Causality and Strategic Reasoning

James Fox

Keble College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2024

Acknowledgements

Honouring the subject of this thesis, ‘gratitude’ is itself a (counterfactual) causal concept. In expressing thanks, I acknowledge the possibility of an alternative world devoid of the encouragement and guidance offered by a constellation of brilliant friends, family, teachers, and colleagues, without whom this thesis would have remained an unrealised dream. It has been the privilege of a lifetime to harness the power of collaboration, the strength of shared intellectual curiosity, and the stability of unwavering support to satisfy my itch for research by pursuing this DPhil.

To my supervisors, Alessandro Abate and Mike Wooldridge, I extend my deepest gratitude for your steadfast advice and instruction throughout my DPhil journey, which began amidst the tumultuous backdrop of the COVID-19 pandemic. Your insistence on research purpose, precision, and direction has been the guiding beacon illuminating my path towards completion. To Tom Everitt, I will be eternally grateful that you took a chance on me for the countless hours we have spent immersed in debates about causality, agency, and philosophy. It has been an honour to have learned so much from your exceptional mentorship and boundless generosity of time.

To my thesis examiners, Kobi Gal and Giuseppe De Giacomo, thank you for carefully reading this thesis. I enjoyed my DPhil defence, and your insights and suggestions have significantly improved the clarity of my contributions.

To all my numerous research collaborators, the opportunity to work with each of you has been an extraordinary privilege, a true testament to the potency of collaboration. In particular, Lewis Hammond has been an equal first author behind much of this thesis's work. I am also grateful to Will Lee, Anna Gautier, Jiarui Gan, Paul Harrenstein, and David Hyland – other researchers in my group – for transforming our workspace and meetings into an oasis of curiosity and inquiry, punctuated by frequent formals filled with wine and Hanabi. To all other members of the Causal Incentives Working Group, especially Ryan Carey, Matt MacDermott, Francis Rhys Ward, and Jon Richens, I thank you for all our thought-provoking debates and your shared belief in the power of causality and game theory for contributing to AI safety and alignment.

To my housemates – Lewis Anderson, George Robinson, Phil Winchester, Liam Elbourne, Alicia Curth, and Laura Stewart– I will forever cherish the memories of our endless musings scribbled upon the blackboard, our sacred ‘standard’ gatherings, and the many happy memories we created beyond the confines of work.

To Wendy Poole, I am eternally grateful for your tireless efforts in shielding me from the burdens of all administrative tasks, for caring for and nurturing all AIMS students, and for being the cornerstone pivotal for our CDT's wonderful existence. Your assistance has been invaluable, a true lifeline in the often-turbulent waters of academic life.

To Oxford, 'that sweet city with her dreaming spires', it has been an honour to call such a beautiful city my home. Your enchanting landscape, from the historic embrace of your cobblestone streets to the tranquil waters of the Cherwell, has been the canvas of my intellectual journey. I will miss cycling with so many friends at OUCC, come rain or shine. I will miss getting lost in the pages of countless books within your magnificent libraries. And, I will miss finding solace in the tranquility of your lush parks and the unique intellectual camaraderie found within your ancient halls.

Finally, to my family, the foundation upon which I owe everything. To my parents, Nigel and Caroline, your unwavering belief and unconditional support have been the bedrock of my every endeavour, the source of strength that has propelled me forward. And to my brothers (and best friends), Daniel and Sebastian, I am so grateful I get to share my life with you; you are my constant inspirations to evolve and aspire towards the best version of myself.

Abstract

This thesis explores the intersection of causal and strategic reasoning, both fundamental concepts in AI, among many other disciplines. Historically, the absence of a unified formal framework accommodating both reasoning approaches has inhibited the formalisation of many important concepts. We offer a solution as *(structural) causal games*. This framework can be seen as extending Judea Pearl's causal hierarchy to the game-theoretic domain or as extending Daphne Koller and Brian Milch's multi-agent influence diagrams (MAIDs) to the causal domain. This dual extension paves the way for a holistic understanding and application of both causal and game-theoretic reasoning in complex decision-making scenarios.

This thesis makes several contributions toward our overall goal. First, we introduce mechanised MAIDs, an extension of MAIDs that enables one to make explicit the dependencies between agents' decision rules and the distributions governing the game. Second, we introduce subgames as well as several equilibrium refinements to MAIDs; in particular, our subgame perfectness refinement can rule out more non-credible threats than in extensive form games (EFGS) and can be exploited to find a Nash equilibrium in a MAID (in best case) exponentially faster than in an EFG. Third, we examine MAIDs in imperfect recall settings, where a Nash equilibrium in behavioural policies may not exist. We overcome this by showing how to solve MAIDs with forgetful and absent-minded agents using mixed policies and two types of correlated equilibrium. We also analyse the computational complexity of key decision problems in MAIDs, and explore tractable cases. Fourth, we present definitions of predictions, interventions, and counterfactuals in causal games and discuss the assumptions required for each. Finally, we showcase a variety of applications of causal games, from designing safe AI systems to analysing economic policy interventions.

Contents

Appendices

x

List of Figures

Notation

Symbol	Object	Page
$\mathbf{Anc}_V$	Ancestors of V	16
$\mathsf{c}_{\mathcal{R}}\mathcal{G}$	Condensed $\mathcal{R}$-relevance graph	133
$\mathbf{Ch}_V$	Children of V	16
$\mathbf{Desc}_V$	Descendants of V	16
do	Do operator	21
$dom(V)$	Domain of V	15
E_V	Exogenous variable for V	23
$\mathscr{E}$	Edges	18
$\mathcal{E}$	Extensive-form game	25
$\mathbf{Fa}_V$	Family of V	16
$\mathcal{G}$	Graph	18
$\mathcal{I}$	Intervention	21
J	Intervention set	58
$\mathcal{M}$	Model	18
$\mathcal{M}(\zeta_k)$	Perturbed model	55
M_V	Mechanism variable for V	34
$\mathsf{m}\mathcal{G}$	Mechanised graph	36
$\mathsf{m}\mathcal{M}$	Mechanised model	36
N	Agents	25
$\boldsymbol{\mathcal{P}}$	Set of primitive interventions	119
$\boldsymbol{P}^i$ $(\dot{\boldsymbol{P}}^i)$	Sets of agent i's behavioural (pure) policies	79
$\boldsymbol{P}$ $(\dot{\boldsymbol{P}})$	Sets of (pure) policy profiles	79
$\mathbf{Pa}_V$	Parents of V	16
Pr or P	Probability distribution	18, 25
$\Pr^{\boldsymbol{\pi}}$ or P^{σ}	Probability distribution combining Pr with $\boldsymbol{\pi}$ or P with σ	27, 30
$\mathcal{Q}$	Queries	44
$\mathcal{R}$	Rationality relations	36
$\mathcal{R}(\mathsf{m}\mathcal{M})$	$\mathcal{R}$-Rational outcomes of $\mathcal{M}$	38
r_D	Rationality relation for D	36
$\mathsf{r}_{\mathcal{R}}\mathcal{G}$	$\mathcal{R}$-relevance Graph	40
V	Variable	16
v	Value of V	16
$\boldsymbol{V}$	Variables	16
$\boldsymbol{v}$	Values of $\boldsymbol{V}$	16
δ	Kronecker delta function	21
Δ	Probability simplex	16
Θ_V	Parameter variable for V	34
$\boldsymbol{\theta}$	Parameters	18
μ^i	Mixed policy for Agent i	79

$\boldsymbol{\pi}^i$	(Behavioural) Policy for Agent i	30
$\dot{\boldsymbol{\pi}}^i$	Pure (behavioural) policy for Agent i	30
$\boldsymbol{\pi}$	(Behavioural) policy profile	30
Π_D	Decision rule variable for D	34
σ	(Behavioural) strategy profile	27
$\perp_{\mathcal{G}}$	d-separated in $\mathcal{G}$	19
$\perp\!\!\!\perp$	Independent	19
$\prec$	Topological ordering	133

1
Introduction

Contents

1.1 Motivation

Both causal and game-theoretic reasoning are core capabilities for intelligent systems and fundamental research topics in artificial intelligence (AI). Causal reasoning is concerned with identifying causal relationships and estimating the effects of interventions (i.e., What would be the effect of doing X?). Game-theoretic reasoning, on the other hand, focuses on the strategic behaviour of rational decision-makers interacting in multi-agent settings, taking into account each other agents' incentives. While formal treatments of causality [1–4] and the strategic foundations of multi-agent systems [5–7] have individually led to many recent applications in AI, our focus in this thesis is a formalism that *combines* causal and strategic reasoning.

Models that support both these kinds of reasoning offer a wide range of possible applications for designing safe and beneficial AI systems, including analysing incentives [8–13], fairness [14–18], deception [19], agent discovery [20], generalisation [21], and intention [22–24]. As systems of multi-agent systems become increasingly ubiquitous and sophisticated, the problem of how to formally reason about these notions becomes increasingly important. A framework that supports causal analysis of systems containing multiple self-interested agents would therefore appear to be of great importance [8, 24, 25].

Causal questions are often more challenging to answer in multi-agent settings because one must consider not only the causal dependencies present in the environment but also the dependencies between agents' strategies. For example, in a multi-agent supply chain, the effect of a disruption in one agent's production process on the overall system performance depends on how the other agents adapt their strategies in response (and, indeed, whether each other agent knows about this intervention). Similarly, reasoning strategically about the effects of an action or what other agents would have done under different circumstances naturally leads one to consider counterfactual possibilities when analysing games. However, these causal concepts are typically not explicitly represented in game-theoretic models, limiting their ability to capture and reason about these complex interactions.

The central purpose of this thesis is to introduce a unifying framework for modelling games that supports both causal and strategic reasoning. This framework – *(structural) causal games* – can be interpreted in two different ways. First, it lifts Koller and Milch's (K&M) multi-agent influence diagrams (MAIDs) [26] from the probabilistic models at level one of Pearl's 'causal hierarchy' [3] 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 can employ existing game-theoretic concepts such as *strategic relevance* [26] and *sufficient recall* [27], which can be elicited purely from the structure of the game. These graphical criteria, as shown in Chapter 3 and Chapter 5, help identify variables relevant to an agent's decision-making and characterise different types of imperfect recall, such as forgetfulness and absent-mindedness.

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 objective(s) 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.

The connections between the models discussed in this thesis and their various acronyms are displayed in Fig. 1.1. Notably, the models in each row are a particular 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 13 and the equilibrium refinements we introduce in Section 4.3) are also well-defined in both causal games (CGs) and structural causal games (SCGs).

level ↑			
	SCM	SCIM	SCG
	CBN	CID	CG
	BN	ID	MAID
	agents →		

B	Bayesian
C	Causal
G	Game
I	Influence
M	Model
MA	Multi-Agent
N	Network
S	Structural

Figure 1.1: In this thesis, we introduce CGs and SCGs. Judea Pearl's 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, i.e., the distributions governing the object level variables can be explicitly represented in the graph using mechanism nodes.

Motivating example

To make this all more concrete, consider the following example – a variation of Michael Spence's Job Market Signaling game [28] – that will be used throughout this thesis.

Example 1 (Job Market)**.** *A worker who is either hard-working or lazy is hoping to be hired by a firm. They can pursue 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 uses an automated hiring system that can only observe the worker's education, not their temperament.*

The worker receives a payoff of 5 *if given a job offer, but they incur a cost of* 1 *for going to university if they are hard-working and a* 2 *cost 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*.*

This scenario involves two agents, the worker and the firm, where each agent's strategic decisions significantly influence the other's optimal strategy. Such interdependencies squarely position our analysis within the domain of game theory. In Section 2.2.1, we explore how this scenario can be modelled as an extensive form game – a traditional representation of dynamic games [29]. This modelling approach is both straightforward and powerful, enabling a clear visualisation of the game's structure. It depicts the available information for each agent, the consequences of their chosen actions in terms of payoffs, and more. However, beyond the strategic dynamics, Example 1 raises questions of a causal nature, which cannot easily be answered using an extensive form game representation. We now give some examples of causal questions that sit at different levels of Pearl's causal hierarchy of association ('seeing'), intervention ('doing'), and counterfactuals ('imagining'). Later, in Chapter 6, we will demonstrate how (structural) causal games enable these queries to be answered.

- **Association**. This category covers 'What is?' questions, aiming to identify potential correlations without asserting causality. An example question might be, 'Is there a correlation between obtaining a university degree and the likelihood of being hired by the firm?' Such questions seek to uncover associations between variables—like education and employment outcomes—without suggesting a direct causal link.

- **Intervention** These are ‘What if?’ questions, probing the potential effects of deliberate changes to the system from outside the system. For instance, ‘What would be a worker’s expected well-being if forced to pursue a university education?’ Here, the focus is on understanding the impact of specific interventions (e.g., obtaining a university degree) on desired outcomes (such as the worker’s expected well-being). This inquiry implies a causal relationship by exploring the consequences of intentional alterations, underlining the need for a model which faithfully captures the causal relationships between the variables we care about.

- **Counterfactuals** These questions delve into ‘What if things had been different?’ scenarios, offering a window into alternate realities. An example might be, ‘Had the lazy worker pursued a university education, contrary to their actual decision not to, what would have been the likely hiring decision?’ This level of questioning engages with counterfactual reasoning, speculating how outcomes might have differed under alternative circumstances. It specifically investigates the causal impact of hypothetical interventions (like pursuing education) on outcomes (e.g., hiring decisions), thereby imagining how changing one element might have altered the result.

1.2 Related Work

Causal games build on Pearl’s hierarchy of causal models [3], and are a refinement of influence diagrams (IDs) – a kind of graphical model used to capture single-agent decision-making scenarios [30]. Recent works have explicitly considered *causal* IDs, or CIDs [8, 31]. Indeed, Heckerman and Shachter’s proposal to unify causal modelling and decision analysis via CIDs [32] 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.

Multi-agent IDs (MAIDs) generalise IDs [26, 27], and were originally developed by K&M to efficiently represent and solve games. One helpful 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, to find an equilibrium, it is often possible to remove some of the edges (i.e. the *ignorable information*) before solving the game [27]. Many other graphical models of games, often inspired by MAIDs, have been introduced [33], including networks of influence diagrams (NIDs) [34], interactive dynamic influence diagrams (I-DIDs) [35, 36], and temporal action graph games [37]. Like MAIDs, however, none of these works incorporate 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 [38, 39], chain graphs [40], probabilistic relational models [41, 42], and credal networks [43], among others. In our 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 [44] but without any reference to the game-theoretic scenarios that are our focus.

Perhaps the most similar work to causal games is on *settable systems*[1] which are partly inspired by generalised structural causal models and can be used to capture optimisation, equilibria, and learning [46, 47]. To deal with cycles, settable systems duplicate intervenable variables into a 'response' and 'setting' variable 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

[1] We note that Gonzalez-Soto et al. also define a kind of causal game [45]. However, 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.

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. In contrast, we focus on identifying subgames and equilibrium refinements. Despite these differences, settable systems are 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 [48] – for deterministic [49] and probabilistic [50] concurrent games. Other related approaches to concurrent games use simpler mathematical formalisms [51, 52]. 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. Instead, we are interested in more high-level models of strategic interactions that can be applied to various 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) [22–24], harm [53], incentives to control or respond to certain variables [8, 9], [14–17], social influence [54, 55], and reasoning patterns such as manipulation, signalling, and deception [19, 25]. In this thesis, we show that causal games subsume the models in these works and allow for even richer concepts.

Besides analysing AI systems, economic analysis is another key application domain, particularly mechanism design. For example, Toulis and Parkes use a behavioural causal model to determine the long-term effects of policy interventions on multi-agent economies [56] – though their emphasis is on dynamical systems and behavioural analysis – and some regulators such as the UK's Financial Conduct Authority include an informal 'causal chain' in their cost-benefit analyses when proposing policy analyses [57]. In this thesis, we show how causal games provide a framework to formalise *causal mechanism design*, which enables the rigorous study of economic interventions in an easily readable graphical model.

1.3 Thesis Overview

Chapter 2. To provide the necessary background for the contributions of this thesis, we start with a short overview of Pearl's causal hierarchy [3] and two representations of strategic settings (games): extensive form games (EFGs) [58, 59] and multi-agent influence diagrams (MAIDs) [26]. Every representation has its weaknesses, so the usefulness of a particular representation rests on its relative strengths at performing the task of interest. Although EFGs are the most common representation of dynamic non-cooperative games, we shall see in this thesis that MAIDs are uniquely suited to serve as a robust, complete, and flexible foundation for answering causal queries in multi-agent settings.

Chapter 3. In this chapter, we begin this thesis' original contributions by introducing *mechanised MAIDs* to allow us to model the fact that agents' choices of policy may be dependent on some parameters of the game (as well as other agents' policies). In particular, we care about which of these dependencies are *relevant* for each agent's choice of each of its decision rules. This depends on the rationality assumptions employed to predict an agent's decision choices; we capture this with *rationality relations*, which generalises K&M's notion of strategic relevance to $\mathcal{R}$-*relevance*. The contributions of this chapter serve as a foundation for much of the work in later chapters.

Chapter 4. In this chapter, we show that mechanised graphs and $\mathcal{R}$-relevance allow us to introduce the concept of *subgames* in MAIDs, which, analogously to their EFG counterparts, allow us to analyse and solve parts of the full game independently. We then introduce two equilibrium refinements to MAIDs – subgame perfectness and trembling hand perfectness. Since more subgames can be identified in MAIDs than in EFGs, our subgame perfectness equilibrium refinement in MAIDs can rule out more non-credible threats than in EFGs. We also empirically show that this feature can be exploited to find an NE in a MAID (in the best case) exponentially faster than in an EFG using our open-source Python codebase, PyCID [60]. Finally, to place MAIDs on a more equal footing compared with EFGs as a model for game-theoretic analysis, we formalise procedures for transforming between any EFG and its corresponding MAID (and vice versa), and we prove several equivalence results.

Chapter 5. In this chapter, we turn to imperfect recall in MAIDs. Both types of imperfect recall – forgetfulness and absent-mindedness – can prevent the existence of an NE in behavioural policies. Yet, imperfect recall is essential for handling bounded rationality, teams with imperfect communication, or memoryless policies in Markov games. To overcome this, we show how to handle mixed and correlated equilibria in MAIDs. We also show that a MAID's mechanised graph (introduced in Chapter 3) makes graphically explicit the semantic difference between behavioural and mixed policies (hidden in EFGs), and readily identifies forgetful or absent-minded agents (or teams of agents). We also formulate two definitions of *correlated equilibrium* in MAIDs. The first follows from the normal-form game definition [61]. Still, we show that the second, based on von Stengel and Forges' extensive-form correlated equilibrium [62], is more natural for dynamic settings, can yield greater social welfare, and is easier to compute. Again, mechanised graphs depict the assumptions made in both. Finally, in this chapter, we examine MAIDs from a computational complexity perspective by studying the decision problems of finding a best response, checking whether a policy profile is an NE, and checking whether each type of NE exists. These provide an insight into what makes particular instances hard, when computations can be made tractable, and rigorously identifies which problems are suitable for analysis as MAIDs. All our results also apply to refinements of MAIDs, such as the *(structural) causal games* we introduce in the next chapter.

Chapter 6. In this chapter, we lift MAID's up Pearl's causal hierarchy by introducing *(structural) causal games*. We describe how these models can answer conditional, interventional, and counterfactual queries in multi-agent settings. 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. We show how some causal concepts can be defined directly in EFGs whilst demonstrating the shortcomings of this approach relative to causal games. Finally, we demonstrate the promise of causal games for economic analysis.

Chapter 7. Finally, we summarise some of the advantages and disadvantages of causal games compared with EFGs. We also mention a variety of future work that can build upon this thesis's predominantly theoretical contributions.

1.4 Publications

The bulk of this thesis is based on the following peer-reviewed conference and journal papers:

(* = equal first-author contribution)

1. 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.

2. 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.

3. Hammond, L.*, Fox, J.*, Everitt, T., Carey, R., Abate, A., and Wooldridge, M. "Reasoning about Causality in Games". *Artifical Intelligence* (2023) 320, 103919.

4. Fox, J., MacDermott, M., Harrenstein, P., Abate, A., and Wooldridge, M. "On Imperfect Recall in Multi-Agent Influence Diagrams". *Proceedings of the 19th Conference on Theoretical Aspects of Rationality and Knowledge* (2023) pp 201-220. (Best paper award)

5. Mishra, M., Fox, J., and Wooldridge, M. "Characterising Interventions in Causal Games". *Proceedings of the 40th Conference on Uncertainty in Artificial Intelligence* (2024).

Content from these other technical papers that I have completed during my DPhil were not included in this thesis since they do not fit into the narrative:

- Abate, A., Almulla, Y., Fox, J.*, Hyland, D., and Wooldridge, M. "Learning Task Automata for Reinforcement Learning Using Hidden Markov Models." *Proceedings of the 26th European Conference on Artificial Intelligence (ECAI)* (2023).

- MacDermott, M., Fox, J., Belardinelli, M., Everitt, T. "Measuring Goal-Directedness". *Accepted (spotlight) for NeurIPS* (2024).

- Ioana Georgescu, L., Fox, J., Gautier, A., Wooldridge, M. "Fixed-budget and Multiple-issue Quadratic Voting". *Under Review for ECAI* (2024).

2
Background

Contents

In this chapter, we review the necessary background for this thesis' contributions. We first review Pearl's causal hierarchy [3] and two representations of games: extensive form games (EFGs) [58, 59] and multi-agent influence diagrams (MAIDs) [26]. EFGs are the most common representation of dynamic games, but in the course of this thesis we will demonstrate why, for many classes of problems, MAIDs (and our causal variants) offer several advantages over EFGs, especially for answering causal queries in games.

Notation

Throughout this thesis, 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 (finite) domain of V 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. For example, for the directed acyclic graph in Fig. 2.1, $\mathbf{Pa}_{X_4} = \{X_2, X_3\}$, $\mathbf{Ch}_{X_4} = \{X_5\}$, $\mathbf{Anc}_{X_4} = \{X_1, X_2, X_3\}$, and $\mathbf{Desc}_{X_3} = \{X_4, X_5\}$.

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 the values of $\boldsymbol{A}$ 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 $\boldsymbol{D}^i = \{D^i_1, \ldots, D^i_m\}$.

2.1 Pearl's Causal Hierarchy

Judea Pearl's *causal hierarchy* consists of three kinds of model: associational, interventional, and counterfactual [3]. 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).

2.1.1 Association

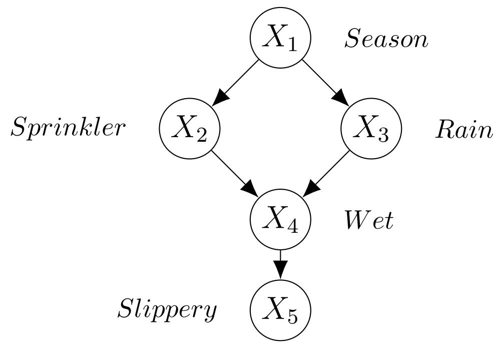


Figure 2.1: A simple (causal) Bayesian network.

A Bayesian network (BN) [63] is situated at the first level of Judea Pearl's causal hierarchy, the level of association, where the focus is on observing and quantifying statistical correlations between variables. This level deals with "What is?" type questions and does not inherently imply causal relationships but rather identifies and quantifies the strength of associations.

BNs are probabilistic graphical models that represent the conditional dependencies present among a collection of random variables. Doing this typically leads to exponential efficiency savings in representation space and inference time costs. Consider the fact that storing an arbitrary joint distribution $\Pr(v_1, v_2, \ldots, v_n)$ over n variables explicitly would require a table with 2^n entries. Bayesian networks can enable enormous efficiency savings by identifying that many variables depend on the state of just a small subset of other variables.

