ICI on X for some $\mathbf{pa}_D$.

These definitions can help us to understand what algorithmic frameworks lead to agents having undesirable incentives, and thus to build safer AI systems. One such proposal for building safe AI systems is cooperative inverse reinforcement learning (CIRL) [36], which seeks to formalise the alignment problem [9, 83] as an *assistance game* in which a human H and a robot R attempt to cooperatively perform a task in an unknown environment, but where the robot is uncertain about the human's true reward function (parameterised by the value of P^H) and so must infer it through observing the human's actions. This setup is formalised using a decentralised partially observable Markov decision process (Dec-POMDP) and is shown as a MAID in Figure 14 for a finite horizon game of three timesteps. At each timestep t, both the human and the robot perform an action A_t^H and A_t^R from their shared state S_t, after which they transition to a new state S_{t+1} and receive a reward R_{t+1}, which is only observable by the human.

If we assume that, in the example given by Figure 14, actions A_1^H, A_1^R, and A_2^H have already been taken (and thus that the MAID reduces to a single-decision ID with decision variable A_2^R), then we may apply the sound and graphical criteria derived in previous work to detect RIs and ICIs in CIRL [22]. A variable X in a graph admits an RI (with respect to decision variable D) if and only if there is a directed path $X \dashrightarrow D$ in the graph that results when all information links to D from variables Y satisfying $Y \not\perp_{\mathcal{G}} \boldsymbol{U} \cap \mathbf{Desc}_D \mid D, \mathbf{Pa}_D$ are removed.[30] As A_2^H does not satisfy this criterion, the path $P^H \to A_2^H \to A_2^R$ exists in this new graph, and hence we see that

---

[30]The observations available at D satisfying this criterion are known as *non-requisite* observations [58]. Note the similarity to *s*-reachability, defined in Proposition 1.
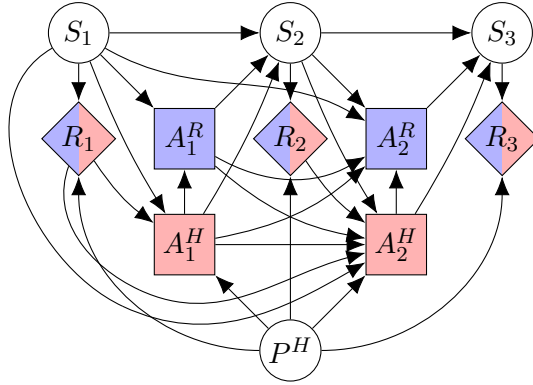
Figure 14: A MAID representing an assistance game [23], played by agents H (a human) and R (a robot); both want to maximise the same rewards, indicated by the shared utility variables.

the robot R has an RI to act according to P^H, as we would hope. A variable X admits an ICI (with respect to decision variable D) if and only if there exists a directed path $D \dashrightarrow X \dashrightarrow U$ for some utility variable U. Hence we see, due to the path $A_2^R \to S_3 \to R_3$, that the robot has an ICI to influence S_3 as, again, we would expect.

C Codebase

In this section, we briefly describe PyCID [27], our open source Python library that implements MAIDs (and their causal variants). PyCID has a range of classes, methods, and functions for handling IDs at level one and level two of the causal hierarchy. For our work in this paper, the `MAID` class is of primary interest; however, note that PyCID also has significant other functionality including functions for computing incentives in single-agent causal IDs [22] and reasoning patterns in MAIDs [76]. This makes the codebase well-suited as a testbed for future research and applications.

We begin by showing how to instantiate MAIDs for Examples 1 and 2 in PyCID. We then provide empirical results demonstrating how MAIDs can be used to compute NEs faster than in the equivalent EFGs. We refer the reader to an existing tool paper [27] and our online codebase for further details, including several tutorials.

C.1 Creating MAIDs

Listings 1 and 2 show how to instantiate a MAID for Examples 1 and 2 as instances of PyCID's `MAID` class, which inherits from pgmpy's `BayesianModel` class [2]. A MAID is initialised using a list of edges as its first argument, and then dictionaries to specify each agent's decision and utility variables. The method `draw` plots the graphs of these MAIDs (shown in Figure 15a and 15b) with chance variables as grey circles, decision variables as rectangles, utility variables as diamonds, and colouring to denote different agents (each agent is assigned a unique colour). Recall that MAIDs are syntactically the same as CGs. Therefore, both can be defined as `MAID` objects using PyCID. In the case of MAIDs, a level one model, all class methods are permitted except those that involve causal interventions.