A BN is structured as a directed acyclic graph (DAG), where each node denotes a distinct variable, and directed edges depict the conditional dependencies among these variables. We use a straightforward example from Pearl [64] in Definition 2, which describes relationships between the season of the year (X_1), whether rain falls (X_2), whether the sprinkler is on (X_3), whether the pavement is wet (X_4), and whether the pavement is slippery (X_5). All variables are binary (taking a value of either true or false) except for the root variable X_1, which can take one of four values: spring, summer, autumn, or winter. The absence of a direct link between X_1 and X_5 captures the understanding that the influence of seasonal variation on the slipperiness of the pavement is mediated by other conditions (e.g., the wetness of the pavement).

The graph structure of a BN encodes the Markov property, which states that each variable is conditionally independent of its non-descendants given its parents. This property allows for the compact representation of the joint distribution as a product of conditional probability distributions (CPDs) associated with each variable, where each CPD depends only on the variable's parents in the graph. The parameters of the CPDs are denoted by θ, and they quantify the strengths of the conditional dependencies. In maths, the Markov property says $\Pr(\boldsymbol{v};\boldsymbol{\theta}) = \prod_{V\in\boldsymbol{V}} \Pr(v \mid \mathbf{pa}_V;\theta_V)$. So, the graph in Fig. 2.1 says that the complete joint distribution over the domain can be factorised as $\Pr(X_1, X_2, X_3, X_4, X_5) = \Pr(X_1)\Pr(X_2 \mid X_1)\Pr(X_3 \mid X_1)\Pr(X_4 \mid X_2, X_3)\Pr(X_5 \mid X_4)$.

Definition 1 (3)**.** *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 $\Pr(\boldsymbol{v};\boldsymbol{\theta}) = \prod_{V\in\boldsymbol{V}} \Pr(v \mid \boldsymbol{pa}_V;\theta_V)$. We drop the parameters $\boldsymbol{\theta} = \{\theta_V\}_{V\in\boldsymbol{V}}$ of the CPDs from our notation where unambiguous.*

Graphical Criteria

A graphical criterion in the context of BNs refers to a set of rules or conditions that are used to determine the independence or dependence between variables within the network, based on its graphical structure alone. BNs, as directed acyclic graphs (DAGs), visually represent the conditional dependencies among variables. Graphical criteria, such as *d-separation* (in Definition 2) [63], provide a method for examining these graphical structures to infer whether a set of variables is independent of another set, given a third set of conditioning variables.

The usefulness of graphical criteria lies in their ability to simplify the process of assessing relationships between variables without necessarily computing the actual probabilities. This simplification is particularly advantageous in complex networks with a large number of variables, where calculating conditional independencies directly from probability distributions would be computationally intensive and practically infeasible. By applying graphical criteria, one can quickly identify the structure of the dependencies and independencies within the network, which is essential for both the construction and interpretation of BNs.

Definition 2 (3)**.** *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 $W \notin \boldsymbol{Anc_Y} \cup \{Y\}$.*

*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 $\boldsymbol{X} \not\perp_{\mathcal{G}} \boldsymbol{Z} \mid \boldsymbol{Y}$.*

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 $\mathcal{G}$ [65].

Now, consider the independence relationships that d-separation reveals in Fig. 2.1. Some of these are obvious. For example, knowing that the pavement is wet (X_4) renders the pavement's slipperiness (X_5) independent of $\{X_1, X_2, X_3\}$ because X_4 'blocks' the flow of information along all relevant chain ($X \to W \to Z$) paths. Colliders ($X \to W \leftarrow Z$) act the opposite way from forks ($X \leftarrow W \to Z$). Once we know the season (X_1), then X_3 and X_2 are independent (because $X_2 \leftarrow X_1 \to X_3$

is blocked when conditioning on X_1); whereas, finding that the pavement is wet or slippery renders X_2 and X_3 dependent, because conditioning on X_4 or X_5 causes X_2 and X_3 to become probabilistically dependent (i.e., refuting one of these explanations increases the probability of the other).

We have seen how d-separation provides for a formal definition of the probabilistic relationship that holds between nodes in the network. The following definition describes a relationship that holds between CPDs of nodes in the network. For example, we can ask whether changing the CPD of one node in a BN will affect some probabilistic query.

Definition 3 (66)**.** *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 $\Pr(\boldsymbol{x} \mid \boldsymbol{y}; \boldsymbol{\theta}) \neq \Pr(\boldsymbol{x} \mid \boldsymbol{y}; \boldsymbol{\theta}')$.*

Gieger et al. [67] proposed a graphical criteria for evaluating requisite probability nodes in BNs. This will be useful for some results in Section 3.3.

Lemma 1 (67)**.** *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 $\mathsf{M}_V \not\perp_{\mathsf{m}_\perp \mathcal{G}} \boldsymbol{X} \mid \boldsymbol{Y}$, where M_V is a new parent added to V.*

2.1.2 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. Instead, a causal model, or 'level two' model, such as a causal Bayesian network (CBN), is required. Of course, we almost always use our implicit causal understanding of the world when we draw a BN's DAG (for example, in Fig. 2.1); however, to be a CBN, the DAG must obey certain formal conditions outlined in Definition 4. The intuition behind the additional assumption made for CBNs is that each parent-child relationship in the BN represents a stable and autonomous physical mechanism (how the values of the child variable depend on the values of its parents, which is often captured by a CPD) – i.e., that it is conceivable to change one such relationship without changing the others. It is this modularity that permits one to predict the effect of external interventions (such as for economic policy analysis in Chapter 6, identifying an agent's adaptability [20], or spotting intentional deception [68]) because the directed edges in a CBN 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, $\Pr(\boldsymbol{V}_{\boldsymbol{y}})$ [3].[1]

Definition 4 (3)**.** *A **causal Bayesian network (CBN)** is a BN $\mathcal{M} = (\mathcal{G}, \boldsymbol{\theta})$ such that $\mathcal{G}$ is Markov compatible with $\Pr_{\boldsymbol{y}}$ for every $\boldsymbol{Y} \subseteq \boldsymbol{V}$ and $\boldsymbol{y} \in dom(\boldsymbol{Y})$, and that:*

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

By the Law of Total Probability (i.e., $P(A) = \sum_n P(A \cap B_n)$), $\Pr_{\boldsymbol{y}}(v \mid \mathbf{pa}_V) = 0$ 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 $\mathbf{Pa}_Y$.[2] 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.

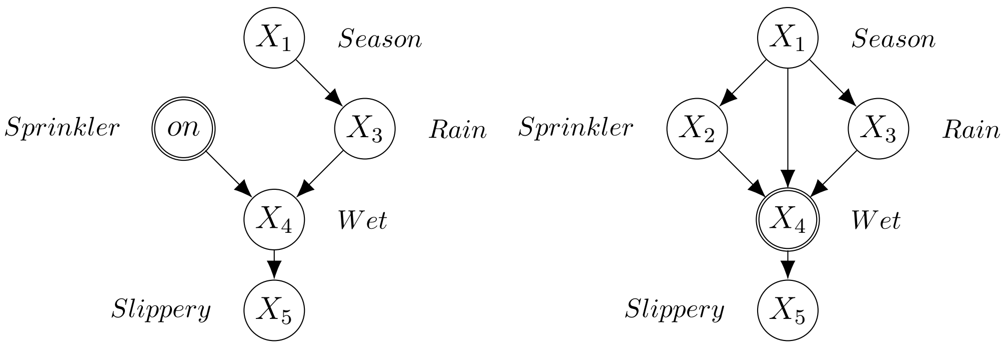
[1] This is also known as an atomic [69], structural [70], surgical [3], or independent [71] intervention.

[2] A soft intervention is also known as a parametric [70] or dependent [71] intervention, and is referred to as conditional or stochastic when deterministic or stochastic, respectively [69]. 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.

When $\mathcal{I}$ is a hard intervention, we simply sever all incoming edges to $\boldsymbol{Y}$, setting their values to $\boldsymbol{y}$, and write $\boldsymbol{V_y}$, $\mathcal{G}_{\boldsymbol{y}}$, and $\mathcal{M}_{\boldsymbol{y}}$ respectively.

Examples of hard and soft interventions are shown in Fig. 2.2a and Fig. 2.2b, respectively. In Fig. 2.2a, we illustrate the DAG $\mathcal{G}_g$, which represents the CBN in Fig. 2.1 after a hard intervention do($X_2 = on$) (forcing the sprinkler to be on). Notice that the edge $X_1 \to X_2$ has been pruned because this intervention only affects things downstream of it. In Fig. 2.2b, we represent the effect of a soft intervention on X_4 which results in the wetness of the pavement now being directly causally dependent on the season (perhaps because the wetness of the pavement also depends on whether the gazebo is set up – which is decided directly by the season).

Although in this background section, we have illustrated these interventions on a normal CBN, in Chapter 6, this thesis will study how interventions are handled in game-theoretic situations.



(a) $\mathcal{G}_{X_1=on}$ **(b)** $\mathcal{G}_{\mathcal{I}}$

Figure 2.2: (a) The graph $\mathcal{G}_{X_2=on}$ representing the CBN in Fig. 2.1 after a hard intervention do($X_1 = on$). (b) A similar graph $\mathcal{G}_{\mathcal{I}}$ where $\mathcal{I}(X_4 \mid \mathbf{Pa}^*_{X_4})$ is a soft intervention with $\mathbf{Pa}^*_{X_4} = \mathbf{Pa}_{U^2} \cup \{X_1\}$.

2.1.3 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 $\Pr(\boldsymbol{x_y} \mid \boldsymbol{z})$. To answer such questions in general, one must appeal to level three of the causal hierarchy, such

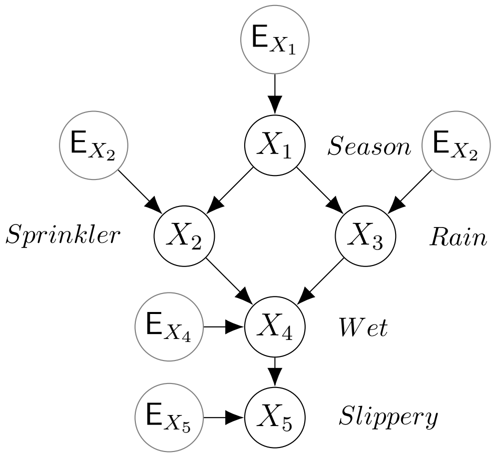


Figure 2.3: An SCM representing the same system as represented by the (C)BN in Fig. 2.1. Each variable V is now associated with an exogenous parent $\mathsf{E}_V \in \mathbf{E}$.

as by using a structural causal model (SCM). In SCMs, variables are partitioned into exogenous and endogenous sets $\mathbf{E}$ and $\boldsymbol{V}$ respectively, where each endogenous variable $V \in \boldsymbol{V}$ is deterministically related to its parents via a structural function $f_V : dom(\boldsymbol{V} \setminus \{V\}) \times dom(\mathbf{E}) \to dom(V)$ that specifies the causal mechanism governing the values of the variable, and where all stochasticity is relegated to the distribution $\Pr(\mathbf{E}; \boldsymbol{\theta})$ over the exogenous variables.

In this thesis, 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 Chapter 6. In Fig. 2.3, we show an SCM representing the same system as represented by the (C)BN in Fig. 2.1. Each variable V is now associated with an exogenous parent $\mathsf{E}_V \in \mathbf{E}$.

Definition 5 (3)**.** *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* $\boldsymbol{Pa}_V \cap \mathbf{E} = \{\mathsf{E}_V\}$*. The parameters* $\boldsymbol{\theta}$ *assign deterministic distributions* $\Pr(v \mid \boldsymbol{pa}_V; \theta_V)$ *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 [3]:

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

2. Apply the intervention $\mathcal{I}(\boldsymbol{Y} \mid \mathbf{Pa}^*_{\boldsymbol{Y}})$ 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 diagrams (MAIDs). Both of these models are illustrated using Example 1.

2.2.1 Extensive Form Games

Extensive form games (EFGs) are characterised by a game tree, which consists of a directed graph in which the set of vertices represents positions in the game. The root vertex represents the starting position of the game and the terminal vertices represent where play ends. Non-terminal vertices denote either chance events, with associated probability distributions for outcomes, or agent moves. *Imperfect information* games introduce *information sets* to represent decision points

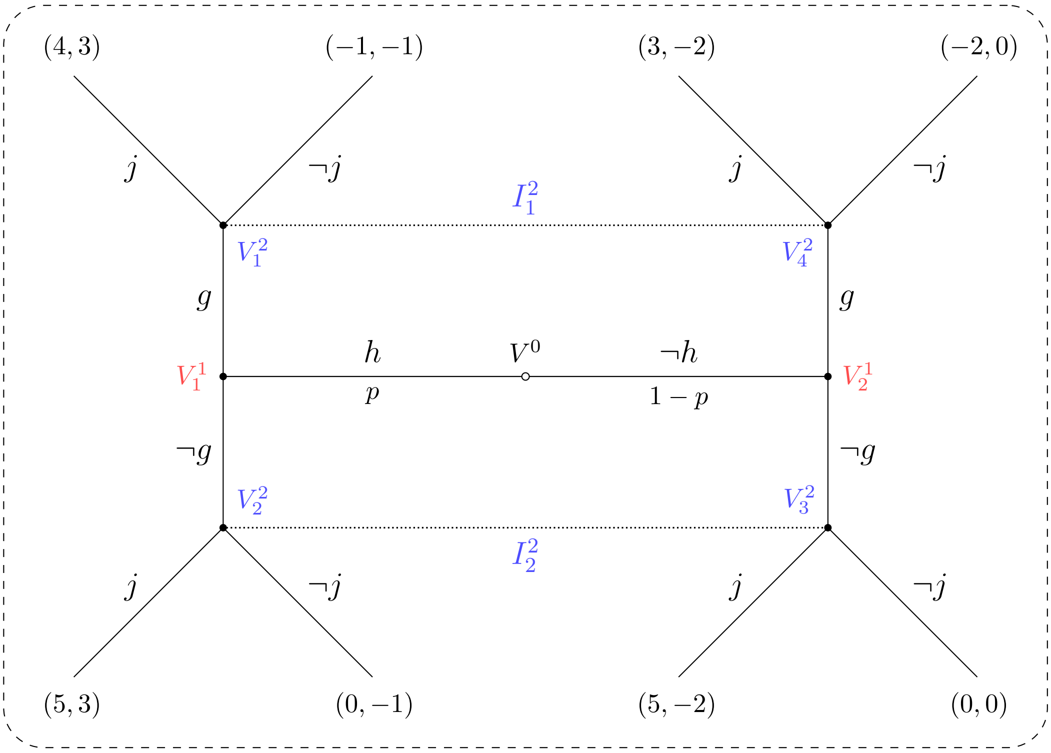


Figure 2.4: An EFG representing Example 1. V^0 represents the initial move made by nature. V_1^1, V_2^1 and V_1^2, V_2^2, V_3^2, & V_4^2 represent moves made by Agent 1 (the worker) and Agent 2 (the firm), respectively. I_1^2 and I_2^2 represent Agent 2's non-singleton information sets.

indistinguishable from an agent given their information at that stage of the game. A *perfect information* game is one in which all information sets are singletons; in such a game, whenever an agent makes an action, they know the full history of actions and chance moves that led to that position.

In an EFG, an agent's *strategy* is a function that assigns to each of their information sets an action available to them at that information set. A path from the root to a terminal vertex is called a play of the game. When the game has no chance moves, any vector of strategies (one for each agent) determines the play of the game, and hence the outcome. In a game with chance moves, any vector of strategies determines a probability distribution over the possible outcomes of the game. We now give the formal definition of an EFG (see Maschler et al. [29] for more details).

Definition 6 (59)**.** *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 $\mathscr{E} \subseteq \boldsymbol{V} \times \boldsymbol{V}$.*
- *$P = \{P_1, \ldots, P_{|\boldsymbol{V}^0|}\}$ is a set of probability distributions $P_j(\boldsymbol{Ch}_{V^0_j})$ over the children of each chance node V^0_j.*
- *A is a set of actions, where $A^i_j \subseteq A$ denotes the set of actions available at $V^i_j \in \boldsymbol{V}^i$.*
- *$\lambda : \mathscr{E} \to A$ is a labelling function mapping each edge (V^i_j, V^k_l) to an action $a \in A^i_j$.*
- *$I = \{I^1, \ldots, I^n\}$ 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 $I^i_j \in I^i$ is defined such that for all $V^i_k, V^i_l \in I^i_j$, the available actions $A^i_j \coloneqq A_{V^i_k} = A_{V^i_l}$ 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.*

Fig. 2.4 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 $V^1_1, V^1_2 \in \boldsymbol{V}^1$ and the automated hiring system's decisions are given by $V^2_1, V^2_2, V^2_3, V^2_4 \in \boldsymbol{V}^2$, 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 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 $\sigma^i_j(A^i_j)$ over the actions available to the agent at each of their information sets I^i_j. A strategy is* ***pure*** *when each $\sigma^i_j(a) \in \{0, 1\}$ and* ***fully stochastic*** *when $\sigma^i_j(a) > 0$, for all $a \in A^i_j$. A* ***strategy profile*** *$\sigma = (\sigma^1, \dots, \sigma^n)$ is a tuple of strategies, and $\sigma^{-i} = (\sigma^1, \dots, \sigma^{i-1}, \sigma^{i+1}, \dots, \sigma^n)$ denotes the partial strategy profile of all agents other than i, hence $\sigma = (\sigma^i, \sigma^{-i})$.*

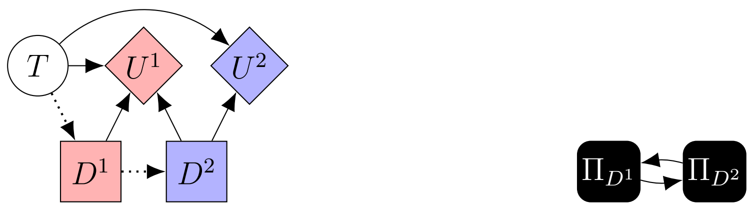
Combining a strategy profile σ with the distributions in P defines a probability distribution P^σ over paths ρ in $\mathcal{E}$. For each path ρ beginning from R and terminating in a leaf node $\rho[\boldsymbol{L}] \in \boldsymbol{L}$, agent i receives utility $U(\rho[\boldsymbol{L}])[i]$ – the i^{th} 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^i_j in $\mathcal{E}$, if there exists $V_k \in I^i_j \cap \boldsymbol{V}'$, then $I^i_j \subseteq \boldsymbol{V}'$; 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 thesis, we denote subgames in EFGs by enclosing them within dashed boxes. For instance, the EFG in Fig. 2.4 has no proper subgames, but the EFG given later in Fig. 4.2c has two.

2.2.2 Multi-Agent Influence Diagrams

Influence diagrams (IDs) generalise BNs to the decision-theoretic setting by adding decision and utility variables [30, 72]. The agent in an ID selects a conditional distribution over actions at decision variables to maximise the expected cumulative



(a) MAID **(b)** K&M's strategic relevance graph

Figure 2.5: (a) The graph of a MAID representing Example 1. Recall that T represents the worker's temperament, D^1 is the worker's decision whether to go to university. D^2 is the firm's decision as to whether to hire the worker. U^1 and U^2 are the worker and firm's utility, respectively. (b) K&M's strategic relevance graph.

value of their utility variables. An instantiation $\mathbf{pa}_D$ of the decision variable's parents represents a *decision context* for D – the information available when an action is chosen at D.

Multi-agent influence diagrams (MAIDs) generalise IDs by introducing multiple agents [26, 27]: they offer a graphical representation of dynamic non-cooperative games that can be more compact and expressive than extensive-form games (EFGs) [26]. MAIDs can be viewed as a BN over a graph without parameters for the decision variables, but technically they lie between levels one and two of Pearl's causal hierarchy as the decisions are effectively modelled as causal interventions [32], but paths between non-decision variables need not encode causal relationships [8].

Thanks in part to the contributions of the work presented in this thesis, (MA)IDs, and their causal variants [73], have been used in business and medical decision-making settings [74, 75], for the design of safe and fair AI systems [10, 12, 13, 18, 21, 22, 55, 76–78], to explore reasoning patterns and deception [19, 25, 79], and to identify agents from data [20].

Definition 9 (26)**.** *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 $N = \{1, \dots, n\}$ 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 $\Pr(V \mid \boldsymbol{Pa}_V; \theta_V)$ 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.*

Fig. 2.5 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. The parameters $\boldsymbol{\theta}$ define conditional distributions for variables T, U^1, and U^2, in accordance with the values shown in Fig. 2.4.

There are two types of directed edge in a MAID. Full edges leading into $\boldsymbol{X} \cup \boldsymbol{U}$ represent probabilistic dependence, as in a Bayesian network. Dotted edges leading into $\boldsymbol{D}$ represent information that is available to the agent at the time a decision is made (e.g. the edge $T \to D^1$). In this way, the values of the parents $\mathbf{pa}_D$ of a decision node D represent the decision context for D. The CPDs of decision nodes are not defined when a MAID is constructed because they are instead chosen by the agents playing the game. In general, a player's decision CPD need not be optimal.

This example demonstrates two clear advantages of MAIDs compared with EFGs. First, in many real-world cases, MAIDs make it possible to explicitly represent aspects of game structure that are obscured in the extensive form. For example, in the EFG, information sets were drawn to reflect the fact that the algorithm does not know whether the worker is naturally hard-working or lazy when it selects its action. However, in the corresponding MAID, this incomplete information is represented simply by the fact that there is no edge $T \to D^2$. Moreover, the company's utility U^2 isn't a function of whether the applicant went to university or not – it only cares whether the applicant is hard-working and whether or not they hired them. We can infer this from the EFG payoffs in Fig. 2.4, but in the MAID this is shown explicitly by the fact that there is no edge $D^1 \to U^2$.

Second, MAIDs can provide a more compact graphical representation of games [26, 27]. In fact, the MAID representation of a game need never be bigger than the corresponding EFG and can be smaller in many cases. For example, there are four nodes $V^2_1, V^2_2, V^2_3, V^2_4$ in Fig. 2.4 which correspond to the company's decision. In the MAID, these are combined into one node, D^2.

A further strength of MAIDs derives from them being probabilistic graphical models, and so probabilistic dependencies between chance and strategic variables can be exploited. We can utilise *d-separation* (Definition 2), a graphical criterion for determining the independence properties of the probability distribution associated with the graph. This is necessary for many of the contributions in this thesis (including $\mathcal{R}$-relevance in Section 3.3 and subgames in Section 4.2).

Definition 10. *Given a MAID $\mathcal{M} = (\mathcal{G}, \boldsymbol{\theta})$, a* ***decision rule*** *π_D for $D \in \boldsymbol{D}$ is a CPD $\pi_D(D \mid \boldsymbol{Pa}_D)$ 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 $\pi_D(d \mid \boldsymbol{pa}_D) \in \{0, 1\}$ and* ***fully stochastic***[3] *if $\pi_D(d \mid \boldsymbol{pa}_D) > 0$ for all $d \in dom(D)$ and each* ***decision context*** *$\boldsymbol{pa}_D \in dom(\boldsymbol{Pa}_D)$; this holds for a policy (profile) if it holds for all decision rules in the policy (profile) and is denoted by $\dot{\boldsymbol{\pi}}^i(\dot{\boldsymbol{\pi}})$.*

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

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

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) [80] to be defined, which identifies outcomes of a game where every agent is simultaneously playing a best response.[4]

Definition 11 (26)**.** *A (behavioural) policy profile $\boldsymbol{\pi}$ is a* ***Nash equilibrium (NE)*** *in a MAID if, for every agent $i \in N$, $\boldsymbol{\pi}^i \in \arg\max_{\hat{\boldsymbol{\pi}}^i \in dom(\boldsymbol{\Pi}^i)} \sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{(\hat{\boldsymbol{\pi}}^i, \boldsymbol{\pi}^{-i})}[U]$.*

[3] We reserve the usage of 'mixed' for a *mixed policy*, defined in Section 5.3.

[4] In Chapter 5, we build upon what's already known about NEs in MAIDs, by explaining the difference between mixed policies and behavioural policies in MAIDs and clarifying under what conditions different types of NE are guaranteed to exist.

For Example 1, if we assume that $p = \Pr(T = h) = \frac{1}{2}$, then the NEs of this game are:

1. The worker always chooses $\neg g$. The hiring system chooses j if the worker chooses $\neg g$. 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 $q \in [0, \frac{3}{5}]$.

K&M also define *strategic relevance (s-relevance)* [26] 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 3.2.

Definition 12 (26)**.** *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 $\boldsymbol{pa}_{D_k}$ such that* $\Pr^{\boldsymbol{\pi}'}(\boldsymbol{pa}_{D_k}) > 0$.[5]

[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.

Proposition 1 (26). $\Pi_{D'}$ *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* $\Pi_{D'} \to D'$*.*

K&M use s-reachability to construct a strategic relevance graph (shown in Fig. 2.5b), which displays which decision rules strategically rely on each other.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 the next chapter, we generalise these to $\mathcal{R}$-relevance so that we can also consider the parameterisation of non-decision variables, which will be key to reasoning about causality in games. Our s-relevance graphs (e.g.,) are the same as K&M's strategic relevance graphs (they are just the dependencies of non-decision rules). Our $\mathcal{R}$-relevance graphs, in general, extend strategic relevance graphs to different rationality assumptions (specified by $\mathcal{R}$).

3
Mechanised MAIDs and Relevance

Contents

3.1 Introduction

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

In this chapter, 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 Chapter 6), for the introduction of subgames in MAIDs (in Section 4.2) and for handling imperfect recall (in Chapter 5). In the following sections, we first define mechanised MAIDs before formally introducing the concept of *relevance* and corresponding graphical criteria.

3.2 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.

3.2.1 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, and the variables $\boldsymbol{V}$ in the original MAID *object-level variables.* To distinguish between different types of mechanism variables, we often refer to those for decisions $\mathbf{\Pi} = \mathbf{M}_{\boldsymbol{D}}$ as *decision rule variables* and those for non-decisions $\mathbf{\Theta} = \mathbf{M}_{\boldsymbol{V} \setminus \boldsymbol{D}}$ as *parameter variables.*

In the mechanised graph $\mathsf{m}\mathcal{G}$, we also add new edges $\mathscr{E}' \subseteq \bigcup_{D \in \boldsymbol{D}} \Big((\mathbf{M} \setminus \Pi_D) \times \Pi_D \Big)$ 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 $\mathbf{\Theta}$) 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 (introduced in Chapter 2) 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., $\mathbf{Pa}_{\Pi_D} = \mathsf{M}_{\boldsymbol{V}\setminus\{D\}}$) – 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 3.3. A mechanised graph $\mathsf{m}\mathcal{G}$ for Example 1 is shown in Fig. 3.1a, where decision rule and parameter variables are represented using black and white rounded squares, respectively.

3.2.2 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 $\mathsf{m}_{\boldsymbol{V}\setminus\boldsymbol{D}} = \boldsymbol{\theta} \in dom(\boldsymbol{\Theta})$ of the game and any policy profile $\mathsf{m}_{\boldsymbol{D}} = \boldsymbol{\pi} \in dom(\boldsymbol{\Pi})$, the resulting joint distribution over $\boldsymbol{V}$ can be written as:

$$\overset{\boldsymbol{\pi}}{\Pr}(\boldsymbol{v};\boldsymbol{\theta}) = \Pr(\boldsymbol{v} \mid \mathsf{m}) := \prod_{V\in\boldsymbol{V}} \Pr(v \mid \mathbf{pa}_V, \mathsf{m}_V),$$

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 $\Delta(dom(V) \mid dom(\mathbf{Pa}_V))$. This says that any particular game induced over the graph corresponds with a particular instantiation of the parameter variables in the mechanised game, the values of which are known by all agents.[1]

[1]This simply amounts to a common prior assumption. Though such an assumption can be relaxed [81], doing so introduces additional complexities that we do not address in this thesis.

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* $r_D \subseteq dom(\mathbf{Pa}_{\Pi_D}) \times dom(\Pi_D)$.[2] 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 this thesis, 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 $r_D(\mathbf{pa}_{\Pi_D}) = \pi_D$.[3]

Definition 13. *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** $\mathsf{m}\mathcal{G} = (N, \boldsymbol{V} \cup \mathbf{M}, \mathsf{m}\mathscr{E})$ is a directed (possibly cyclic) graph over $\boldsymbol{V}$ and $\mathbf{M}$ with edges $\mathsf{m}\mathscr{E} := \mathscr{E} \cup \{(\mathsf{M}_V, V)\}_{V \in \boldsymbol{V}} \cup \mathscr{E}'$, where $\mathscr{E}' \subseteq \bigcup_{D \in \boldsymbol{D}} \Big((\mathbf{M} \setminus \Pi_D) \times \Pi_D\Big)$; and $\mathcal{R} = \{r_D\}_{D \in \boldsymbol{D}}$ is a set of **rationality relations**, where each $r_D \subseteq dom(\boldsymbol{Pa}_{\Pi_D}) \times dom(\Pi_D)$ 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 $\mathcal{R}^{\text{BR}} = \{r^{\text{BR}}_{D^1}, r^{\text{BR}}_{D^2}\}$ such that:

$$\pi_D \in r^{\text{BR}}_D(\mathbf{pa}_{\Pi_D}) \quad \Leftrightarrow \quad \pi_D \in \operatorname*{arg\,max}_{\hat{\pi}_D \in dom(\Pi_D)} \sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{(\hat{\pi}_D, \boldsymbol{\pi}_{-D})}[U], \tag{3.1}$$

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* $\mathbf{pa}_{\Pi_D} = \mathbf{m}_{\boldsymbol{V} \setminus \{D\}}$ and π_D. Note that for games in which each agent has only one decision rule, this gives rise to a Nash equilibrium (Definition 11) (so the set of $\mathcal{R}^{\text{BR}}$-rational outcomes is identical to the set of Nash equilibria. 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 (so it is solved just like a MAID is).

[2]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 $(a, b) \in r$.

[3]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.

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 4.3). 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.

It will sometimes be visually useful to restrict a mechanised graph to just the mechanism variables, as can be seen in Fig. 3.1b. 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 $V \in \boldsymbol{V}$.

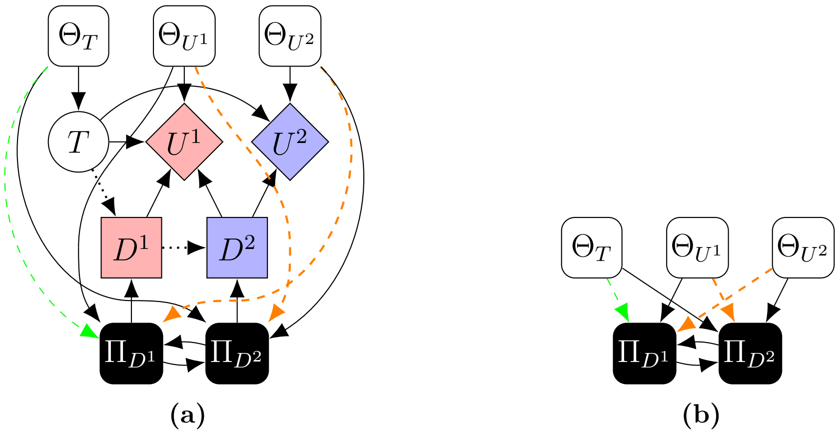


Figure 3.1: (a) A mechanised graph $\mathsf{m}\mathcal{G}$ representing Example 1. Orange dashed edges represent those that are present in neither the $\mathcal{R}^{\text{BR}}$-minimal nor the s-minimal mechanised graph. Green dashed edges represent just those that are not present in the s-minimal mechanised graph. (b) The $\mathcal{R}^{\text{BR}}$-relevance graph is formed by removing the dotted edges, and the s-relevance graph by further removing the dashed edge.

3.2.3 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 $\mathcal{R}^{\text{BR}}$-rational outcomes of the MAID $\mathcal{M}$ 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 14. *Given a mechanised MAID* $\mathsf{m}\mathcal{M}$*, we say that any* $\pi_D \in r_D(\boldsymbol{pa}_{\Pi_D})$ *is an* $\mathcal{R}$-***rational response*** *to* $\boldsymbol{pa}_{\Pi_D}$*, and that a (partial) policy profile* $\boldsymbol{\pi}_{\boldsymbol{D}'}$ *is* $\mathcal{R}$-***rational*** *if* $\pi_D \in r_D(\boldsymbol{pa}_{\Pi_D})$ *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 $\{\Pr^{\boldsymbol{\pi}}\}_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M})}$ over the variables $\boldsymbol{V}$ where $\Pr^{\boldsymbol{\pi}}(\boldsymbol{v};\boldsymbol{\theta}) = \Pr(\boldsymbol{v} \mid \mathbf{m})$ for $\mathbf{m}_{\boldsymbol{D}} = \boldsymbol{\pi}$ and $\mathbf{m}_{\boldsymbol{V} \setminus \boldsymbol{D}} = \boldsymbol{\theta}$, which can be extended over the mechanised MAID as $\Pr(\boldsymbol{v}, \mathbf{m}) = \Pr(\boldsymbol{v} \mid \mathbf{m})\delta(\mathbf{M},\mathbf{m})$. We can thus view a mechanised MAID as a set of BNs induced by the $\mathcal{R}$-rational outcomes.[4] 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 14 is related to the notion of a *solution* in cyclic causal models [39, 44]. 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 Chapter 6). Similarly, we can view the rational responses at a decision variable as a *credal set* [82], and therefore the resultant model as a form of credal network [43], where the imprecise probabilities arise from the fact that there may be more than one rational outcome in a game.

Remark. *Often, agents might only be* boundedly *rational [83]. 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*

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

each agent cannot improve their expected utility by more than $\epsilon > 0$ by deviating [84]. These bounded rationality conditions can also be captured using rationality relations.

3.3 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 Fig. 3.1a, 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*[5] $\mathsf{m}_{\perp}\mathcal{G}$. Game-theoretic models thereby break the *independent causal mechanism* assumption, which states that mechanisms should be causally and probabilistically independent of each other [4].

Formally, we have the following definitions, which are similar in spirit to those proposed in earlier work [26, 27, 85], but differ in that they consider the case of non-deterministic rationality relations that encode the solutions of a game.

Definition 15. *Given a mechanised MAID $\mathsf{m}\mathcal{M}$ with rationality relations $\mathcal{R}$, $\mathsf{M}_V \in \boldsymbol{Pa}_{\Pi_D}$ is $\mathcal{R}$-**relevant** to Π_D if there exists $\boldsymbol{pa}_{\Pi_D} \neq \boldsymbol{pa}'_{\Pi_D}$ such that $r_D(\boldsymbol{pa}_{\Pi_D}) \neq r_D(\boldsymbol{pa}'_{\Pi_D})$, where $\boldsymbol{pa}_{\Pi_D}$ and $\boldsymbol{pa}'_{\Pi_D}$ differ only on M_V.*

Definition 16. *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*

[5][31] call this diagram the extended influence diagram, but they do not investigate the dependencies and relationships between mechanisms.

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 15 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 $\mathcal{R}^{\text{BR}}$-reachability in Proposition 2 enable us to remove the dotted edges $\Theta_{U^1} \to \Pi_{D^2}$ and $\Theta_{U^2} \to \Pi_{D^1}$ in Fig. 3.1a. We now prove this proposition.

Proposition 2. 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 15, that $\mathsf{M}_V \in \mathbf{Pa}_{\Pi_D}$ is $\mathcal{R}^{\text{BR}}$**-relevant** to Π_D if there exists $\mathbf{pa}_{\Pi_D} \neq \mathbf{pa}'_{\Pi_D}$ such that $r^{\text{BR}}_D(\mathbf{pa}_{\Pi_D}) \neq r^{\text{BR}}_D(\mathbf{pa}'_{\Pi_D})$, where $\mathbf{pa}_{\Pi_D}$ and $\mathbf{pa}'_{\Pi_D}$ differ only on M_V. Moreover, recall that for each $D \in \boldsymbol{D}^i$, $\pi_D \in r^{\text{BR}}_D(\mathbf{pa}_{\Pi_D})$ if and only if:

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

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:

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

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 $r^{\text{BR}}_D(\mathbf{pa}_{\Pi_D})$ satisfies:

$$\underset{\hat{\pi}_D \in dom(\Pi_D)}{\arg\max} \left[\sum_{U \in \boldsymbol{U}^i \cap \mathbf{Desc}_D} \sum_{d \in dom(D)} \sum_{\mathbf{pa}_D \in dom(\mathbf{Pa}_D)} \overset{\pi}{\Pr}(u \mid d, \mathbf{pa}_D) \hat{\pi}_D(d \mid \mathbf{pa}_D) \overset{\pi}{\Pr}(\mathbf{pa}_D) \cdot u \right]. \tag{3.2}$$

Therefore, if the expression (Eq. (3.2)) is independent of the value m_V, then M_V is not $\mathcal{R}^{\text{BR}}$-**relevant** to Π_D. As explained in Section 3.3, this is equivalent to asking whether V is a requisite probability node for any distribution in $\mathcal{Q}^{\text{BR}}_D = \{\Pr^{\boldsymbol{\pi}}(\boldsymbol{u}^i \cap \mathbf{desc}_D \mid d, \mathbf{pa}_D), \Pr^{\boldsymbol{\pi}}(\mathbf{pa}_D)\}$. 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 $\mathcal{R}^{\text{BR}}$-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 $\mathcal{R}^{\text{BR}}$-reachable from Π_D, then M_V is not $\mathcal{R}^{\text{BR}}$-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 $\mathcal{R}^{\text{BR}}$-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 (Eq. (3.2)) is independent of the value of M_V and so $r^{\text{BR}}_D(\mathbf{pa}_{\Pi_D}) = r^{\text{BR}}_D(\mathbf{pa}'_{\Pi_D})$, meaning that M_V is not $\mathcal{R}^{\text{BR}}$-relevant to Π_D, as required.

We now turn to completeness, i.e., that if M_V is $\mathcal{R}^{\text{BR}}$-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 $\mathsf{m}\mathcal{M} = (\mathsf{m}\mathcal{G}, \boldsymbol{\theta}, \mathcal{R}^{\text{BR}})$, M_V is $\mathcal{R}^{\text{BR}}$-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 $\mathcal{R}^{\text{BR}}$-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 $r^{\text{BR}}_D(\mathbf{pa}_{\Pi_D}) = r^{\text{BR}}_D(\mathbf{pa}'_{\Pi_D})$ for any $\mathbf{pa}_{\Pi_D}$ and $\mathbf{pa}'_{\Pi_D}$; in this case, M_V is never $\mathcal{R}^{\text{BR}}$-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 $\pi_D \in r^{\text{BR}}_D(\mathbf{pa}_{\Pi_D})$, note that the decision rule $\bar{\pi}_D \in r^{\text{BR}}_D(\mathbf{pa}_{\Pi_D})$ 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 $\Pr^{\boldsymbol{\pi}}(V = 0 \mid \mathbf{pa}_V) > 0$ 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, $\bar{\pi}_D \notin r^{\mathrm{BR}}_D(\mathbf{pa}'_{\Pi_D})$, and hence $r^{\mathrm{BR}}_D(\mathbf{pa}_{\Pi_D}) \neq r^{\mathrm{BR}}_D(\mathbf{pa}'_{\Pi_D})$, 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'. Then, we replace all references in the proof to $V = 0$ (where V is some variable) with $V \neq 1$. The probability mass assigned to $V \neq 1$ in a CPD is uniformly distributed over the values in $dom(V) \setminus 1$. It is straightforward to check that the proof still holds. $\square$

3.3.1 $\mathcal{R}$-minimal mechanised graphs and $\mathcal{R}$-relevance graphs

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 $\mathcal{R}$.[6] 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 variables, we can see in Fig. 3.1a, for example, that Θ_T is not s-relevant to Π_{D^1}, though it is $\mathcal{R}^{\mathrm{BR}}$-relevant.

[6] Note that the direction of the edges of the relevance graphs in this thesis are the same as in Koller et al., [26], but are reversed compared with Koller et al., [86] and Hammond et al., [87]. 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.

Our generalisation of the idea of s-relevance to different rationality relations parallels our development of various equilibrium refinements within MAIDs (introduced in Section 4.3), as opposed to the singular concept introduced by K&M. For instance, we shall see in Section 4.3 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 (in Chapter 6) where agents may adapt their policies in response to interventions on any mechanism variable (not just those representing decision rules). The mechanised graph also makes an agent's type of imperfect recall explicit (see Chapter 5).

3.3.2 Generalising the Soundness and Completeness Results

We conclude this section by noting that it is possible to generalise the proof procedures for Proposition 2 and Proposition 1 to other rationality assumptions (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 $\mathsf{m} = (\boldsymbol{\pi}, \boldsymbol{\theta})$, completely determine each r_D. Formally, let $\mathcal{Q}(\mathsf{m})$ be the set of probabilistic queries $\Pr^{\boldsymbol{\pi}}(\boldsymbol{x} \mid \boldsymbol{y}; \boldsymbol{\theta})$ over $\boldsymbol{X}, \boldsymbol{Y} \subseteq \boldsymbol{V}$ that can be formed in a MAID given $\mathsf{m} = (\boldsymbol{\pi}, \boldsymbol{\theta})$. 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:

$$\pi_D \in r_D(\mathsf{m}_{\boldsymbol{V}\setminus\{D\}}) \quad \Leftrightarrow \quad g_D\Big(\mathcal{Q}_D(\mathsf{m}), dom(\boldsymbol{V})\Big). \tag{3.3}$$

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 $\Pr^{\boldsymbol{\pi}}(\boldsymbol{x} \mid \boldsymbol{y}) \in \mathcal{Q}_D$, which is, in turn, equivalent to asking whether V is a *requisite probability node* for $\Pr^{\boldsymbol{\pi}}(\boldsymbol{x} \mid \boldsymbol{y})$ [66], introduced in Definition 3. Whether a variable is a requisite probability node can be determined using a well-established graphical criterion [67], 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 Proposition 1 (s-reachability) and Proposition 2 ($\mathcal{R}^{\mathrm{BR}}$-reachability) correspond to choices:

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

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

	$\mathsf{r}_{\mathcal{R}}\mathcal{G}$ contains an edge $\mathsf{M}_V \to \Pi_D$	
$\Leftrightarrow$	M_V is $\mathcal{R}$-relevant to Π_D	by Definition 16
$\Leftrightarrow$	$r_D(\mathbf{pa}_{\Pi_D})$ may vary with the value of M_V	by Definition 15
$\Leftrightarrow$	$\mathcal{Q}_D(\mathbf{pa}_{\Pi_D}, \pi_D)$ may vary with the value of M_V	by (Eq. (3.3))
$\Leftrightarrow$	V is a requisite probability node for some $\Pr^{\boldsymbol{\pi}}(\boldsymbol{x} \mid \boldsymbol{y}) \in \mathcal{Q}_D$	by Definition 3
$\Leftrightarrow$	$\mathsf{M}_V \not\perp_{\mathsf{m}_\perp\mathcal{G}} \boldsymbol{X} \mid \boldsymbol{Y}$ for some $\Pr^{\boldsymbol{\pi}}(\boldsymbol{x} \mid \boldsymbol{y}) \in \mathcal{Q}_D$	by Lemma 1

3.4 Conclusions

In this chapter, we have introduced mechanised MAIDs. These allow us to cleanly model the dependencies of decision rules on each other and on the parameters of the game. We then generalised K&M's notion of strategic relevance to $\mathcal{R}$-relevance to enable us to model agents behaving according to many different rationality assumptions, and to determine which other CPDs or decision rules are relevant to an agent when they are deciding on their decision rule.

The contributions of this chapter will serve as a foundation for much of the work in later chapters. In the next chapter, we shall see that $\mathcal{R}$-relevance enables one to identify more subgames in a MAID than in its corresponding EFG; in Chapter 5, we show that the mechanised graph enables us to explicitly reason about different types of imperfect recall and the assumptions behind behavioural, mixed, and correlated policies; and, in Chapter 6, we show how the formalism introduced in this chapter was essential for providing the framework to answer causal queries in games.

4
Subgames and Equilibrium Refinements

Contents

4.1 Introduction

In Chapter 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 chapter, we build on these ideas in four sections. First, we introduce the concept of *subgames* within MAIDs, which, analogously to their EFG counterparts, allow us to analyse and solve parts of the full game independently. Second, we introduce several *equilibrium refinements* for MAIDs. Third, we formalise transformations between MAIDs and EFGs and provide several equivalence results. With these three contributions, we aim to place MAIDs on a more equal footing with EFGs as a tool for game-theoretic analysis. Note that as (structural) causal games are refinements of MAIDs, the results in this section also apply to those models (introduced in Chapter 6). In this chapter's final section, we show how our theoretical contributions result in empirical applications, as we can use the fact that more subgames can be identified from a MAID's graph compared with an EFG graph of the same game to find an NE in a MAID exponentially faster (in the best case).

4.2 Subgames

Subgames in EFGs (defined in Section 2.2.1) represent parts of the game that can be solved independently of 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 4.3); they can reduce the cost of computing equilibria (see Section 4.5); 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.

However, there are two key differences between subgames in EFGs and those in MAIDs. First, because MAIDs explicitly represent conditional independencies between variables, we can often find more subgames in a MAID than in its corresponding EFG. Second, 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 3.3, 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 17. *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 4.4.2). 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 18. *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* $\Pr'(\boldsymbol{v}'; \boldsymbol{\theta}') := \Pr(\boldsymbol{v}' \mid \boldsymbol{z}; \boldsymbol{\theta})$*, where* $\boldsymbol{z}$ *is some instantiation of the variables* $\boldsymbol{Z} = \boldsymbol{V} \setminus \boldsymbol{V}'$*.*[1] *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 4.4.2). Therefore, in the remainder of this chapter, unless otherwise specified, we 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

[1]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.

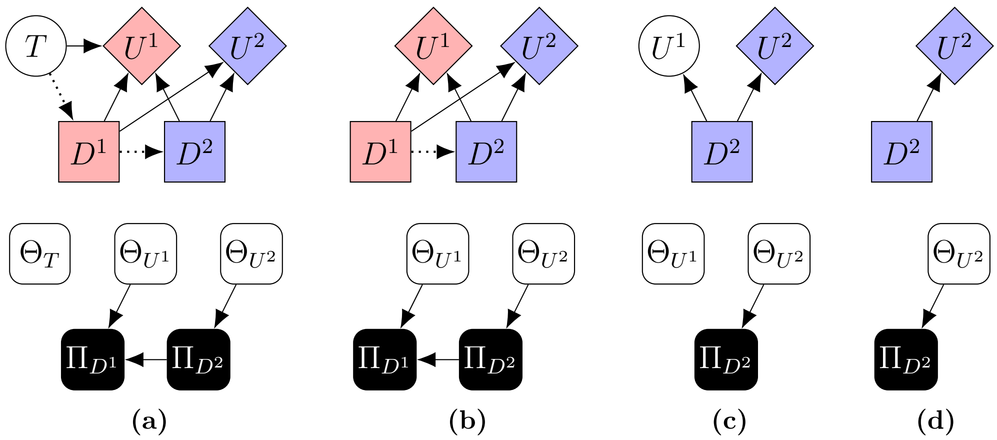


Figure 4.1: (a) A MAID $\mathcal{M} = (\mathcal{G}, \theta)$ for the modified version of Example 1 – in which the firm's profits are also a function of the worker's decision but not of their temperament – and resulting s-relevance graph. The graph $\mathcal{G}$ is also an (improper) s-subdiagram and 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.

(shown in Fig. 2.5) has no proper $\mathcal{R}^{\text{BR}}$-subgames because Π_{D^1} and Π_{D^2} are both $\mathcal{R}^{\text{BR}}$-reachable from one another.

K&M used s-reachability to find decision rules that can be optimised separately from other decision rules in the game [26]. s-subgames are a generalisation of this to find the full sub-structures of the game that are independent (including chance and utility variables). This is crucial for causal reasoning in games (in Chapter 6.