A MAID is parameterised by assigning domains to the decision variables and CPDs to every chance and utility variable. CPDs in PyCID are `StochasticFunctionCPD` objects, and there are multiple ways to define them. Listing 1 shows how to instantiate a MAID for Example 1. The CPD for T follows a Bernoulli distribution with success probability $\frac{1}{2}$; D^1 and D^2 are decision variables for the worker and firm's hiring system respectively with binary domains; and U^1 and U^2 are defined as described in Section 2.2.

```
1  import pycid
2
3  job_market = pycid.MAID(
4      [
5          ("T", "D1"),
6          ("T", "U1"),
7          ("T", "U2"),
8          ("D1", "D2"),
9          ("D1", "U1"),
10         ("D2", "U1"),
11         ("D2", "U2"),
12     ],
13     agent_decisions={
14         1: ["D1"],
15         2: ["D2"],
16     },
17     agent_utilities={
18         1: ["U1"],
19         2: ["U2"],
20     },
21 )
22
23 job_market.draw()
24
25 prob = 1/2
26
27 job_market.add_cpds(
28     T = pycid.bernoulli(prob), # T = 1 corresponds to T = h (hard-
    working)
29     D1 = [0,1], # D1 = 1 corresponds to D1 = g (going to university)
30     D2 = [0,1], # D2 = 1 corresponds to D2 = j (offering a job)
31     U1 = lambda d1, d2, t: 5*d2 - t*d1 - 2*d1*(1-t),
32     U2 = lambda d2, t: 3*t*d2 - 2*(1-t)*d2 - (1-d2)*t,
33 )
```

Listing 1: An instantiation of Example 1 in PyCID.

Listing 2 shows how to instantiate a MAID for Example 2. D^1 and D^2 are decision variables for the two robots with binary domains; U^1 and U^2 are defined as described in Section 5; and B is a chance variable defined as a function of its parent D^1 – it takes value $\neg b$ with probability 1 if $D^1 = \neg q$ and probability $\frac{2}{3}$ if $D^1 = q$.

```
1  warehouse_robots = pycid.MAID(
2      [
3          ("D1", "D2"),
4          ("D1", "U1"),
5          ("D1", "B"),
6          ("B", "U2"),
7          ("B", "U1"),
8          ("D2", "U2"),
9          ("D2", "U1"),
10     ],
11     agent_decisions={
12         1: ["D1"],
13         2: ["D2"],
14     },
15     agent_utilities={
16         1: ["U1"],
17         2: ["U2"],
18     },
19 )
20
21 warehouse_robots.draw()
```

```
22
23  warehouse_robots.add_cpds(
24      B = lambda d1 : {0: 1 if d1==0 else 2/3, 1: None}, # B = 1
        corresponds to B = b (breaking something)
25      D1 = [0,1], # D1 = 1 corresponds to D1 = q (moving quickly)
26      D2 = [0,1], # D2 = 1 corresponds to D2 = p (patrolling)
27      U1 = lambda d1, d2, b: (1 - 0.5*d2)*((1-d1)*2 + d1*5 - 3*b),
28      U2 = lambda d2, b: 6*(1 - (1-d2)*b) - d2,
29  )
```
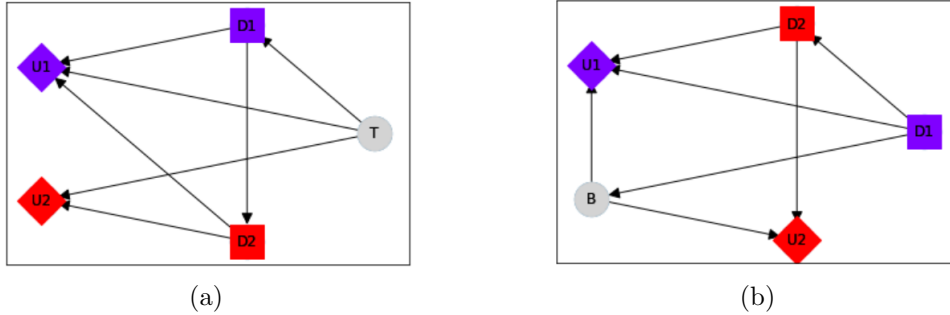
Listing 2: An instantiation of Example 2 in PyCID.

(a)   (b)

Figure 15: MAIDs for (a) Example 1 and (b) Example 2 drawn in PyCID.

C.2 Computing Equilibria

PyCID finds all pure NEs and SPEs in a MAID natively, and finds all behavioural NEs and SPEs in two agent games by converting the MAID into a normal form game and interfacing with Nashpy.[31] We refer the interested reader to our online codebase for up-to-date syntax showing how to compute NEs. All SPEs in a MAID are found by adapting Algorithm 6.2 from [51] to iterate backwards through a topological ordering of the MAID's s-subdiagrams, finding NEs in every s-subgame in turn. More specifically, the method follows the procedure outlined in the proof of Proposition 8. In contrast to K&M's algorithm, our implementation finds all (pure) SPEs, as opposed to just one.