Identifying more Subgames in MAIDs than in EFGs

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 Fig. 4.1a. Note that the *structure* of the EFG (shown in Fig. 2.4) 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 Fig. 4.1b, Fig. 4.1c, and Fig. 4.1d) 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 Fig. 4.1b) and the two values that D^1 can take (for the s-subdiagrams in Fig. 4.1c and Fig. 4.1d).

4.3 Equilibrium Refinements

Concepts in this section will be explained with the help of the following Warehouse Robots example (represented in MAID and 'reduced EFG' form in Fig. 4.2a and Fig. 4.2c, respectively.).

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 when the warehouse is completely tidy 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 Fig. 4.2c 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 explained in the MAID to EFG transformation explained in Section 4.4.1).

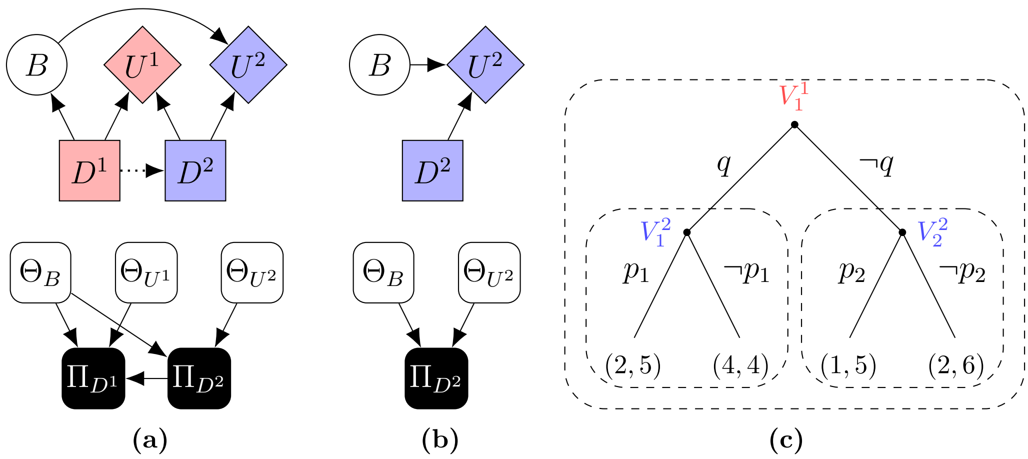


Figure 4.2: (a) A MAID $\mathcal{M} = (\mathcal{G}, \theta)$ representing Example 2 and its s-relevance graph. (b) The smallest proper s-subdiagram $\mathcal{G}'$ of $\mathcal{G}$ and its s-relevance graph. (c) A reduced EFG where each of the proper subgames corresponds to the two possible s-subgames for $\mathcal{G}'$, where $D^1 = q$ or $D^1 = \neg q$ respectively.

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) [80], a policy profile such that no agent can benefit by unilaterally deviating (see Definition 11 in Section 2.2.2). 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.

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 [88] and trembling hand perfect equilibria [89] – within MAIDs. The first rules out 'non-credible' threats, and the second requires that each player is still playing a best response when other players make small mistakes.

Later, in Section 4.4.2, we provide proofs regarding various equivalences between these definitions and those in EFGs.

4.3.1 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 [88, 89]. 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 (we will see what this means later in Definition 27, Chapter 5), 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 Section 4.5, that allows an NE to be computed much more efficiently in a MAID 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 Proposition 12 in Chapter 5 for more on this).

Definition 19. *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.*[2]

Recall the aforementioned $\boldsymbol{\pi}^{\text{NE}}$, in which robot one chooses $D_1 = q$ and robot two chooses $D_2 = p$ whatever the value of D_1. 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 Fig. 4.2b – 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 $\boldsymbol{\pi}^{\text{SPE}}$ 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

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

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 $\mathcal{R}^{\text{BR}}(\mathsf{m}\mathcal{M})$ are simply the NEs of $\mathcal{M}$, as can easily be seen via inspection of Eq. (3.1) 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:

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

expressing that each agent plays a best response for their decision rule in every s-subgame containing that decision. In other words, $\boldsymbol{\pi} \in \mathcal{R}^{\text{SP}}(\mathsf{m}\mathcal{M})$ if $\boldsymbol{\pi}' \in \mathcal{R}^{\text{BR}}(\mathsf{m}\mathcal{M}')$ 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, $\mathcal{R}^{\text{SP}}(\mathsf{m}\mathcal{M})$ 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 3. *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. 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 $1 \leq j < m$, $\mathcal{G}_{j+1}$ 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:

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

for each $\mathcal{M}_1$, i.e., just in case $\pi_D \in r_D^{\text{SP}}(\mathbf{pa}_{\Pi_D})$. This is akin to the result that the $\mathcal{R}^{\text{BR}}$-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 11 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 $\pi_D \in r_D^{\text{SP}}(\mathbf{pa}_{\Pi_D})$. 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 $\pi_D \in r_D^{\text{SP}}(\mathbf{pa}_{\Pi_D})$, i.e., just in case $\boldsymbol{\pi} \in \mathcal{R}^{\text{SP}}(\mathsf{m}\mathcal{M})$. $\square$

4.3.2 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 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* [89]. 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 20. *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).

4.4 Connections to EFGs

EFGs are 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 thesis. With that said, a natural question is whether the game-theoretic definitions for MAIDs introduced in this chapter successfully capture the familiar concepts of their EFG counterparts, or whether we might lose something by working with MAIDs – and hence (Structural) Causal Games (in Chapter 6) – 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 the previous section. Following this, we briefly discuss how tree-based models can also be used for causal reasoning in certain cases.

4.4.1 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 [26, 63]. 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. *This procedure can be made more efficient by marginalising out all variables not in $\boldsymbol{U} \cup \boldsymbol{Fa_D}$, such as in the reduced EFG for Example 2 (shown in Fig. 4.2c), which only has $2^2 = 4$ leaves, as opposed to the $2^3 = 8$ leaves that would have resulted had we retained all the variables in the original MAID (in Fig. 4.2a). 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 PyCID codebase [60], 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 [90].*

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 [26]. In general, there are many MAIDs that can represent a given EFG. For instance, upon converting the EFG representation (in Fig. 2.4) 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 Fig. 2.5, 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.1.2), which maps an EFG to a *unique*, canonical MAID.

Definition 21. *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.

4.4.2 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 4.3) are *preserved* when converting between representations. First, we provide the relevant well-known definitions of an NE, SPE, and THPE in an EFG.

Definition 22. *A strategy profile σ is a **Nash equilibrium (NE)** in an EFG if, for every agent $i \in N$, $\sigma^i \in \arg\max_{\hat{\sigma}^i \in \Sigma^i} \mathbb{E}_{(\hat{\sigma}^i, \sigma^{-i})}\Big[U(\rho[\boldsymbol{L}])[i]\Big]$.*

Definition 23. *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.*

Definition 24. *A perturbation vector η_k contains perturbations $\epsilon_a^{i,j} \in (0,1)$ with $\sum_{a \in A^i_j} \epsilon_a^{i,j} \leq 1$ for every information set I^i_j such that in a perturbed game $\mathcal{E}(\eta_k)$, each strategy is forced to have $\sigma^i_j(a) \geq \epsilon_a^{i,j}$. 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 $\lim_{k \to \infty} \sigma_k = \sigma$.*

Now, we turn to define what it means 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 25. *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 $f^i : \Sigma^i \rightarrow dom(\boldsymbol{\Pi}^i)/\sim$ between their strategies in $\mathcal{E}$ and a partition of their policies in $\mathcal{M}$ (the quotient set of $dom(\boldsymbol{\Pi}^i)$ 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 $f(\sigma) \coloneqq \bigtimes_{i \in N} f^i(\sigma^i)$. We refer to any such f as a* ***natural mapping*** *between $\mathcal{E}$ and $\mathcal{M}$.*

Lemma 2. *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 25. 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. □

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 3. *If* $\mathcal{E} \in$ `maid2efg`$(\mathcal{M})$ *or* $\mathcal{M} =$ `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 without affecting anything important about the game), 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 `maid2efg` and `efg2maid` respectively.

To see this, first suppose we have a MAID, $\mathcal{M}$, and $\mathcal{E}$ is an EFG resulting from `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 `maid2efg` procedure, each feasible decision context is an instantiation of a set of variables which corresponds to one information set I^i_j where $S_V = D$ for all $V \in I^i_j$, 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^i_j 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 $\sigma^i_j(d) \coloneqq \pi_D(d \mid \mathbf{pa}_D)$ 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 $\sigma^i_j(D), \pi_D(D \mid \mathbf{pa}_D) \in \Delta(dom(D))$, 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 $f : \Sigma \to dom(\mathbf{\Pi})/\sim$. 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 the 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 $f : \Sigma \to \mathit{dom}(\mathbf{\Pi})/\sim$. 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. □

The above proof 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. Next, the following result is a direct corollary of Lemma 2 and Lemma 3.

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 3 we have that $\mathcal{E}$ and $\mathcal{M}$ are equivalent. Thus, let $f : \Sigma \to \mathit{dom}(\mathbf{\Pi})/\sim$ be a natural mapping between $\mathcal{E}$ and $\mathcal{M}$ as defined in Definition 25. By applying Lemma 2, we have that σ is an NE in $\mathcal{E}$ if and only if every $\boldsymbol{\pi} \in f(\sigma)$ is an NE in $\mathcal{M}$. □

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 4. *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:

- $\mathsf{M}_Z \rightarrow Z \rightarrow \cdots \boldsymbol{U}^i \cap \mathbf{Desc}_D$. 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 $Z \rightarrow D$ 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.

- $\mathsf{M}_Z \rightarrow Z \leftarrow \cdots \boldsymbol{U}^i \cap \mathbf{Desc}_D$. 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 $\cdots \leftarrow W \rightarrow \cdots$. 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 $\mathcal{E} \in$ `maid2efg`$(\mathcal{M})$ or $\mathcal{M} =$ `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 3. 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:

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

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. □

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 4 and then Lemma 2 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 $\mathcal{E}$.*

Proof. The corollary can be seen to follow immediately from combining the definitions of SPEs in EFGs and MAIDs with Proposition 4 and Lemma Lemma 2. Concretely, if $\mathcal{E} \in$ `maid2efg`$(\mathcal{M})$ or $\mathcal{M} =$ `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 4) 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 2). $\square$

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 3 and Lemma 2 to each of these perturbed games.

Proposition 5. *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 $\mathcal{M}$.*

Proof. Given Lemma 3 and Lemma 2, it suffices to show that the perturbed MAID resulting from `efg2maid`$(\mathcal{E}(\eta_k))$ is equivalent to $\mathcal{E}(\eta_k)$, and similarly that any perturbed EFG resulting from `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 `efg2maid` construction. Clearly for any $\boldsymbol{\pi} \in f(\sigma)$ where f is a natural mapping between $\mathcal{E}$ and $\mathcal{M}$ then

$\sigma^i_j(a) \geq \epsilon^{i,j}_a$ if and only if $\pi_D(d \mid \mathbf{pa}_D) \geq \epsilon^{\mathbf{pa}_D}_d$. 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 $\epsilon^{i,j}_a := \epsilon^{\mathbf{pa}_D}_d$ 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. □

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

4.5 Computing Equilibria

Our open-source Python codebase, PyCID, finds all pure NEs and SPEs in a MAID directly and finds behavioural NEs and SPEs in N-agent games by employing a number of algorithms implemented in PyGambit [90]. We refer the interested reader to Appendix A.2 and our online codebase `https://github.com/causalincentives/pycid` for up-to-date syntax showing how to compute NEs. All SPEs in a MAID are found by adapting Algorithm 6.2 from Koller et al., [26] 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 Proposition 11 in Section 5.4.1. In contrast to K&M's algorithm, our implementation finds almost[3] all SPEs, as opposed to just one.

We now demonstrate the computational efficiency advantages 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 to

[3]At the time of writing, PyCID is guaranteed to compute all pure SPE in any MAID; all NE in MAIDs where the smallest proper subgame contains at most 2 agents; and at least 1 NE in arbitrary N-agent games.

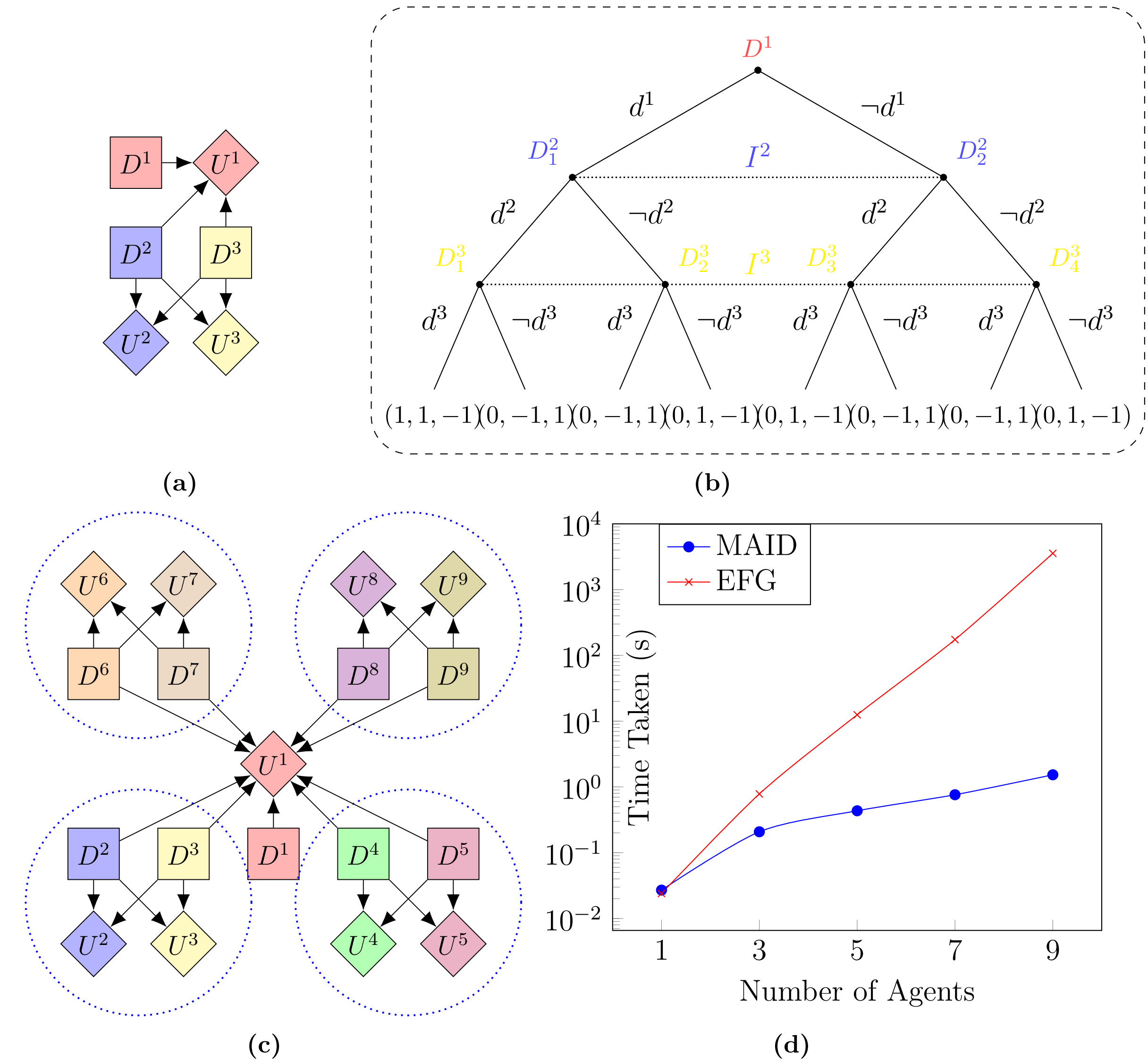


Figure 4.3: The (a) MAID and (b) EFG for the 3-agent version of the matching-pennies-like game described in the text. (c) The graph of the MAID for the 9-agent version of the game. (d) A plot of the time taken to find an NE in the MAID and EFG representations of this game for varying numbers of agents.

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. MAIDs for the 3-agent and 9-agent variants of this game are shown in Fig. 4.3a, Fig. 4.3c, respectively.

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 Fig. 4.3b) has no proper subgames.[4] Since the complexity of finding even an approximate mixed NE is PPAD-complete [91], 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 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. Fig. 4.3d 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.[5]

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 4.4.2) 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.

4.6 Conclusions

The contributions of this chapter have extended previous results on MAIDs by introducing the concept of a subgame in MAIDS and a range of key equilibrium refinements. These contributions built upon the contributions of Chapter 3, namely mechanised graphs and $\mathcal{R}$-relevance. By formalising the transformation process between EFGs and MAIDs and proving several equivalence results, we have

[4]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.

[5]These calculations were performed using PyCID [60] and Gambit [90] on an NVIDIA Tesla K80 GPU and the time taken is the mean of seven runs.

contributed to the underlying theoretical foundations of MAIDs as a game-theoretic model. Finally, we demonstrated that an NE can be computed more efficiently in a MAID than in a corresponding EFG, demonstrating that there are practical applications of this theoretical work.

5
Imperfect Recall

Contents

5.1 Introduction

In previous work, agents in MAIDs are assumed to have perfect (or, at least, 'sufficient') recall [26]. This assumption is often unreasonable. For example, MAIDs must allow imperfect recall to handle bounded rationality, teams with imperfect communication [92], or memoryless policies in Markov games. However, forgetfulness (when an agent forgets an observation or the *outcome* of one of their previous decisions) or absent-mindedness (when an agent cannot even remember whether they have previously made a decision) can prevent the existence of an NE in behavioural policies. To overcome this, one can consider other solution concepts, such as mixed or correlated equilibria.

In this chapter, we focus on imperfect recall in MAIDs. Imperfect recall has already been extensively studied in EFGs [59, 93, 94], but we first show how a MAID's mechanised graph (introduced in Chapter 3) makes graphically explicit the semantic difference between behavioural and mixed policies (hidden in EFGs) and readily identifies forgetful or absent-minded agents (or teams). Next, we prove results giving the conditions required for the existence of various types of equilibria. After this, we show how our insights inspire two definitions of *correlated equilibrium* in MAIDs. The first follows from the normal-form game definition [61]. The second, based on von Stengel and Forges' extensive-form correlated equilibrium [62], is more natural for dynamic settings, can yield greater social welfare, and is easier to compute. Again, mechanised graphs clearly depict the assumptions made in both. Finally, we examine MAIDs from a computational complexity perspective by studying the decision problems of finding a best response, checking whether a policy profile is an NE, and checking whether each type of NE exists. These provide an insight into what makes particular instances hard, when computations can be made tractable, and rigorously identify which problems are suitable for analysis as MAIDs. Our results also apply to refinements of MAIDs, such as the *(structural) causal games* we introduce in Chapter 6. We assume familiarity with EFGs [29], BNs [95], and the complexity classes P, NP, and PP [96].

5.2 Information conditions

Agents may possess different degrees of information about the state of a game. A game has **perfect recall** if each agent remembers all their past moves and observations, and it has **perfect information** if each agent, when making a decision, is perfectly informed of all previous events.

Definition 26. *Agent i in a MAID $\mathcal{M}$ is said to have **perfect recall** [26] if there exists a total ordering $D_1 \prec \cdots \prec D_m$ over $\boldsymbol{D}^i$ such that $(\boldsymbol{Pa}_{D_j} \cup D_j) \subseteq \boldsymbol{Pa}_{D_k}$ for*

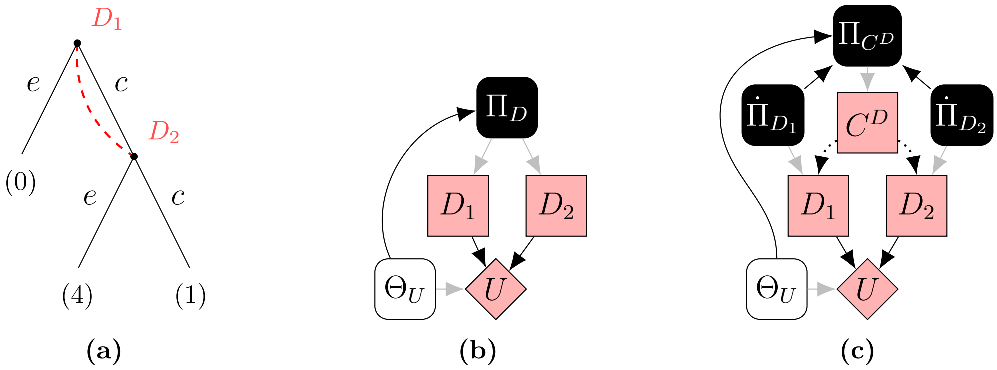


Figure 5.1: The EFG (a) and the mechanised graphs for an absent-minded driver choosing behavioural (b) or mixed (c) policies.

any $1 \leq j < k \leq m$*.* $\mathcal{M}$ *is a perfect recall game if all agents in* $\mathcal{M}$ *have perfect recall.* $\mathcal{M}$ *is a* ***perfect information*** *game if there exists a total ordering* $\prec$ *over* $\boldsymbol{V}$ *such that: (i) if there is a directed path from* V_i *to* V_j*, then* $V_i \prec V_j$*; and (ii) if* $V \prec D$*, then* $V \in \boldsymbol{Pa}_D$*.*

A MAID with perfect information (recall) can be transformed into an EFG with perfect information (recall), and vice versa (Chapter 4, Section 4.4.1). Hence, these information conditions also guarantee the existence of an NE in pure (behavioural) policies in the MAID ([59] gives the equivalent results in EFGs). However, the mechanised representation of a MAID enables weaker criteria to be defined – **sufficient information** and **sufficient recall**. Later, in Proposition 10, we will see that these criteria are 'sufficient' to preserve the NE existence results of perfect information and perfect recall games, respectively.

Definition 27. *Agent* i *in a MAID* $\mathcal{M}$ *has* ***sufficient recall*** *[27] if the subgraph of the mechanised graph* $\mathsf{m}\mathcal{G}$ *restricted to just agent* i*'s decision rule nodes* $\boldsymbol{\Pi}_{\boldsymbol{D}^i}$ *is acyclic.* $\mathcal{M}$ *is a sufficient recall game if all agents in* $\mathcal{M}$ *have sufficient recall.* $\mathcal{M}$ *is a* ***sufficient information*** *game if the subgraph of* $\mathsf{m}\mathcal{G}$ *restricted to contain only and all decision rule nodes* $\boldsymbol{\Pi}_{\boldsymbol{D}}$ *is acyclic.*[1]

5.2.1 Forgetfulness and Absent-Mindedness

Previous work on MAIDs has assumed perfect or sufficient recall. We now begin the contributions of this chapter by distinguishing between two types of imperfect recall in MAIDs. Recall that **Forgetfulness** applies when an agent forgets an observation or the *outcome* of one of their previous decisions. **Absent-mindedness** applies when an agent cannot even remember whether they have previously made a decision. To make this distinction, we leverage the following insight: *mechanism nodes represent the CPDs governing object-level variables. Every edge between a mechanism and object-level node represents an independent draw from the mechanism's distribution.* We now provide formal definitions.

Definition 28. *Agent i has **imperfect recall** in a MAID $\mathcal{M}$ if for every total ordering $D_1 \prec \cdots \prec D_m$ over $\boldsymbol{D}^i$ there exists some $j < k$ such that $(\boldsymbol{Pa}_{D_j} \cup D_j) \not\subseteq \boldsymbol{Pa}_{D_k}$ (i.e., if agent i does not have perfect recall). Agent i is **forgetful** if such a D_j and D_k have distinct decision rules and is **absent-minded** if in $\mathcal{M}$'s mechanised graph, a decision rule node has more than one outgoing edge to a decision node.*

To motivate our definition of absent-mindedness in MAIDs, we revisit Piccione and Rubinstein's absent-minded driver game [93] (its EFG is in Fig. 5.1a). A driver on a highway may take one of two exits. Taking the first, second, or no exit yields a payoff of 0, 4, or 1, respectively. Adopting Aumann [97]'s *modified multi-selves approach* (i.e., that the driver should only be able to control her current action, not her future actions), the driver does not know which junction she is facing, so she must have the same decision rule at both junctions. We make absent-mindedness explicit with a shared decision rule node Π_D for D_1 and D_2 in the mechanised graph (Fig. 5.1b) (note this is consistent with our mechanised graph definition). Π_D's *two outgoing edges now represent two independent draws from the same distribution.* For D_i and D_j to share a decision rule, it is necessary that $dom(D_i) = dom(D_j)$ and $dom(\boldsymbol{Pa}_{D_i}) = dom(\boldsymbol{Pa}_{D_j})$. Note that perfect recall

[1]Note that since previous work on influence diagrams has not modelled absent-mindedness (see our Definition 28 in Section 5.2.1), this definition implicitly assumes each mechanism variable has a single child.