Finally, we demonstrate the computational usefulness of subgames (and SPEs) in MAIDs. Consider a class of games with agents simultaneously choosing whether to place a coin heads or tails face up. The first agent gets utility 1 if and only if all agents in the game choose heads, otherwise their utility equals zero. All other agents are added to the game in pairs, and receive utility according a standard game of matching pennies with their partner: 1 (or -1) utility for matching (heads or tails), and -1 (or 1) for mismatching, respectively. Matching pennies has no pure NEs, only a mixed NE in which both agents randomise equally between choosing heads or tails. A MAID for the 3-agent variant of this game is shown in Figure 16b.

By the above construction, in any game belonging to this class with more than one agent, there must be no pure NE, as one agent in each pair would always have an incentive to deviate. Moreover, the EFG representation of the game (shown for the 3-agent case in Figure 16a) has no proper subgames.[32] Since the complexity of finding even an approximate mixed NE is PPAD-complete [15], finding an NE in this game is hard for any EFG solver, and soon becomes intractable as the number of agent pairs playing matching pennies increases.

In the MAID representation, however, every matching-pennies-playing pair of agents is identified as a proper s-subgame. This means that the unique behavioural SPE (agents randomising

---

[31]Available at https://github.com/drvinceknight/nashpy.

[32]Technically there are six possible corresponding EFGs because of the six permutations of D^1, D^2, and D^3, but none of them have any proper subgames.
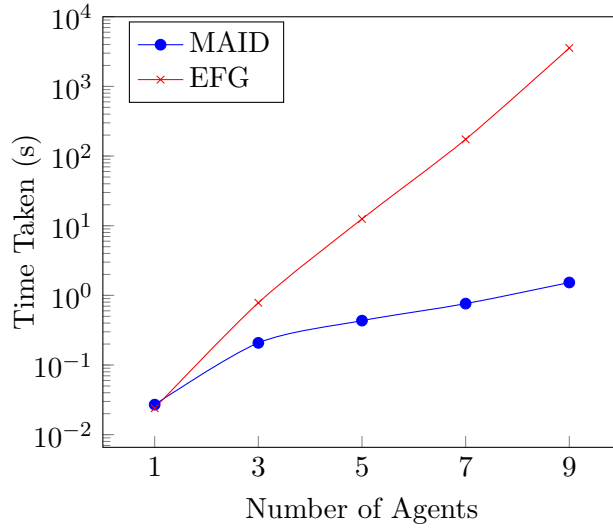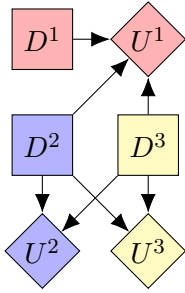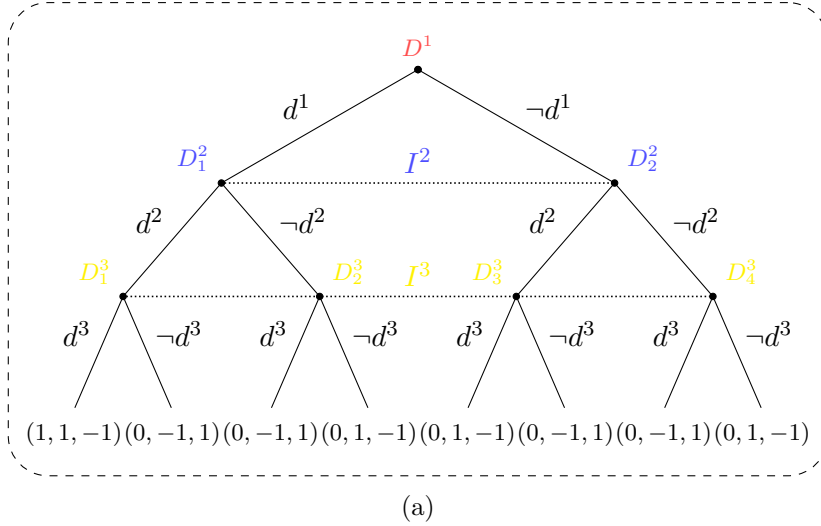
Figure 16: The (a) MAID and (b) EFG for the 3-agent version of the matching-pennies-like game described above. (d) A plot of the time taken to find an NE in the MAID and EFG representations of this game for varying numbers of agents.

equally between choosing heads or tails) can be found in these subgames (with fewer agents) before returning to the full game where, finally, the best response of agent one will be to always play heads. Figure 16c compares the time taken to find an NE in the MAID and EFG representations of games in this game class with between one and nine agents; for the nine-agent game, the difference is three orders of magnitude.[33]

There are games in which neither a MAID nor a corresponding EFG has any proper subgames. In these games, the time taken to compute an NE will be comparable. Nevertheless, many games will have more subgames in the MAID than the EFG (as explained in Section 6.2) and every subgame in an EFG is guaranteed to be a subgame in the corresponding MAID (as proven in Corollary 2). Thus, subgames may often allow for the more efficient computation of equilibria in MAIDs (and hence causal games) compared to EFGs.

---

[33]These calculations were performed using PyCID [27] and Gambit [63] on an NVIDIA Tesla K80 GPU and the time taken is the mean of seven runs.