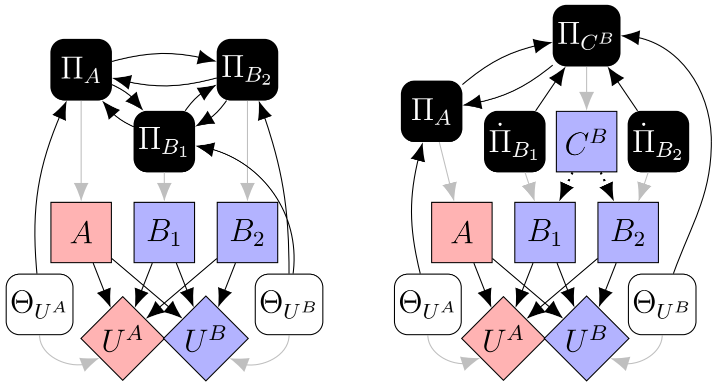
implies that for any two decisions belonging to the same agent, one's set of parents is a strict superset of the other's, so their decision rules have a different type signature, which rules out absent-mindedness.

In the following examples, used just to explain forgetfulness and absent-mindedness, Alice and Bob play variations of matching pennies with the usual payoffs given according to the *final* state of their two coins (where a/b and $\bar{a}/\bar{b}$ represent heads and tails, respectively). Example 3 illustrates a consequence of Bob being forgetful – meaning he cannot remember the *outcome* of his previous decision. In Example 4, Bob is absent-minded – he cannot remember whether he has made a decision at all.

Example 3 (Fig. 5.2a, Fig. 5.2b, and Fig. 5.2c)**.** *Bob is told he must submit a move in advance (B_1) and then confirm it on game day (B_2). If his moves agree, payoffs correspond with normal matching pennies, but if his moves disagree, he must forfeit and always loses (these payoffs are shown in Fig. 5.2c). Bob is forgetful, so on game day he cannot remember his advance choice (i.e., the edge $B_1 \to B_2$ is missing in Fig. 5.2a).*

Example 4 (Fig. 5.2d, Fig. 5.2e, and Fig. 5.2f)**.** *In a new game, the pennies start heads up, and Bob decides whether or not to turn the coin over (B_1). He is absent-minded, so when he sees heads he cannot remember whether he has already made his move, and he decides again (B_2). If he turns the coin having previously chosen to keep heads, Bob gets a -2 penalty and Alice a $+2$ bonus. In all other cases, the payoffs correspond with normal matching pennies (payoffs are shown at the leaves of the EFG in Fig. 5.2e).*

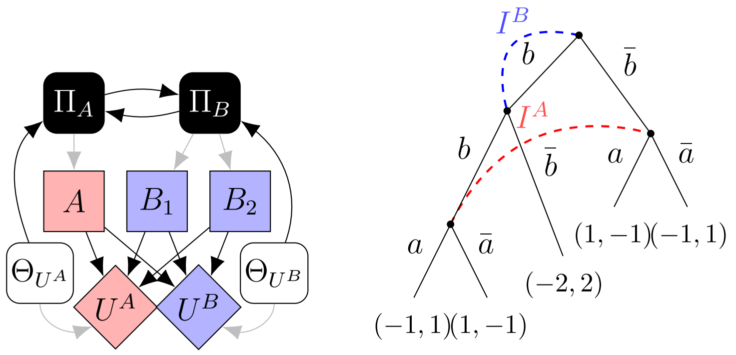
Observe that the MAID's regular graph (just the object-level variables) is identical for both Fig. 5.2a and Fig. 5.2d with the missing $B_1 \to B_2$ edge implying imperfect recall. The difference between forgetfulness and absent-mindedness is only revealed by the mechanised graph. Forgetful Bob has two independent decision rules Π_{B_1} and Π_{B_2} for B_1 and B_2. Absent-minded Bob only has one shared decision rule Π_B.



(a) (b)

	a	$\bar{a}$
b_1b_2	$-1, 1$	$1, -1$
$b_1\bar{b}_2$	$-1, 1$	$-1, 1$
$\bar{b}_1b_2$	$-1, 1$	$-1, 1$
$\bar{b}_1\bar{b}_2$	$1, -1$	$-1, 1$

(c)



(d) (e)

	a	$\bar{a}$
b	$-1, 1$	$1, -1$
$\bar{b}$	$1, -1$	$-1, 1$

(f)

Figure 5.2: The mechanised graphs for forgetful Bob (Example 3) using (a) behavioural or (b) mixed policies, with normal-form in (c). (d) The mechanised graph for absent-minded Bob (Example 4) using a behavioural policy, with EFG and normal-form representations in (e) and (f).

Example 3 and Example 4 demonstrate that both types of imperfect recall can mean an NE in behavioural policies may not exist, even in zero-sum two-agent MAIDs with binary decisions. This arises due to the grand best response function being non-convex valued, which violates a condition of Kakutani's fixed point theorem [98] (an extension of Brouwer's Fixed Point Theorem that accommodates set-valued functions):

Theorem 1 (**Kakutani's Fixed Point Theorem** [98])**.** *Suppose $\mathcal{S}$ is a non-empty, convex, closed, bounded, non-empty subset of $\mathcal{R}^n$, and suppose $f : \mathcal{S} \to 2^{\mathcal{S}}$ is such that $f(s)$ is non-empty and convex, and f has a closed graph. Then f has a fixed point.*

For mixed policies (introduced in the next section), the grand best response function satisfies all conditions of Kakutani's fixed point theorem, which is why an NE is always guaranteed to exist (as this is just a fixed point of the grand best response function) in finite games. These examples show that absent-mindedness and forgetfulness can mean that the restriction to behavioural policies results in a non-convex valued grand best response function.

The normal-form games (in Fig. 5.2c and Fig. 5.2f) show that neither contains an NE in pure policies. We now prove non-existence in behavioural policies. c

Proposition 6. *Both forgetfulness and absent-mindedness can prevent the existence of an NE in behavioural policies.*

Proof. Example 3 (Fig. 5.2a, Fig. 5.2b, and Fig. 5.2c) and Example 4 (Fig. 5.2d, Fig. 5.2e, and Fig. 5.2f) are counterexamples for each case.

Proof for Example 3 (forgetfulness): The normal-form game showing the payoffs for each agent is shown in Figure Fig. 5.2c. First, observe that there are no NE in pure policies. Now, suppose that there does exist an NE in behavioural policies. If Alice always plays a or always $\bar{a}$ – i.e., $\pi^A(a) = 1$ or $\pi^A(a) = 0$ – then Bob's best response is always $\bar{b}_1\bar{b}_2$ or always b_1b_2, respectively. However, this does not form an NE. So, Alice must select a stochastic decision rule π_A and be indifferent (by the principle of indifference) between a and $\bar{a}$.

Letting Π_{B_1} and Π_{B_2} be parameterised by $p, q \in [0, 1]$ where $\pi_{B_1}(b_1) = p$ and $\pi_{B_2}(b_2) = q$, we obtain two constraints on p and q. On the one hand, by virtue of Alice's indifference, Bob's behavioural policy $\boldsymbol{\pi}^B$ must result in $\pi^B(\neg b_1, \neg b_2) = \pi^B(b_1, b_2)$, and so: $(1-p)(1-q) = pq \implies p + q = 1$. On the other hand, Bob receives utility -1 if his policy $\boldsymbol{\pi}^B$ results in any outcome with $B_1 = \neg b_1$ and $B_2 = b_2$, or $B_1 = b_1$ and $B_2 = \neg b_2$, whatever the choice of $\boldsymbol{\pi}^A$. Therefore, we must have that $\pi^B(\neg b_1, b_2) + \pi^B(b_1, \neg b_2) < \pi^B(b_1, b_2) + \pi^2(\neg b_1, \neg b_2)$ and thus, by substituting in the result that $p+q = 1$: $(1-p)q+p(1-q) < pq+(1-p)(1-q) \implies (2p-1)^2 < 0.$. This contradiction implies that the MAID for Example 3 has no NE in behavioural policies.

To further understand what this example, let us again write Bob's policy as a tuple (p, q), and suppose $\pi_A(a) = 0.5$. Then, either pure policy $(1, 1)$ and $(0, 0)$ is a best response for Bob with $\mathbb{E}_{\boldsymbol{\pi}}[U^B] = 0$. But, consider the convex combination of these best responses $0.5 \cdot (1, 1) + 0.5 \cdot (0, 0) = (0.5, 0.5)$. Under this policy, each of the eight outcomes in the payoff matrix is equally likely and so Bob's expected payoff drops to $(-1-1-1+1+1-1-1-1)/8 = -0.5$. Since a convex combination of best responses is no longer a best response, Bob's best response function is not convex-valued, and so nor is the grand best response function. The conditions of Kakutani's fixed point theorem are not satisfied, which explains why a Nash equilibrium need not exist.

Proof for Example 4 (absent-mindedness): First, observe from the normal-form game in Fig. 5.2f that there is no NE in pure policies in this game. Next, suppose there exists a NE in behavioural policies and let Π_B be parameterised by $p \in [0, 1]$, where $\pi_B(b) = p$ for $p \in [0, 1]$. Alice's payoff only depends on her policy π^A when Bob plays bb or $\bar{b}\bar{b}$, for which Alice has pure best responses. This implies that at an NE, $p^2 = (1 - p)^2 \implies p = 0.5$. Therefore, Alice's policy is irrelevant and $\mathbb{E}_{\boldsymbol{\pi}}[U^B] = -1$ ($\mathbb{E}_{\boldsymbol{\pi}}[U^B] = 0$) if he does (doesn't) forfeit, which happens with probability 0.5. Therefore, Bob's policy is dominated by his pure policies, with worst-case payoff $\mathbb{E}_{\boldsymbol{\pi}}[U^B] = -1$. This contradicts the assumption of an NE in behavioural policies.

Explanation: If $\pi_A(a) = 0.5$, then $p = 0$ and $p' = 1$ are both best responses for Bob with $\mathbb{E}_{\boldsymbol{\pi}}[U^B] = 0$. However, the convex combination $0.5p + 0.5p'$ gives expected payoff to Bob $\mathbb{E}_{\boldsymbol{\pi}}[U^B] = 0.25 \cdot 1 + 0.25 \cdot (-1) + 0.5 \cdot (-10) = -5$ and is therefore not a best response. Again this is due to the fact that under behavioural policies, in situations of imperfect recall, a convex combination of pure policies can introduce outcomes that could not occur under either pure policy. Under a mixed combination of pure policies, Alice will always follow one or the other, and so no new outcomes are introduced. However, under a behavioural combination, two independent absent-minded draws from the same distribution over actions can come out differently, introducing new potential outcomes—in this case forfeit. $\square$

5.3 Mixed Policies and Behavioural Mixtures

To overcome the fact that a behavioural policy NE may not exist in imperfect recall MAIDs, one can use mixed or correlated policies. These ensure that the grand best response function always satisfies the conditions of Kakutani's fixed point theorem, so a fixed point of the grand best response function (and therefore a Nash equilibrium) always exists. In this section, we show how the assumptions behind behavioural, mixed, and behavioural mixtures policies (well-studied in EFGs [62, 99], but unexplored in MAIDs), are made graphically explicit in mechanised graphs and prove several important theoretical results. In the next section, we will turn to correlated equilibria.

For convenience, for the remainder of this chapter we will denote the set of Agent i's behavioural policies as $\boldsymbol{P}^i := dom(\boldsymbol{\Pi}^i)$, sets of Agent i's pure policies as $\dot{\boldsymbol{P}}^i$, and (pure) policy profiles by $\boldsymbol{P}$ $(\dot{\boldsymbol{P}})$.

Behavioural policies allow agents to randomise independently at every decision node. By contrast, a **mixed policy** $\mu^i \in \Delta(\dot{\boldsymbol{P}}^i)$ is a distribution over pure policies. It allows an agent to coordinate their choice of decision rules at different decisions by randomising once at the game's outset and then committing to the assigned pure policy. More generally, **behavioural mixtures** in $\Delta(\boldsymbol{P}^i)$ are distributions over all behavioural policies. They allow agents to randomise *both* at the outset of the game and before each decision. The outcome of the first randomisation determines the distributions for the others.

A behavioural mixture changes the specification of the game because it can require a correlation between different decision rules. At the object-level, a behavioural mixture for Agent i requires a new (correlation) decision variable C^i with $\mathbf{Pa}_{C^i} = \varnothing$, $\mathbf{Ch}_{C^i} = \boldsymbol{D}^i$, and $dom(C^i) = \boldsymbol{P}^i$ (the set of all behavioural policies). The decision rules for each D^i become conditional on C^i, so each value of C^i determines a behavioural policy. This explains why C^i and still every $D \in \boldsymbol{D}^i$ are decision nodes – the agent chooses the CPDs for both. Even in the mixed policy case, where each D^i depends deterministically on C^i, the agent chooses the dependence independently from choosing the distribution over C^i. In the mechanised graph (see Fig. 5.1c), C^i gets an associated mechanism variable Π_{C^i} for the distribution C^i is drawing from (its mechanism parents are again determined by s-reachability).

In EFGs, the mechanism by which agents decide on their decision rules is not explicitly shown. Mechanised graphs, however, show clearly when an agent chooses to randomise. Behavioural and mixed policies are the limiting cases of behavioural mixtures: the former where the distribution over $\boldsymbol{P}^i$ is deterministic; the latter where the decision rules $\boldsymbol{\Pi}_{\boldsymbol{D}^i}$ are deterministic. The difference between forgetful Bob in Example 3 using a behavioural or mixed policy is shown in Fig. 5.2a and Fig. 5.2b. For Bob's behavioural policy, C^B and Π_{C^B} are omitted as the decision rules Π_{B_1} and Π_{B_2} are independent. This leaves a normal mechanised graph. Whereas, if Bob uses a mixed policy, he only randomises once from Π_{C^B} at the start of the game to select a pure policy at C^B. This fixes deterministic decision rules at $\dot{\Pi}_{B_1}$ and $\dot{\Pi}_{B_2}$.

Proposition 7. *Given a MAID $\mathcal{M}$ with any partial profile $\boldsymbol{\pi}^{-i}$ for agents $-i$, then if agent i is not absent-minded, for any behavioural policy $\boldsymbol{\pi}^i$ there exists a pure policy $\dot{\boldsymbol{\pi}}^i$ which yields a payoff at least as high against $\boldsymbol{\pi}^{-i}$. On the other hand, if agent i is absent-minded in $\mathcal{M}$ across a pair of decisions with descendants in $\boldsymbol{U}^i$, then there exists a parameterisation of $\mathcal{M}$ and a behavioural policy $\boldsymbol{\pi}^i$ which yields a payoff strictly higher than any payoff achievable by a pure policy.*

Proof. Let $\boldsymbol{\pi}^i$ be a behavioural policy and begin with any decision node $D \in \boldsymbol{D}^i$ with decision rule $\pi_D \in \boldsymbol{\pi}^i$. Now $\pi^i_D(d \mid \mathbf{pa}_D)$ is the probability of choosing $d \in dom(D)$ at D when $\mathbf{Pa}_D = \mathbf{pa}_D$ according to $\boldsymbol{\pi}^i$. Since agent i is not absent-minded, the expected payoff for agent i can be written $\sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{(\boldsymbol{\pi}^i, \boldsymbol{\pi}^{-i})}[U] = \sum_{d \in dom(D)} \pi^i(d \mid \mathbf{pa}_D)\lambda_d + \nu$, where each coefficent λ_d and ν are independent of $\pi^i_D(d \mid \mathbf{pa}_D)$. Consider the action $\hat{d} \in dom(D)$ which achieves the highest λ_d (i.e., contributes most the expected utility) Setting $\pi^i_D(\hat{d} \mid \mathbf{pa}_D) = 1$, therefore, yields a payoff at least as high. The first claim, therefore, follows by repeating this argument for every $D \in \boldsymbol{D}^i$.

For the converse claim, agent i is absent-minded, which means that at least two of agent i's decision nodes must draw from an identical distribution. Without loss of generality, call these D_l and D_m. Recall that for this to be the case,

$dom(D_l) = dom(D_m)$ and $dom(\boldsymbol{Pa}_{D_l}) = dom(\boldsymbol{Pa}_{D_m})$. Now consider an outcome of the game $\hat{\boldsymbol{v}} \in dom(\boldsymbol{V})$ where $\mathbf{pa}_{D_l} = \mathbf{pa}_{D_m}$, but $d_l \neq d_m$. Since D_l and D_m have descendants in $\boldsymbol{U}^i$, Parameterise the MAID $\mathcal{M}$ such that agent i's expected utility is 1 if and only if $\boldsymbol{V} = \hat{\boldsymbol{v}}$. For all other game outcomes $\boldsymbol{v} \neq \hat{\boldsymbol{v}}$, let agent i's expected utility be zero. The claim follows since the outcome $\hat{\boldsymbol{v}}$ cannot be instantiated by any pure policy for agent i, but can be instantiated by any behavioural policy for agent i that has a (shared) decision rule for D_l and D_m that assigns a positive probability to both actions d_l and d_m. $\square$

Proposition 7 says that a non-absent-minded agent cannot achieve more expected utility by using a behavioural rather than a pure (or mixed) policy, but an absent-minded agent often can. Consider Fig. 5.1c, where $dom(C^D) = \dot{\boldsymbol{P}}^D$, the set of all the driver's pure policies. Π_{C^D} represents the distribution over $dom(C^D)$, so D_1 and D_2 must both be e or both be c. Therefore, the driver's expected utility is less than or equal to 1 under any mixed policy. Whereas, under the behavioural policy $\pi^1_D(e) = \frac{1}{3}$, their expected utility is $\frac{4}{3}$.

This highlights an important difference between absent-mindedness and forgetfulness. In the next section, we give a result (Lemma 4) that says that under perfect recall, every mixed policy has an equivalent behavioural policy, in the sense of inducing the same distribution over outcomes against every partial policy profile (policies for the other agents). Under forgetfulness, whilst a mixed policy might not have an equivalent behavioural policy, a behavioural policy always has an equivalent mixed policy [59], so there must exist a pure policy which performs just as well. On the other hand, under absent-mindedness, neither mixed nor behavioural policies are guaranteed to have an equivalent of the other type, so there can be a behavioural policy which outperforms every mixed policy against a given policy profile.

5.4 Conditions for NE existence

We introduce mixed policies (and behavioural mixtures) to MAIDs to allow more generality in modelling when agents randomise and to guarantee the existence of an

NE. However, a mixed policy can require exponentially more parameters $\mathcal{O}(2^{2^{|\boldsymbol{V}|}})$ than a behavioural policy $\mathcal{O}(2^{|\boldsymbol{V}|})$ to define. Moreover, single agents are often more naturally modelled as randomising once they encounter decision points [59] (this changes for team situations described in Section 5.7). It is, therefore, important to know when the existence of different types of NE is guaranteed.

As a first step, we prove that if all agents in the MAID have perfect recall (i.e., agents never forget their past moves nor any of the information they knew previously, as introduced in Definition 26), then an NE in behavioural policies is guaranteed to exist. First, we prove a lemma that says when an agent has perfect recall, then for every mixed policy of theirs, there exists a behavioural policy that yields the same probability distribution over outcomes of the game.

Lemma 4. *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:

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

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 $1, \dots, m$ 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:

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

where:

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

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:

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

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

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

But the first antecedent above also implies that $\mu^i(\boldsymbol{d}^i \mid \mathbf{pa}_{\boldsymbol{D}^i}) \coloneqq \sum_{\dot{\boldsymbol{\pi}}^i \,:\, \dot{\boldsymbol{\pi}}^i(\mathbf{pa}_{\boldsymbol{D}^i}) = \boldsymbol{d}^i} \mu^i(\dot{\boldsymbol{\pi}}^i) = 0$, 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:

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

$\square$

Proposition 8. *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 4 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 [80], we know that every MAID is guaranteed to have an NE μ in mixed policies. Given $\mu = (\mu^1, \dots, \mu^n)$, we can therefore apply Lemma 4 n times to result in a behavioural policy $\boldsymbol{\pi} = (\boldsymbol{\pi}^1, \dots, \boldsymbol{\pi}^n)$ 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 $\hat{\mu}^i \in \Delta(\bigtimes_{D \in \boldsymbol{D}^i} dom(\dot{\Pi}_D))$ that results in the same joint distribution as $\hat{\boldsymbol{\pi}}^i$ in the following manner. We define $\hat{\mu}^i(\dot{\boldsymbol{\pi}}^i) \coloneqq \prod_{j=1}^m \hat{\mu}^i(\dot{\pi}^i_{D_j})$ where we set:

$$\hat{\mu}^i(\dot{\pi}^i_{D_j}) \coloneqq \prod_{\mathbf{pa}'_{D_j} \in dom(\mathbf{Pa}_{D_j})} \hat{\pi}^i_j\Big(\dot{\pi}^i_{D_j}(\mathbf{pa}'_{D_j}) \mid \mathbf{pa}'_{D_j}\Big).$$

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 $\hat{\pi}^i_j\Big(\dot{\pi}^i_{D_j}(\mathbf{pa}'_{D_j}) \mid \mathbf{pa}'_{D_j}\Big) \geq 0$, and we have:

$$\begin{aligned}
\sum_{\dot{\boldsymbol{\pi}}^i} \hat{\mu}^i(\dot{\boldsymbol{\pi}}^i) &= \sum_{\dot{\pi}^i_{D_1}} \cdots \sum_{\dot{\pi}^i_{D_m}} \prod_{j=1}^m \hat{\mu}^i(\dot{\pi}^i_{D_j}) \\
&= \prod_{j=1}^m \sum_{\dot{\pi}^i_{D_j}} \hat{\mu}^i(\dot{\pi}^i_{D_j}) \\
&= \prod_{j=1}^m \sum_{\dot{\pi}^i_{D_j}} \left(\prod_{\mathbf{pa}'_{D_j} \in dom(\mathbf{Pa}_{D_j})} \hat{\pi}^i_j\Big(\dot{\pi}^i_{D_j}(\mathbf{pa}'_{D_j}) \mid \mathbf{pa}'_{D_j}\Big) \right) \\
&= \prod_{j=1}^m \left(\prod_{\mathbf{pa}'_{D_j} \in dom(\mathbf{Pa}_{D_j})} \left(\sum_{d'_j \in dom(D_j)} \left(\sum_{\dot{\pi}^i_{D_j} \,:\, \dot{\pi}^i_{D_j}(\mathbf{pa}'_{D_j}) = d'_j} \hat{\pi}^i_j\Big(\dot{\pi}^i_{D_j}(\mathbf{pa}'_{D_j}) \mid \mathbf{pa}'_{D_j}\Big) \right) \right) \right) \\
&= \prod_{j=1}^m \left(\prod_{\mathbf{pa}'_{D_j} \in dom(\mathbf{Pa}_{D_j})} 1 \right) \\
&= 1.
\end{aligned}$$

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 $j \in \{1, \dots, m\}$:

$$
\begin{aligned}
\hat{\mu}^i_j(d_j \mid \mathbf{pa}_{D_j}) &:= \sum_{\dot{\pi}^i_{D_j} \,:\, \dot{\pi}^i_{D_j}(\mathbf{pa}_{D_j}) = d_j} \hat{\mu}^i(\dot{\pi}^i_{D_j}) \\
&= \sum_{\dot{\pi}^i_{D_j} \,:\, \dot{\pi}^i_{D_j}(\mathbf{pa}_{D_j}) = d_j} \left(\prod_{\mathbf{pa}'_{D_j} \in dom(\mathbf{Pa}_{D_j})} \hat{\pi}^i_j\Big(\dot{\pi}^i_{D_j}(\mathbf{pa}'_{D_j}) \mid \mathbf{pa}'_{D_j}\Big) \right) \\
&= \sum_{\dot{\pi}^i_{D_j} \,:\, \dot{\pi}^i_{D_j}(\mathbf{pa}_{D_j}) = d_j} \hat{\pi}^i_j(d_j \mid \mathbf{pa}_{D_j}) \left(\prod_{\mathbf{pa}'_{D_j} \in dom(\mathbf{Pa}_{D_j}) \setminus \{\mathbf{pa}_{D_j}\}} \hat{\pi}^i_j\Big(\dot{\pi}^i_{D_j}(\mathbf{pa}'_{D_j}) \mid \mathbf{pa}'_{D_j}\Big) \right) \\
&= \hat{\pi}^i_j(d_j \mid \mathbf{pa}_{D_j}) \cdot \prod_{\mathbf{pa}'_{D_j} \in dom(\mathbf{Pa}_{D_j}) \setminus \{\mathbf{pa}_{D_j}\}} \left(\sum_{\dot{\pi}^i_{D_j} \,:\, \dot{\pi}^i_{D_j}(\mathbf{pa}_{D_j}) = d_j} \hat{\pi}^i_j\Big(\dot{\pi}^i_{D_j}(\mathbf{pa}'_{D_j}) \mid \mathbf{pa}'_{D_j}\Big) \right) \\
&= \hat{\pi}^i_j(d_j \mid \mathbf{pa}_{D_j}) \cdot \prod_{\mathbf{pa}'_{D_j} \in dom(\mathbf{Pa}_{D_j}) \setminus \{\mathbf{pa}_{D_j}\}} 1 \\
&= \hat{\pi}^i_j(d_j \mid \mathbf{pa}_{D_j}).
\end{aligned}
$$

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 4 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. □

Going further than Proposition 8, because MAIDs reveal conditional independencies between variables, actually only the weaker criterion of *sufficient recall* (Definition 27) is sufficient for the existence of an NE in behavioural policies. First, we prove that sufficient recall and perfect recall are not identical concepts.

Proposition 9. *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 a counterexample that gives a MAID $\mathcal{M}$ (with mechanised MAID in Fig. 5.3a) where an agent has sufficient but imperfect recall and where there exists a mixed policy profile with no behavioural policy resulting in the same distribution over outcomes.

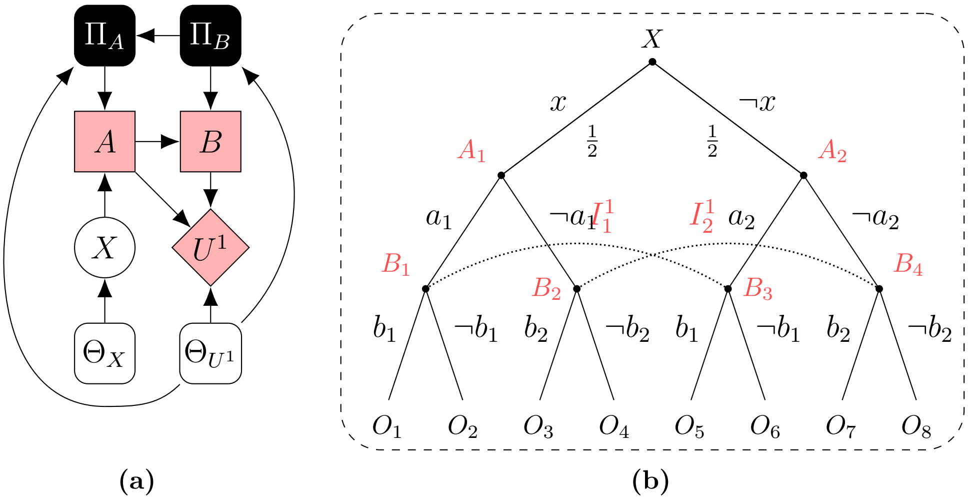


Figure 5.3: (a) An s-minimal mechanised MAID showing that when an agent has sufficient but imperfect recall, then there can exist a mixed policy with no equivalent behavioural policy resulting in the same distribution over outcomes. (b) An equivalent EFG representing the same game.

□

First, we assume that all variables are binary, and further that $\Pr(x) = \frac{1}{2}$. To aid our reasoning, Fig. 5.3b 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 $\{O_1, \dots, O_8\}$, e.g., O_3 represents $(x, \neg a, b, u_3^1)$. We identify each pure policy with the decisions chosen, for example, $\dot{\boldsymbol{\pi}}^1 = a_1 a_2 b_1 \neg b_2$ 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:

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

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

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 $\{O_1, \dots O_8\}$ as μ^i, and so the following equalities must hold:

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

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

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

(O_4) $\frac{1}{2} \cdot (1 - p) \cdot (1 - s) = 0$ $\qquad$ (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 $\mathcal{M}$.

5.4.1 Sufficient NE existence conditions

Proposition 10 gives the sufficient conditions needed for each type of NE. Since both sufficient recall and sufficient information (Definition 27) can be checked in poly-time[2], they expand the class of games that have simple NEs beyond those

[2]The mechanised graph is constructed using s-reachability, which uses the poly-time graphical criterion d-separation [66].

identifiable using an EFG. For example, we can check in poly-time that the MAID for Example 1 (in Fig. 2.5) is an imperfect but sufficient information game, and hence we know that there must exist an NE in pure policies.

Proposition 10. *A MAID with sufficient information always has an NE in pure policies, a MAID with sufficient recall always has an NE in behavioural policies, and every MAID has an NE in mixed policies.*

Proof. The mixed policies case follows from Nash's theorem since all the finite number of random variables in a MAID have finite domains [80].

K&M implicitly prove the sufficient recall case in Koller et al., [26, Theorem 6.2] when proving that their Algorithm 6.2 for finding an NE always converges under certain conditions (as these conditions correspond exactly with those of sufficient recall).

We are therefore left with proving the sufficient information case, where we show that a NE in pure policies must exist. Begin with an arbitrary policy profile across all decision nodes in the original MAID, $\mathcal{M}$. Decision rules associated with each $D \in \boldsymbol{D}$ can be optimised by iterating backwards through a subdiagram ordering $\mathcal{G}_1 \prec \cdots \prec \mathcal{G}_m$ of $\mathcal{M}$'s subdiagrams such that $\mathcal{G}_j \prec \mathcal{G}_k$ implies that $\mathcal{G}_j$ is *not* a subdiagram of $\mathcal{G}_k$. When $\mathcal{M}$ is a sufficient information game, this means that $\mathcal{G}_m$ contains just one decision node for some agent $i \in N$, and, for each subdiagram $\mathcal{G}_j$ where $1 \leq j < m$, $\mathcal{G}_{j-1}$ contains *at most* one additional decision variable. Several subdiagrams can have the same set of decisions, $\boldsymbol{D}_k$, so we choose a single subdiagram $\mathcal{G}_k$ (one with the fewest nodes $\boldsymbol{V}'$) for each $\boldsymbol{D}_k$ and discard the others. Each subdiagram in this ordering has an associated subgame for each setting of the nodes which have a child in $\boldsymbol{V}'$.

When considering each subgame $\mathcal{M}_{m-j}$ for $\mathcal{G}_{m-j}$, the decision rules for all decision nodes in proper subgames of $\mathcal{M}_{m-j}$ will have already been optimised and fixed in previous iterations, so these are now chance nodes in $\mathcal{M}_{m-j}$. In addition, the decision node D_{m-j} in $\mathcal{M}_{m-j}$ does not strategically rely on any of the decision nodes outside of $\mathcal{M}_{m-j}$. Therefore, this step is localised to computing only the optimal decision rule for D_{m-j}. Since this is a single-agent single-decision optimisation, we know that there must exist a pure decision rule best response. In the case of a tie, pick one arbitrarily. After repeating this optimisation process for all subgames in the MAID, we know that every decision node must have a pure decision rule, so we have found a NE in pure policies, as required. □

What about the conditions for the existence of an SPE? Proposition 11 proves that when the MAID has sufficient recall, then there is at least one SPE in behavioural policies. However, Proposition 12 gives a negative result for the case where even just one agent has insufficient recall. This negative result holds even if we permit mixed policies.

Proposition 11. *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 $1 \leq j < m$, $\mathcal{G}_{j+1}$ 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 [26] 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. □

Proposition 12. *If one agent in a MAID has insufficient recall, then an SPE may not exist.*

Proof. This proceeds via the use of a counterexample.

First, we adapt an example from Wichardt et al., [100] to devise a MAID with one NE in mixed policies, but no NE in behavioural policies. We will then use this as a subgame of our full counterexample.

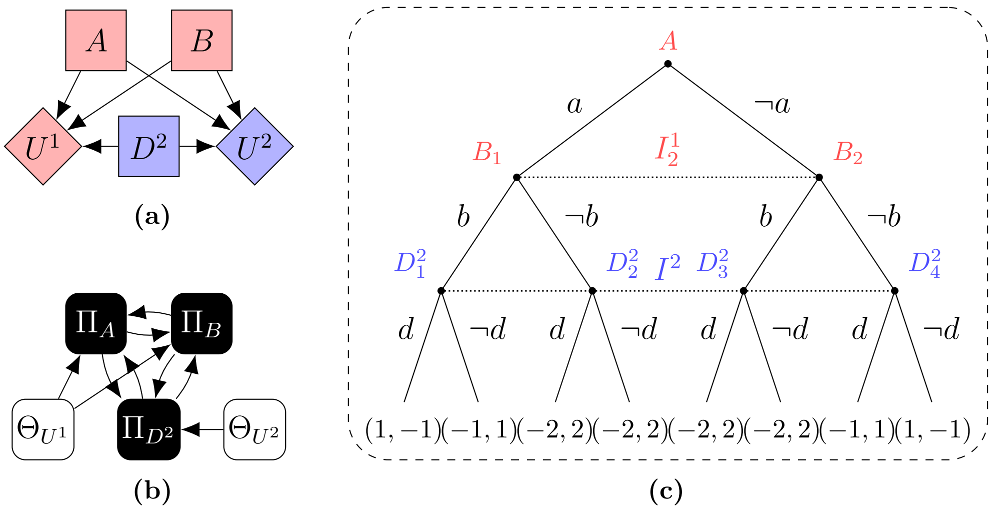


Figure 5.4: (a) A MAID representing the NE non-existence example and (b) its s-relevance graph. (c) An equivalent EFG representing the same game.

The example is shown as both a MAID $\mathcal{M}$ and an EFG $\mathcal{E}$ in Fig. 5.4, in which Agent 1 has two binary decision variables (A and B), Agent 2 has one binary decision variable (D), and the 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 is 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 $\pi^1_A(a) = p$ and $\pi^1_B(b) = q$. 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 $\pi^2_{D^2}$ 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 [80], 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^1_r, which we define as ab with probability r and $\neg a \neg b$ with probability $1 - r$, then $\mathbb{E}_{(\mu^1_r, d)}[U^2] = \mathbb{E}_{(\mu^1_r, \neg d)}[U^2]$ implies $r = \frac{1}{2}$. Furthermore, assuming that agent 2 plays mixed policy μ^2_s, which we define as d with probability s and $\neg d$ with probability $1 - s$, then $\mathbb{E}_{(ab, \mu^2_s)}[U^1] = \mathbb{E}_{(\neg a \neg b, \mu^2_s)}[U^1]$ implies $s = \frac{1}{2}$. Thus, the policy profile $\mu = (\mu^1_{\frac{1}{2}}, \mu^2_{\frac{1}{2}})$ is an NE of $\mathcal{M}$ in mixed policies.

We now extend this MAID to construct a MAID which 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.

Fig. 5.5a shows the graph $\mathcal{G}$ of the MAID $\mathcal{M}$ for our example and Fig. 5.5d 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 Fig. 5.5b, Fig. 5.5c, and Fig. 5.5e 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 Fig. 5.4c). 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 $D^3 = 0$.

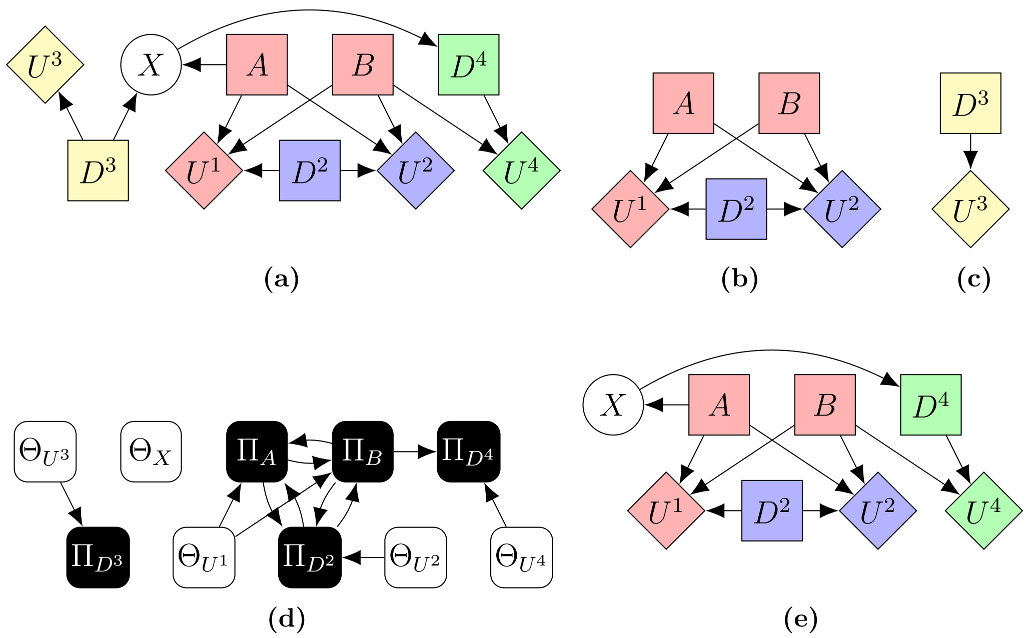


Figure 5.5: (a) A MAID $\mathcal{M}$ representing the SPE non-existence example. (b), (c), and (e) show the three proper s-subdiagrams of this MAID, $\mathcal{G}_1$, $\mathcal{G}_2$, and $\mathcal{G}_3$, respectively. (d) The s-relevance graph of $\mathcal{M}$.

From above, we know that any s-subgame defined over $\mathcal{G}_1$ has no behavioural NEs, and that the only mixed NE is given by $\mu_1 = (\mu^1_{\frac{1}{2}}, \mu^2_{\frac{1}{2}})$ where $\mu^1_{\frac{1}{2}}$ plays the pure strategies ab and $\neg a \neg b$ each with probability $\frac{1}{2}$, and $\mu^2_{\frac{1}{2}}$ 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 $\mu^1 = \mu^1_{\frac{1}{2}}$ and $\mu^2 = \mu^2_{\frac{1}{2}}$.

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. In Section 5.3, we showed that 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 $\mathbf{Fa}_{D^4} = \{X, D^4\}$.

□

5.5 Correlated Equilibria

In Section 5.3, we showed how mechanised graphs can explicitly represent the assumption behind mixed policies: a *single* agent uses a source of randomness to correlate their decision rules. In this section, we do the same for when *multiple* agents can use the same source of randomness, so the choice of pure policy made by each agent may be correlated. An equilibrium in such a game is called a *correlated equilibrium (CE)* [61], which is a distribution κ over the set of all pure policy profiles, i.e., $\kappa \in \Delta(\dot{\boldsymbol{P}})$. A mediator samples $\dot{\boldsymbol{\pi}}$ according to κ, then recommends to each agent i the pure policy $\dot{\boldsymbol{\pi}}^i$. The distribution κ is a CE if no agent, given their information, has an incentive to unilaterally deviate from their recommended policy $\dot{\boldsymbol{\pi}}^i$.

Definition 29. *In a MAID, $\kappa \in \Delta(\dot{\boldsymbol{P}})$ is a* ***correlated equilibrium (CE)*** *if and only if $\forall i$, $\forall \dot{\boldsymbol{\pi}}^i, \dot{\boldsymbol{\varpi}}^i \in \dot{\boldsymbol{P}}^i$:*

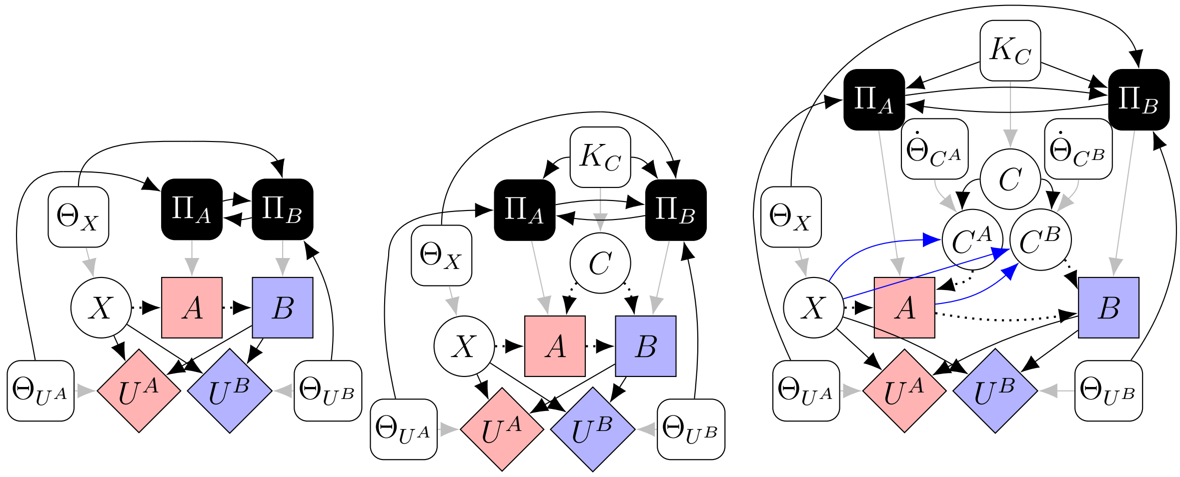
$$\sum_{\dot{\boldsymbol{\pi}}^{-i} \in \dot{\boldsymbol{P}}^{-i}} \kappa(\dot{\boldsymbol{\pi}}^i, \dot{\boldsymbol{\pi}}^{-i}) \sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{(\dot{\boldsymbol{\pi}}^i, \dot{\boldsymbol{\pi}}^{-i})}[U] \geq \sum_{\dot{\boldsymbol{\pi}}^{-i} \in \dot{\boldsymbol{P}}^{-i}} \kappa(\dot{\boldsymbol{\pi}}^i, \dot{\boldsymbol{\pi}}^{-i}) \sum_{U \in \boldsymbol{U}^i} \mathbb{E}_{(\dot{\boldsymbol{\pi}}^{-i}, \dot{\boldsymbol{\varpi}}^i)}[U]$$

	$b_a b_{\bar{a}}$	$b_a \bar{b}_{\bar{a}}$	$\bar{b}_a b_{\bar{a}}$	$\bar{b}_a \bar{b}_{\bar{a}}$
$a_x a_{\bar{x}}$	$5, 5$	$5, 5$	$0, 6$	$0, 6$
$a_x \bar{a}_{\bar{x}}$	$5, 5$	$2, 8$	$3, 3$	$0, 6$
$\bar{a}_x a_{\bar{x}}$	$5, 5$	$3, 3$	$2, 8$	$0, 6$
$\bar{a}_x \bar{a}_{\bar{x}}$	$5, 5$	$0, 6$	$5, 5$	$0, 6$

(a)

	$b_a b_{\bar{a}}$	$b_a \bar{b}_{\bar{a}}$	$\bar{b}_a b_{\bar{a}}$	$\bar{b}_a \bar{b}_{\bar{a}}$
$a_x a_{\bar{x}}$	α_1	α_2	α_3	α_4
$a_x \bar{a}_{\bar{x}}$	β_1	β_2	β_3	β_4
$\bar{a}_x a_{\bar{x}}$	γ_1	γ_2	γ_3	γ_4
$\bar{a}_x \bar{a}_{\bar{x}}$	δ_1	δ_2	δ_3	δ_4

(b)



(c) **(d)** **(e)**

Figure 5.6: (a) and (b) give the expected payoff for each agent under each pure policy profile and the parameterisation of the distribution κ, respectively. The mechanised graph for Example 5's original MAID is shown in (c), and the mechanised graphs for when a trusted mediator gives public or private recommendations to find a CE are shown in (d) and (e), respectively. The blue edges are added to the graph in (e) for a MAID-CE's staggered recommendations.

We illustrate how MAIDs and their mechanised graphs make explicit the assumptions used for a CE using a costless-signal variation of Spence's Job Market game [28].

Example 5. *Alice is hardworking or lazy (X) with equal probability. She applies for a job with Bob by deciding which costless signal (A) to send. Bob can distinguish between the signals but does not know Alice's true temperament. He decides whether to offer the job (B) to Alice. The utility functions for Alice and Bob are* $U^A = (6 - 2X) \cdot B$ *and* $U^B = 6 + (10X - 6) \cdot B$*, respectively.*

The mechanised graph for the original game's MAID is shown in Fig. 5.6c. The cycle between Π_A and Π_B reveals that each agent's decision rule strategically relies on the other agent's decision rule.[3] Therefore, the MAID has insufficient information and no proper subgames, making it difficult to solve.

To find the CE of this game, a trusted mediator is added using a *correlation variable* C with $\mathbf{Pa}_C = \varnothing$, $\mathbf{Ch}_C = \boldsymbol{D}$, and $dom(C) = \dot{\boldsymbol{P}}$. In the mechanised graph, C's associated mechanism variable K_C represents the distribution $\kappa \in \Delta(\dot{\boldsymbol{P}})$ that the mediator draws a pure policy profile according to. This time, since K_C is fixed as κ at the game outset instead of being chosen by any agent, C acts as a chance variable (in contrast to the correlation decision variable introduced for mixed policies and behavioural mixtures).

There is a well-known difference between public and private recommendations. If public, every payoff in the convex hull of the set of NE payoffs can be attained by a CE; however, if the recommendations are private, then the payoffs to each agent in a CE can lie outside this convex hull (e.g., Aumann's game of chicken [61]). This distinction is made explicit in the MAID's graph. If the recommendations are public, then the full outcome of C (the pure policy profile chosen by the mediator) is known by every agent (shown by the dotted edges between C and both A and B in Fig. 5.6b). If the recommendations are private, then each agent only observes their decision rules (action recommendations) in C's outcome, i.e., all recommendations given to other players are hidden (at C^A and C^B in Fig. 5.6c). In this latter case, the agent infers, using Bayes' rule, a posterior over the pure policy profile that was chosen (and also which action was recommended to the other agent(s)). If κ is a CE, then each agent picks for their decision D's decision rule the mediator's recommendation, i.e., $\dot{\pi}_D$ where $c = \dot{\boldsymbol{\pi}}$. The set of variables $\boldsymbol{D}$ remain as decisions because agents are free to deviate from their recommendation and pick any CPDs as decision rules for their decisions.

[3] That Bob strategically relies on Alice's decision rule might be less obvious than the fact that Alice strategically relies on Bob's decision rule. The dependency occurs because since Bob can observe A, this unblocks an active path $\Pi_A \rightarrow A \leftarrow X \rightarrow U^B$ in the independent mechanised graph, so Π_A is s-reachable from Π_B.

This mediator's distribution $\kappa \in \Delta(\dot{\boldsymbol{P}})$ can be parameterised according to that in Fig. 5.6e. Note that $b_a\bar{b}_{\bar{a}}$ denotes the pure policy profile where Bob offers the job (b) to Alice if she selects a and Bob does not offer the job ($\bar{b}$) if Alice selects $\bar{a}$. Using the expected payoff for Alice and Bob under each pure policy profile (Fig. 5.6d), Definition 29's incentive constraints define 24 inequalities that must be satisfied by the CE distribution of a similar form to:

$$\alpha_1\mathbb{E}_{(a_x a_{\bar{x}}, b_a b_{\bar{a}})}[U^1] + \alpha_2\mathbb{E}_{(a_x a_{\bar{x}}, b_a \bar{b}_{\bar{a}})}[U^1] + \alpha_3\mathbb{E}_{(a_x a_{\bar{x}}, \bar{b}_a b_{\bar{a}})}[U^1] + \alpha_4\mathbb{E}_{(a_x a_{\bar{x}}, \bar{b}_a \bar{b}_{\bar{a}})}[U^1] \geq$$
$$\alpha_1\mathbb{E}_{(a_x \bar{a}_{\bar{x}}, b_a b_{\bar{a}})}[U^1] + \alpha_2\mathbb{E}_{(a_x \bar{a}_{\bar{x}}, b_a \bar{b}_{\bar{a}})}[U^1] + \alpha_3\mathbb{E}_{(a_x \bar{a}_{\bar{x}}, \bar{b}_a b_{\bar{a}})}[U^1] + \alpha_4\mathbb{E}_{(a_x \bar{a}_{\bar{x}}, \bar{b}_a \bar{b}_{\bar{a}})}[U^1]$$

This gives the following linearly independent inequalities:

$$\alpha_2 \geq \alpha_3 \tag{5.1}$$
$$\beta_3 \geq \beta_2 \tag{5.2}$$
$$\beta_2 \geq \beta_3 \tag{5.3}$$
$$\gamma_3 \geq \gamma_2 \tag{5.4}$$
$$\gamma_2 \geq \gamma_3 \tag{5.5}$$
$$\delta_3 \geq \delta_2 \tag{5.6}$$
$$-3\beta_1 + 2\gamma_1 - \delta_1 \geq 0 \tag{5.7}$$
$$-\alpha_1 + 2\beta_1 - 3\gamma_1 \geq 0 \tag{5.8}$$
$$-\alpha_1 - \beta_1 - \gamma_1 - \delta_1 \geq 0 \tag{5.9}$$
$$3\beta_2 - 2\gamma_2 + \delta_2 \geq 0 \tag{5.10}$$
$$-\alpha_2 + 5\beta_2 - 5\gamma_2 + \delta_2 \geq 0 \tag{5.11}$$
$$-\alpha_2 + 2\beta_2 - 3\gamma_2 \geq 0 \tag{5.12}$$
$$\alpha_3 - 2\beta_3 + 3\gamma_3 \geq 0 \tag{5.13}$$
$$\alpha_3 - 5\beta_3 + 5\gamma_3 - \delta_3 \geq 0 \tag{5.14}$$
$$-3\beta_3 + 2\gamma_3 - \delta_3 \geq 0 \tag{5.15}$$
$$\alpha_4 + \beta_4 + \gamma_4 + \delta_4 \geq 0 \tag{5.16}$$
$$\alpha_4 - 2\beta_4 + 3\gamma_4 \geq 0 \tag{5.17}$$
$$3\beta_4 - 2\gamma_4 + \delta_4 \geq 0 \tag{5.18}$$

From (2) and (3), we get that $\beta_2 = \beta_3$; from (4) and (5), that $\gamma_2 = \gamma_3$; and since all parameters ≥ 0, (9) implies that $\alpha_1 = \beta_1 = \gamma_1 = \delta_1 = 0$. This reduces the inequalities to:

$$\alpha_2 \geq \alpha_3 \tag{5.19}$$

$$\delta_3 \geq \delta_2 \tag{5.20}$$

$$3\beta_2 - 2\gamma_2 + \delta_2 \geq 0 \tag{5.21}$$

$$-\alpha_2 + 5\beta_2 - 5\gamma_2 + \delta_2 \geq 0 \tag{5.22}$$

$$-\alpha_2 + 2\beta_2 - 3\gamma_2 \geq 0 \tag{5.23}$$

$$\alpha_3 - 2\beta_2 + 3\gamma_2 \geq 0 \tag{5.24}$$

$$\alpha_3 - 5\beta_2 + 5\gamma_2 - \delta_3 \geq 0 \tag{5.25}$$

$$-3\beta_2 + 2\gamma_2 - \delta_3 \geq 0 \tag{5.26}$$

$$\alpha_4 + \beta_4 + \gamma_4 + \delta_4 \geq 0 \tag{5.27}$$

$$\alpha_4 - 2\beta_4 + 3\gamma_4 \geq 0 \tag{5.28}$$

$$3\beta_4 - 2\gamma_4 + \delta_4 \geq 0 \tag{5.29}$$

Adding (23) to (26) yields $-\alpha_2 - \beta_2 - \gamma_2 - \delta_3 \geq 0$, which implies that $\alpha_2 = \beta_2 = \gamma_2 = \delta_3 = 0$. From (19) and (20), this means that also $\alpha_3 = \delta_2 = 0$. This leaves only $\alpha_4, \beta_4, \gamma_4, \delta_4 \geq 0$ and just two (linearly independent) inequalities that need to be satisfied for a CE distribution:

$$\alpha_4 - 2\beta_4 + 3\gamma_4 \geq 0 \tag{5.30}$$

$$3\beta_4 - 2\gamma_4 + \delta_4 \geq 0 \tag{5.31}$$

After some algebra, we find that $\alpha_1 = \alpha_2 = \alpha_3 = \beta_1 = \beta_2 = \beta_3 = \gamma_1 = \gamma_2 = \gamma_3 = 0$; $\alpha_4, \beta_4, \gamma_4, \delta_4 \geq 0$; $\alpha_4 - 2\beta_4 + 3\gamma_4 \geq 0$, and $3\beta_4 - 2\gamma_4 + \delta_4 \geq 0$. Any CE, therefore, has Bob never offering a job to Alice because they play the pure policy $\bar{b}_a\bar{b}_{\bar{a}}$ with probability 1, i.e., Bob's decision rule has $\pi^B(B = \bar{b} \mid A = a) = \pi^B(B = \bar{b} \mid A = \bar{a}) = 1$. The remaining constraints require Alice not to give any incentive for Bob to offer her a job by making the conditional probability of Alice being hardworking too high relative to the conditional probability of her being lazy when he receives the signal a or $\bar{a}$. These constraints find that every CE will result in $\mathbb{E}_{\boldsymbol{\pi}}[U^A] = 0$ and $\mathbb{E}_{\boldsymbol{\pi}}[U^B] = 6$. This is unsurprising because, in a signaling game with costless signals, every CE will be a 'pooling equilibrium' [101] (an equilibrium in a signaling game in which all types of the sender choose the same action regardless of their true type).

Whilst the CE is among the best-known solution concepts for normal-form games, and is efficiently computable in that setting (e.g., via linear programming [102]), this concept becomes intractable to compute in EFGs since the strategic form of the extensive game has typically exponential size (by size we are referring to the size of the algorithmic input of the EFG with its game tree, information sets, moves, chance probabilities and payoffs). Hence, there can be an exponential number of linear incentive constraints in EFGs and even in bounded treewidth MAIDs. It is, therefore, currently unknown if a CE can be found in an EFG or MAID in poly-time. Motivated by these tractability concerns, Von Stengel and Forges proposed an *extensive-form correlated equilibrium (EFCE)* [62]. Along similar lines, we define a *MAID correlated equilibrium.*

Instead of revealing the entire recommendation $\dot{\boldsymbol{\pi}}^i$ to each agent i immediately, we let the mediator *stagger* their recommendations. This is made visible in the mechanised graph by adding the blue edges in Fig. 5.6c. Importantly, if an agent deviates from any recommendation, then the mediator will *cease giving further recommendations to that agent* (but will still give recommendations to all other agents). Thus, the incentive constraints are now tied to the threat of the mediator withholding future information.

Definition 30. *Given a distribution $\kappa \in \Delta(\dot{\boldsymbol{P}})$, consider the MAID with an additional correlation variable C with $\boldsymbol{Pa}_C = \varnothing$, $\boldsymbol{Ch}_C = \{C_D\}_{D \in \boldsymbol{D}}$, and $\boldsymbol{Ch}_{C_D} = \{D\}$ for each D. Let a pure policy profile $\dot{\boldsymbol{\pi}}$ be selected at C according to κ. Then, when each decision context $\boldsymbol{pa}_D$ is reached, agent i receives a recommended move $d \in dom(D)$ specified by $\dot{\boldsymbol{\pi}}_D \in \dot{\boldsymbol{\pi}}$ (C_D hides all other recommendations $\dot{\boldsymbol{\pi}}_{-D} \in \dot{\boldsymbol{\pi}}$). A **MAID correlated equilibrium (MAID-CE)** is an NE of this game in which no agent has an incentive to deviate from their recommendations.*

The localised recommendations in a MAID-CE pose weaker incentive constraints compared to a CE, so the set of MAID-CE outcomes is larger. As such, MAID-CEs can lead to Pareto improvements over the CEs (and NEs) in a game. We now give one such MAID-CE. The mediator chooses a signal s with equal probability for type $X = x$, i.e., $\Pr(c_A = a \mid X = x) = \Pr(c_A = \bar{a} \mid X = x) = 0.5$. Bob is recommended to offer Alice a job (b) when Alice's action matches s and to reject otherwise ($\bar{b}$). If $X = \bar{x}$, then the recommendation to Alice is arbitrary and is independent of the signal s, which is only shown to hardworking Alice. Because the mediator only gives Alice her recommendation once her decision context $\mathbf{Pa}_A$ is set, lazy Alice cannot know s. Therefore, in any situation, lazy Alice's action will match s with probability $\frac{1}{2}$. Consequently, when Bob is called to play (i.e., the decision context $\mathbf{Pa}_B$ is set), and Alice's action matches s, Alice is twice as likely to be hardworking than lazy (so $\mathbb{E}_{\boldsymbol{\pi}}[U^B] = \frac{20}{3}$ for offering Alice a job rather than $expect_{\boldsymbol{\pi}}[U^B] = 6$ for rejecting her). If instead, Alice's action does not match s, then he knows with certainty that Alice is lazy, so his best response is to reject. Overall, Alice's expected payoff in this MAID-CE is 3.5, and Bob's is 6.5 (higher than 0 and 6, respectively, for all CEs).

A MAID-CE can be computed in poly-time if the treewidth is bounded, via a reduction to a linear program. We follow Huang et al. [103]'s method because the information sets in an EFG are in bijection with the decision contexts in a MAID but relax beyond their conditions as MAIDs only require sufficient (rather than perfect) recall [103]. Any distribution over pure policies induced by an NE can be represented using a distribution κ, and hence any mixed NE (or equivalent behavioural NE) is also a CE and MAID-CE. As every MAID has an NE in (mixed) policies, every MAID must also have a CE and a MAID-CE. Like our complexity results later, the complexity is heavily influenced by the **treewidth** of the instance's DAG – treewidth measures a DAG's resemblance to a tree and is given by the number of vertices in the largest clique of the corresponding triangulated moral graph minus one [104].

Proposition 13. *A MAID-CE in bounded treewidth MAIDs with sufficient recall can be found in poly-time.*

Proof sketch. We follow Huang and von Stengel's method for this result [103]. Our result comes from the observation that if there is sufficient recall in a MAID, then: (i) the set of decision contexts of every decision node in the MAID is in bijection with the set of all information sets in a corresponding EFG, and (ii) sufficient recall is sufficient for the ordering of decision contexts analogous to Huang and von Stengel's ordering of information sets. $\square$

5.6 Complexity Results in MAIDs

We are interested in trying to predict the behaviour of agents in environments as a result of their optimisation processes. In MAIDs, it is therefore important to understand rigorously which problems are hard, what makes them hard, and what restrictions can make them tractable. This helps us to understand which problems are suitable for analysis as MAIDs.

Since a full policy profile in a MAID induces a BN, many of our results inherit from that setting, where the decision problem variant of marginal inference is, in general, PP-complete [105]. However, we care about the cases we encounter in practice, not just the worst case. Marginal inference in a BN can be performed in time exponential in the treewidth of the underlying graph [95], which entails a poly-time algorithm when the treewidth is small. Similarly, we will see that tractable results for computations in MAIDs can be found when problems are restricted to certain settings. We also sometimes reduce from partial order games [106], which can be interpreted as MAIDs without chance nodes, with deterministic decision rules, and where each agent has a single utility node as a child of all the decision nodes. Whilst complexity results have been found for single-agent influence diagrams [11, 107–110], to our knowledge, we are the first to define and prove important complexity problems for influence diagrams with multiple agents.

5.6.1 Concise Representations

A concise representation ensures that a decision problem's input does not pack our complexity results (in Section 5.6). By this, we mean that the size of the input accurately reflects the inherent complexity of the decision problem, rather than artificially increasing it through inefficient representation. First, real numbers may obscure the true complexity of the problems [111], so we assume that all probability parameters are given by a fraction of two integers, both of which are expressed in finite binary notation. This is realistic since the probabilities are normally either assessed by domain experts or estimated by a learning algorithm from data instances and means that all CPDs can be read in poly-time. Second, even with binary variables, a joint distribution across $\boldsymbol{V}$ requires $2^{|\boldsymbol{V}|}-1$ parameters. A MAID or BN's graphical Markov factorisation reduces this to $\sum_{V\in\boldsymbol{V}} 2^{|\mathbf{Pa}_V|}$, but this can still be exponential in $|\boldsymbol{V}|$. Therefore, it is standard [112–114] to assume that the maximum in-degree in the graph is much less than $|\boldsymbol{V}|$ (or constant), so that the size of the CPDs are polynomial in $|\boldsymbol{V}|$. This means that the total representation of our MAID (including all CPDs) is polynomial in our chosen complexity parameter $|\boldsymbol{V}|$. Finally, as in BNs, our complexity results are strongly affected by the DAG's **treewidth**. The **treewidth** of a DAG measures its resemblance to a tree and is given by the number of vertices in the largest clique of the corresponding triangulated moral graph minus one [104].

We now give some complexity results in MAIDs. Our first follows from the known result in normal-form games [91]. Any normal-form game $\mathcal{N}$ can be reduced to a MAID where each agent has one utility node (which copies the payoffs in $\mathcal{N}$) and one decision node. The domains of the decision variables are the set of each agent's pure strategies in $\mathcal{N}$. Edges are added from every $D \in \boldsymbol{D}$ to every $U \in \boldsymbol{U}$.

Proposition 14. *In a MAID, finding an NE in mixed policies is* PPAD*-hard.*

Problem	**Input**	**Question**
IS-BEST-RESPONSE	$\mathcal{M}$, i, $\boldsymbol{\pi}^{-i}$, $q \in \mathbb{Q}$	Is there some $\hat{\boldsymbol{\pi}}^i$ such that $EU^i(\hat{\boldsymbol{\pi}}^i, \boldsymbol{\pi}^{-i}) > q$?
IS-NASH	$\mathcal{M}$, $\boldsymbol{\pi}$	Is $\boldsymbol{\pi}$ a (behavioural) NE of $\mathcal{M}$?
NON-EMPTINESS:	$\mathcal{M}$	Does $\mathcal{M}$ have a (behavioural) NE?

Table 5.1: Three decision problems in MAIDs with behavioural policies.

In the following results, we focus on the complexity of the decision problems in Table 5.1.

Proposition 15. *IS-BEST-RESPONSE is* $\mathrm{NP}^{\mathrm{PP}}$*-complete,* NP*-complete when restricted to MAIDs with graphs of bounded treewidth, and* PP*-complete if both* $|\boldsymbol{D}^i|$ *and the in-degrees of* $\boldsymbol{D}^i$ *are bounded.*

Proof sketch. IS-BEST-RESPONSE is in $\mathrm{NP}^{\mathrm{PP}}$ because given $\hat{\boldsymbol{\pi}}^i$, we can verify that $EU^i(\hat{\boldsymbol{\pi}}^i, \boldsymbol{\pi}^{-i}) > q$ in poly-time using a PP oracle for inference in a BN. With bounded treewidth, verification can be done in poly-time. The final setting is in PP by analogy with Kwisthout's PARAMETER TUNING [115]. For the general case's hardness, we can reduce from E-MAJSAT as in [116], where MAP-nodes are replaced by agent i's decision nodes; for bounded treewidth, we can reduce from MAXSAT as in [117]; and for the final case, IS-BEST-RESPONSE with $|\boldsymbol{D}^i| = 0$ is the same as inference in a BN. □

Proposition 15 suggests IS-BEST-RESPONSE is, in general, only tractable if inference is easy *and* $|\boldsymbol{D}^i|$ is bounded by a constant. Proposition 16 then explains the decision problem's name.

Lemma 5. *If IS-BEST-RESPONSE can be solved in poly-time, then agent i's expected utility under a best response to a partial policy profile $\boldsymbol{\pi}^{-i}$ in a MAID can be found in poly-time.*

Proof. This follows immediately from using binary search over agent i's policies and uses the fact that we are restricting parameters in the MAID to be rational numbers. □

Proposition 16. *If the in-degrees of $\boldsymbol{D}^i$ are bounded and IS-BEST-RESPONSE can be solved in poly-time, then a best response policy for agent i to a partial profile $\boldsymbol{\pi}^{-i}$ can be found in polynomial time.*

Proof. Begin by constructing the MAID $\mathcal{M}(\boldsymbol{\pi}^{-i})$ by replacing decision nodes $\boldsymbol{D} \setminus \boldsymbol{D}^{-i}$ as chance nodes with CPDs given by $\boldsymbol{\pi}^{-i}$. Next, use Lemma 5 to compute agent i's expected utility under a best response policy in $\mathcal{M}(\boldsymbol{\pi}^{-i})$ and use this value as q. Take each of agent i's decision variables $D \in \boldsymbol{D}^i$ and build a new MAID $\mathcal{M}(\boldsymbol{\pi}^{-i}, \pi_D)$ for every possible decision rule of D (i.e., replace D as a chance node with CPD π_D). The fact that the in-degrees of agent i's decision nodes are bounded, bounds the number of these MAIDs. For each induced MAID, we can then use a poly-time algorithm for IS-BEST-RESPONSE to determine any decision rule π_D that makes up the best response policy for agent i. □

Proposition 17. *Is-Nash is* $\mathrm{coNP}^{\mathrm{PP}}$*-complete, and* coNP*-complete when restricted to MAIDs with graphs of bounded treewidth. The general problem remains* $\mathrm{coNP}^{\mathrm{PP}}$*-hard in sufficient information MAIDs. In MAIDs without chance variables, the problem remains* coNP*-hard.*

Proof sketch. For membership, we can check that $\boldsymbol{\pi}$ is *not* an NE by guessing an agent i and checking if $\boldsymbol{\pi}^i \in \boldsymbol{\pi}$ is a best response in poly-time using a PP-oracle (this is unnecessary if the graph has bounded treewidth). Hardness comes from the single-agent setting where it is the complement of Is-Best-Response. In MAIDs without chance variables, we reduce from partial order games [106]. □

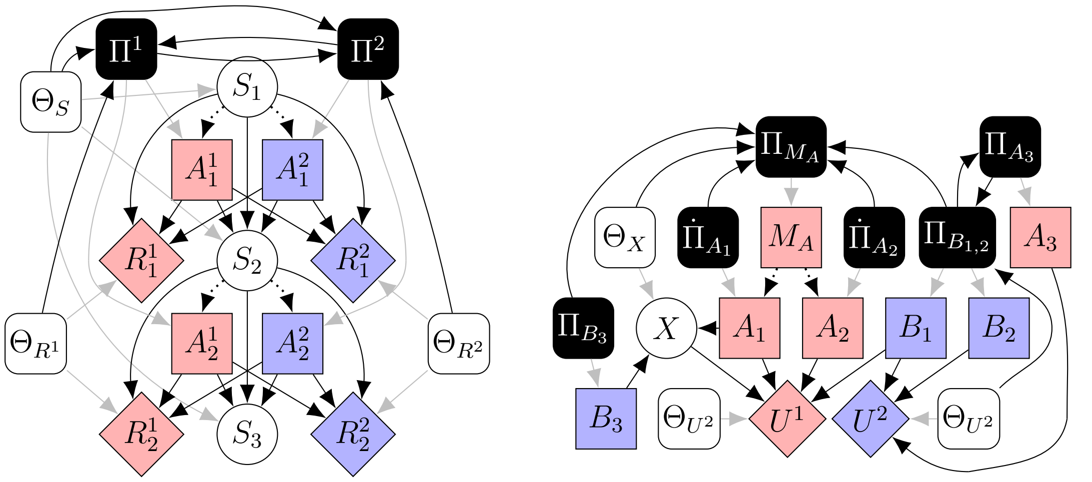
Proposition 10 shows when Non-Emptiness is vacuous. However, in an insufficient recall MAID, Non-Emptiness is, in general, intractable even without chance variables.

Proposition 18. *Non-Emptiness is* NEXPTIME*-hard and becomes* NEXPTIME*-complete if we restrict to MAIDs without chance variables.*

Proof sketch. For hardness, we can reduce from partial order games. Without chance variables, we can determine Non-Emptiness using a similar algorithm to that in Zahoransky et al., [106]. It exploits the setting's determinism: payoffs are poly-time computable and the number of policy profiles is reduced to $\mathcal{O}(2^{|\boldsymbol{V}|})$. □

Proposition 19. *In a MAID with sufficient information, if the in-degrees of* $\boldsymbol{D}$ *are bounded and Is-Best-Response can be solved in poly-time, then a pure NE can be found in poly-time.*

Proof. First, note that we can check whether a MAID is a sufficient information game in poly-time using s-reachability, a graphical criterion based on d-separation [66]. We can then follow the constructive procedure given for the proof of Proposition 10. Given Proposition 16, each optimisation step must take poly-time and since the in-degrees of all decision nodes are bounded by a constant, the number of subgames is also bounded by a constant. Therefore, the entire procedure takes poly-time. □



(a) **(b)**

Figure 5.7: Mechanised graphs for (a) a Markov game and (b) a team setting with imperfect communication.

This result suggests an NE can be found efficiently in certain MAIDs, but even in games without sufficient information, NEs can be found more efficiently in a MAID than in an EFG. The mechanised graph dependencies reveal more 'subgames' – parts of the MAID that can be solved independently from the rest – to which dynamic programming can be applied [26, 87]. As finding an NE in both EFGs and MAIDs depends significantly on the game's size, this can empirically lead to large compute savings [26].

5.7 Applications of Imperfect Recall MAIDs

We introduced forgetfulness and absent-mindedness as properties of individual agents (due to imperfect memory). However, imperfect recall also commonly arises in *team situations*; each team consists of several agents targeting a common goal with imperfect communication. Forgetfulness or absent-mindedness occurs when an agent does not know their teammates' actions (or observations) or whether they have acted at all. Mechanised graphs represent these situations where teams often employ a mix of randomisation strategies (e.g., Fig. 5.7b). For mixed policies, the random seed is chosen at the start before the agents set out following their distinct policies. For behavioural policies, agents pick a new random seed at every decision point. Behavioural mixtures correspond to randomising at both stages.

Another application of imperfect recall in MAIDs is to *Markov (or 'stochastic') games* [118], in which the agents move between different states over time (e.g., Fig. 5.7a). At each time step t, each agent i selects an action A^i_t, and the game probabilistically transitions to a new state S_{t+1}, depending on the previous state S_t and the actions selected, and each agent receives a payoff R^i_t. Each S_{t+1} and R^i_t has parents $\{S_t, A^1_t, \ldots, A^n_t\}$ and must be identically distributed for all t, again represented using shared mechanism variables. Often, the agent must learn a memoryless, stationary policy $\pi^i : S \to \Delta(A^i)$, where S is the set of states and $\Delta(A^i)$ the set of probability distributions over agent i's actions. Hence, the agents are absent-minded (every decision A^i_{t+1} of agent i shares the same decision rule) and use *behavioural* policies (since the action selected in each state is independently stochastic). In light of Proposition 6, it is, therefore, natural to ask whether a Markov game may not have an NE in memoryless stationary policies. It is known that infinite-horizon Markov games might not (for a counterexample, see de et al., [119]). Although infinite games lie outside of the scope of this thesis, it is nonetheless interesting to note that this possible non-existence is due to absent-mindedness: if agents can choose a different decision rule at each time step, a behavioural NE is guaranteed [120].

5.8 Conclusions

In this chapter, we have shown how to handle imperfect recall in MAIDs by overcoming the potential lack of NEs in behavioural policies using mixed and correlated equilibria. EFGs leave many assumptions about how agents play games hidden, but mechanised graphs make explicit the assumptions behind imperfect recall (both forgetfulness and absent-mindedness), mixed policies, and two types of correlated equilibria. Previous work used the concept of an NE in an MAID but did not fully characterise when it is and isn't guaranteed to exist. We show that NE existence guarantees depend on which class of policies agents are permitted to choose from. Our complexity results highlight the importance of restricting the use of MAIDs to those with a limited number of decision variables and bounded treewidth. Finally, our applications to Markov games and team situations show that imperfect recall broadens the scope of what can be modelled using MAIDs.

6
Causal Games

Contents

6.1 Introduction

Many types of queries may be of interest in game-theoretic scenarios. For instance, recalling Example 1 (in Chapter 2), we could ask questions corresponding to:

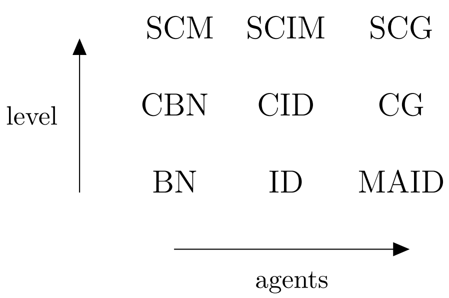
1) Predictions, such as a) 'Given that the worker went to university, what is their expected wellbeing?' or b) 'Given that the worker always decides to go

to university, what is their expected wellbeing?'

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

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

Although these queries are ostensibly similar, they belong to different levels of Pearl's causal hierarchy (see Section 2.1). 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.



B	Bayesian
C	Causal
G	Game
I	Influence
M	Model
MA	Multi-Agent
N	Network
S	Structural

Figure 6.1: In this thesis, 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 Fig. 6.1 (a duplicate of that in Fig. 1.1). 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 13 and the equilibrium refinements we introduce in Section 4.3) 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 $\boldsymbol{\pi}$ 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 6.3 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 and 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; we unify both types of query using the same formalism.

In this chapter, we show how the six types of causal queries represented by the examples above (each of which is written formally in Table 6.1) 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. In Section 6.4, we fully characterise causal interventions so that any arbitrarily complex causal query can be studied in multi-agent settings.

6.2 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

	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	$\Pr(u^1 \mid \bar{\pi}_{D^1})$	$\Pr(u^1_{\hat{\pi}_{D^1}})$	$\Pr(u^1_{\tilde{\pi}_{D^1}} \mid \bar{\pi}_{D^1})$

Table 6.1: 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 2.1, $\Pr(\boldsymbol{x_y})$ 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.

$\Pr^{\boldsymbol{\pi}}(\boldsymbol{V}; \boldsymbol{\theta})$ 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 3.2, there may be many or no such rational outcomes.

Definition 31. *Given a mechanised MAID* $\mathsf{m}\mathcal{M}$ *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*

$$\overset{\mathcal{R}}{\Pr}(\boldsymbol{x} \mid \boldsymbol{z}) \coloneqq \left\{ \overset{\boldsymbol{\pi}}{\Pr}(\boldsymbol{x} \mid \boldsymbol{z}) \right\}_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z})}$$

where

$$\mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z}) \coloneqq \{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M}) : \overset{\boldsymbol{\pi}}{\Pr}(\boldsymbol{z}) > 0\}$$

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* $\mathsf{m}\mathcal{M}$*, where* $\mathbf{M}' = \mathbf{M} \setminus \boldsymbol{Z}$ *and* $\mathbf{m}_{\boldsymbol{D}} = \boldsymbol{\pi}$*.*

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 31. For example, we can bind $\boldsymbol{\pi}$ by quantifying over it as in $\exists \boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z}). \big(\varphi(\boldsymbol{\pi}) \sim q\big)$ 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 [121–123]), 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 [43]). 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 [46, 47], 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 chooses $\neg g$. 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 $\Pr^{\mathcal{R}}(u^1 \mid g)$, 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 31 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 $\Pr^{\mathcal{R}}(u^1 \mid \bar{\pi}_{D^1})$ 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 $\Pr^{\boldsymbol{\pi}}(h \mid g, j) = \Pr^{\boldsymbol{\pi}}(h \mid g)$ 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 $\Pr^{\boldsymbol{\pi}}(h \mid g, j)$ in $\mathcal{M}$ is given by $\Pr(h \mid g, j, \mathbf{m})$ in $\mathsf{m}\mathcal{M}$, where $\mathbf{m}_{\boldsymbol{D}} = \boldsymbol{\pi}$. 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.

6.3 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 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 32. *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* $\Pr^{\boldsymbol{\pi}}(\boldsymbol{V})$ *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 [39], but where the cyclic dependencies are governed by relations instead of functions [44]) we quantify over the set of rational outcomes that are consistent with the given intervention.

Definition 33. *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* $\Pr^{\mathcal{R}}(\boldsymbol{x}_{\mathcal{I}}) := \left\{ \Pr^{\boldsymbol{\pi}}(\boldsymbol{x}_{\mathcal{I}}) \right\}_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M}_{\mathcal{I}})}$ *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 $\Pr^{\mathcal{R}}(u^1_g)$ and thus perform a hard intervention $\mathrm{do}(D^1 = g)$ in the mechanised game (shown in Fig. 6.2a). As the set of rational outcomes does not change under a post-policy intervention, we have that $\Pr^{\mathcal{R}}(u^1_g) = \left\{ \Pr^{\boldsymbol{\pi}}(u^1_g) \right\}_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M})}$, which results in $\left\{ \mathbb{E}_{\boldsymbol{\pi}}[U^1_g] \right\}_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M})} = \{-\frac{3}{2}, \frac{7}{2}\}$. 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 expected well-being 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 $\Pr(u^1_{\hat{\pi}_{D^1}})$ 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 Fig. 6.2b – 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 $\left\{\mathbb{E}_{\pi}[U^1_{\hat{\pi}_{D^1}}]\right\}_{\pi \in \mathcal{R}(\mathsf{m}\mathcal{M}_{\mathcal{I}})} = \{\frac{17}{4}\}$.

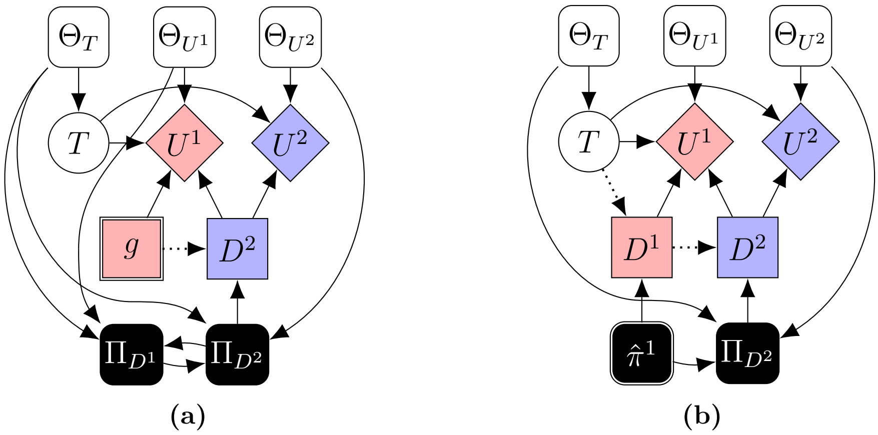


Figure 6.2: (a) A mechanised game showing the hard post-policy intervention $\mathrm{do}(D^1 = g)$, where incoming edges to D^1 are severed. (b) A mechanised game showing the hard pre-policy intervention $\mathrm{do}(\Pi_{D^1} = \hat{\pi}^1)$, where incoming edges to Π_{D^1} are severed.

Remark. *In previous work, soft interventions on object-level variables V have been modelled as hard interventions on its mechanism variable M_V [31]. While these can be viewed as essentially equivalent in the single-agent case,*[1] *possible dependencies between mechanism variables in mechanised games mean 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).*

[1]To be precise, this is true when the agent has *sufficient recall*; see Definition 27.

6.4 Fully Characterising Interventions

In the previous section, we introduced a distinction between *pre-policy* queries, where the intervention occurs before the policy profile is selected, and *post-policy* queries, where the intervention occurs after. In this section, we extend this to accommodate arbitrary queries where each agent makes decisions based on the subset of interventions visible to them.

6.4.1 Primitive Interventions

Given a causal game $\mathcal{M}$ with mechanised graph $\mathsf{m}\mathcal{G}$ and rationality relations $\mathcal{R}$, we saw that an intervention $\mathcal{I}$ is a function that maps a set of joint probability distributions $\{\Pr^{\boldsymbol{\pi}}(\boldsymbol{v})\}_{\boldsymbol{\pi}\in\mathcal{R}}$ to a new set $\{\Pr^{\boldsymbol{\pi}}(\boldsymbol{v}_{\mathcal{I}})\}_{\boldsymbol{\pi}\in\mathcal{R}^*}$ where $\mathcal{R}^*$ are the rationality relations of the intervened game with graph $\mathsf{m}\mathcal{G}_{\mathcal{I}}$ and $\Pr^{\boldsymbol{\pi}}(\boldsymbol{v}_{\mathcal{I}})$ is the joint probability distribution represented by the CBN induced by $\mathsf{m}\mathcal{G}_{\mathcal{I}}$ when parameterised over policy profile $\boldsymbol{\pi}$. We now define four primitive types of intervention.

(1) Fixing an object-level variable: Intervening on variable X replaces $\Pr^{\boldsymbol{\pi}}(x \mid \mathbf{pa}_X)$ with a new CPD $\Pr^{\mathcal{I}}(x \mid \mathbf{pa}^*_X)$. Graphically, when $\mathbf{pa}^*_X \neq \mathbf{pa}_X$, the incoming edges to X are changed such that $V \to X$ exists if and only if $V \in \mathbf{pa}^*_X$. The induced distribution is

$$\Pr^{\boldsymbol{\pi}}(\boldsymbol{v}_{\mathcal{I}}) = \Pr^{\mathcal{I}}(x \mid \mathbf{pa}^*_X) \cdot \prod_{V \in \boldsymbol{V}\setminus\{X\}} \Pr^{\boldsymbol{\pi}}(v \mid \mathbf{pa}_V)$$

A *hard object-level intervention* assigns $\Pr^{\mathcal{I}} = \delta(X, g)$. In Pearl. [3]'s do-calculus, this is written $\mathrm{do}(X = g)$. Any other form of object-level intervention is qualified as *soft*.

(2) Fixing a mechanism variable: A *hard mechanism-level intervention* $\mathrm{do}(\mathsf{M}_V = \mathsf{m}_V)$ sets the distribution over each mechanism M_V to $\delta(\mathsf{M}_V, \mathsf{m}_V)$. Any other form of mechanism-level intervention is qualified as *soft*. A mechanism-level intervention on decision rule Π_D replaces $r_D : dom(\mathbf{Pa}_{\Pi_D}) \to dom(\Pi_D)$ with a new rationality relation $r^{\mathcal{I}}_D : dom(\mathbf{Pa}^*_{\Pi_D}) \to dom(\Pi_D)$. Graphically, when $\mathbf{Pa}^*_{\Pi_D} \neq \mathbf{Pa}_{\Pi_D}$, the incoming edges to variable Π_D are changed such that $V \to \Pi_D$ exists if and only if $V \in \mathbf{Pa}^*_{\Pi_D}$. For a parameter variable Θ_V of $V \in \boldsymbol{V} \backslash \boldsymbol{D}$, an intervention assigns a new distribution from domain $\Delta(dom(V) \mid dom(\mathbf{Pa}_V))$. Note that parameter variables don't have parent mechanism variables as inputs to the choice of distribution.

(3) Adding a new object-level variable: Adding a new object-level variable Y introduces new CPD $\Pr^{\mathcal{I}}(y \mid \mathbf{pa}_Y)$ to the joint distribution factorization. Graphically, this adds a new node Y to $\mathcal{G}$ and adds edges $X \to Y$ for all $X \in \mathbf{Pa}_Y$ and $Y \to Z$ for all $Z \in \mathbf{Ch}_Y$.

The induced distribution is

$$\Pr^{\boldsymbol{\pi}}((\boldsymbol{v} \cup Y)_{\mathcal{I}}) = \Pr^{\mathcal{I}}(y \mid \mathbf{pa}_Y) \cdot \prod_{V \in \boldsymbol{V}} \Pr^{\boldsymbol{\pi}}(v \mid \mathbf{pa}^*_V)$$

where, for $V \in \boldsymbol{V}$,

$$\mathbf{Pa}^*_V = \begin{cases} \mathbf{Pa}_V \cup Y & \text{if } V \in \mathbf{Ch}_Y \\ \mathbf{Pa}_V & \text{otherwise} \end{cases}$$

(4) Removing an existing object-level variable: Removing an existing object-level variable Y removes the CPD $\Pr^{\boldsymbol{\pi}}(y \mid \mathbf{pa}_Y)$ from the joint distribution factorization. Graphically, this removes the node Y from $\mathcal{G}$ and removes edges $X \to Y$ for all $X \in \mathbf{Pa}_Y$ and $Y \to Z$ for all $Z \in \mathbf{Ch}_Y$. The induced distribution is

$$\Pr^{\boldsymbol{\pi}}((\boldsymbol{v} \setminus \{Y\})_{\mathcal{I}}) = \prod_{V \in \boldsymbol{V} \setminus \{Y\}} \Pr^{\boldsymbol{\pi}}(v \mid \mathbf{pa}^*_V)$$

where, for $V \in \boldsymbol{V}$,

$$\mathbf{Pa}^*_V = \begin{cases} \mathbf{Pa}_V \setminus \{Y\} & \text{if } V \in \mathbf{Ch}_Y \\ \mathbf{Pa}_V & \text{otherwise} \end{cases}$$

Remark. *After any intervention of type 1, 3, or 4,* $\mathsf{m}\mathcal{G}$ *must be updated to reflect any changes in* $\mathcal{R}$*-reachability between mechanisms. Note that a type 1 intervention can be considered a type 4 intervention followed by a type 3 intervention, but we include it as a primitive for convenience.*

Theorem 2. *Primitive interventions are a sound and complete formulation of causal interventions.*

Proof. We first prove soundness. This is true using the definitions of the primitive interventions. Let $\Pr(\boldsymbol{v}_\mathcal{I})$ be the induced joint distribution by type 1, 3, and 4 interventions, as per the definitions, and $\Pr(\boldsymbol{v}_\mathcal{I}) = \Pr(\boldsymbol{v})$ for type 2 interventions. Also, let $\mathcal{R}^*$ be the induced rationality relations by type 2 interventions and $\mathcal{R}^* = \mathcal{R}$ for type 1, 3, and 4 interventions. Then, the effect of a primitive intervention P of any type is the function $\{\Pr^{\boldsymbol{\pi}}(\boldsymbol{v})\}_{\boldsymbol{\pi}\in\mathcal{R}} \mapsto \{\Pr^{\boldsymbol{\pi}}(\boldsymbol{v}_\mathcal{I})\}_{\boldsymbol{\pi}\in\mathcal{R}^*}$ which is a valid causal intervention.

We now turn to completeness by showing that any intervention $\mathcal{I}$ can be decomposed into a set of equivalent primitive interventions $\boldsymbol{\mathcal{P}}$. That is to say, the state of the game after applying intervention $\mathcal{I}$ is equivalent to the state of the game after applying interventions $\boldsymbol{\mathcal{P}}$. Suppose,

$$\begin{aligned}
\mathcal{I}(\{\Pr^{\boldsymbol{\pi}}(\boldsymbol{v})\}_{\boldsymbol{\pi}\in\mathcal{R}}, \mathcal{R}) &= (\{\Pr^{\boldsymbol{\pi}}(\boldsymbol{v}_\mathcal{I})\}_{\boldsymbol{\pi}\in\mathcal{R}^*}, \mathcal{R}^*) \\
\text{s.t. } \Pr^{\boldsymbol{\pi}}(\boldsymbol{v}) &= \prod_{V\in\boldsymbol{V}} \Pr^{\boldsymbol{\pi}}(v \mid \mathbf{pa}_V) \\
\text{and } \Pr^{\boldsymbol{\pi}}(\boldsymbol{v}_\mathcal{I}) &= \prod_{V_\mathcal{I}\in\boldsymbol{V}_\mathcal{I}} \Pr^{\boldsymbol{\pi}}(v_\mathcal{I} \mid \mathbf{pa}_V)
\end{aligned}$$

Then the trivial decomposition of $\mathcal{I}$ is $|\boldsymbol{V}_\mathcal{I}|$ type 3 interventions which multiply the joint distribution by each of $\Pr^{\boldsymbol{\pi}}(v_\mathcal{I} \mid \mathbf{pa}_V)$, followed by $|\boldsymbol{V}|$ type 4 interventions that divide the joint distribution by each of $\Pr^{\boldsymbol{\pi}}(v \mid \mathbf{pa}_V)$, then $|\mathcal{R}^*|$ type 2 interventions that attach the appropriate rationality relation to each mechanism variable.

Of course, more concise decompositions are possible if there is an overlap between $\mathcal{R}$ and $\mathcal{R}^*$ as well as an overlap between $\{\mathbf{Pa}_V\}_{V\in\boldsymbol{V}}$ and $\{\mathbf{Pa}_{V_\mathcal{I}}\}_{V_\mathcal{I}\in\boldsymbol{V}_\mathcal{I}}$. □

There are a number of other interesting intervention types that can be constructed by composing these primitives.

Unfixing an object-level variable: For every type 1 intervention $\mathcal{I}$ which fixes variable X, there is a type 1 inverse intervention $\mathcal{I}'$ which unfixes it. It restores the intervened CPD to be based on the original policy profile $\boldsymbol{\pi}$ and parents $\mathbf{Pa}_X$, rather than $\mathcal{I}$ and $\mathbf{Pa}^*_X$.

Unfixing a mechanism variable: Similarly, for every type 2 intervention $\mathcal{I}$ which fixes a variable Π_D, there exists a type 2 inverse intervention $\mathcal{I}'$ which unfixes it. This restores the rationality relation associated with Π_D to its default r_D, rather than $r_D^{\mathcal{I}}$. It also makes the mechanism conditionally dependent on the original parents $\mathbf{Pa}_{\Pi_D}$ rather than $\mathbf{Pa}^*_{\Pi_D}$.

Adding an object-level dependency: Adding a dependency, e.g., $\text{add}(X \to Y)$, is equivalent to a type 1 intervention where $\Pr^{\mathcal{I}}(y \mid \mathbf{pa}_Y) = \Pr^{\boldsymbol{\pi}}(y \mid \mathbf{pa}_Y \cup \{X\})$.

Removing an object-level dependency: Removing a dependency, e.g., $\text{del}(X \to Y)$, is equivalent to a type 1 intervention where $\Pr^{\mathcal{I}}(y \mid \mathbf{pa}_Y) = \Pr^{\boldsymbol{\pi}}(y \mid \mathbf{pa}_Y \setminus \{X\})$.

6.4.2 Interventional Queries

An interventional query concerns the outcome of a game after a set of causal interventions $\boldsymbol{\mathcal{I}}$, where each agent is privy to the state of the game after a subset of these interventions has been performed. We say that an intervention is *visible* to an agent if the agent has an opportunity to adapt their policy to that intervention. Consider Example 1. Unbeknownst to the firm, the worker may have an alternative job offer which changes her best-response policy. Simultaneously, the firm may have new hiring quotas, which change their payoffs and, therefore, their best response, but which are not disclosed to the worker. These two external interventions can be expressed in a unified analysis using our framework.

First, we introduce some new notation: $\boldsymbol{\mathcal{P}}$ denotes a set of primitive interventions ; $\boldsymbol{\mathcal{I}}^i \subseteq \boldsymbol{\mathcal{I}}$ denotes the set of interventions visible to agent i; $\boldsymbol{\mathcal{I}}(\mathcal{M})$ denotes the state of the causal game after applying interventions $\boldsymbol{\mathcal{I}}$ in any order; and the $\circ$ operator denotes ordered composition where $(\boldsymbol{\mathcal{I}}_1 \circ \boldsymbol{\mathcal{I}}_0)(\mathcal{M})$ is the state of the game after applying $\boldsymbol{\mathcal{I}}_0$ then $\boldsymbol{\mathcal{I}}_1$ (as shorthand, $\mathcal{I}_0 \circ \mathcal{I}_1$ means $\{\mathcal{I}_0\} \circ \{\mathcal{I}_1\}$).

Remark. *The order in which interventions are applied is important because interventions are not commutative. Consider, for example, two hard object-level* *interventions on the same variable but to different CPDs,* $\delta(X, a)$ *and* $\delta(X, b)$*. Then clearly* $(\mathrm{do}(X = a) \circ \mathrm{do}(X = b))(\mathcal{M}) \not\equiv (\mathrm{do}(X = b) \circ \mathrm{do}(X = a))(\mathcal{M})$*.*

The $\mathcal{R}$-rational outcomes of the game after each agent i has an opportunity to adapt to her visible interventions $\mathcal{I}^i$, is denoted $\mathcal{R}(\mathcal{M}_I)$. Θ_I denotes the parameterisation of non-decision mechanisms after interventions $\mathcal{I}$. Using this, we define an interventional query which Theorem 3 proves can always be decomposed into primitive intervention sets.

Definition 34 (Interventional Query)**.** *Given a CG* $\mathcal{M}$*, rationality relations* $\mathcal{R}$*, and a set of visible interventions for each agent* $\boldsymbol{\mathcal{I}}^1, \ldots, \boldsymbol{\mathcal{I}}^N$*, an interventional query* $\phi(\boldsymbol{\pi})$ *is a first-order logical formula that acts on the joint probability distribution* $\Pr^{\boldsymbol{\pi}}$ *induced by a* $\mathcal{R}$*-rational outcome* $\boldsymbol{\pi} \in \mathcal{R}(\mathcal{M}_{\boldsymbol{\mathcal{I}}})$ *and a parameterisation* $\Theta_{\boldsymbol{\mathcal{I}}}$ *where* $\boldsymbol{\mathcal{I}} = \boldsymbol{\mathcal{I}}^1 \cup \cdots \cup \boldsymbol{\mathcal{I}}^N$*.*

Theorem 3 says that for any set of interventions $\boldsymbol{\mathcal{I}}$, where the visible set of each agent is an *arbitrary* subset, $\boldsymbol{\mathcal{I}}^i \subseteq \boldsymbol{\mathcal{I}}$, we can construct an *ordered* list of primitive interventions such that, after the first j sets of primitive interventions, the state of the game is the exact state visible to Agent i when choosing her policy.

Theorem 3 (Decomposition of Intervention Sets)**.** *For any set of interventions* $\boldsymbol{\mathcal{I}}$*, where* $\boldsymbol{\mathcal{I}}^i \subseteq \boldsymbol{\mathcal{I}}$ *is the subset of interventions visible to agent* i*, there are primitive intervention sets* $\boldsymbol{\mathcal{P}}_0, \ldots, \boldsymbol{\mathcal{P}}_m$*, such that*

$$\forall i \; \exists j \in \{0, \ldots, m\} : \boldsymbol{\mathcal{I}}^i(\mathcal{M}) = (\boldsymbol{\mathcal{P}}_j \circ \boldsymbol{\mathcal{P}}_{j-1} \circ \cdots \circ \boldsymbol{\mathcal{P}}_0)(\mathcal{M}) \tag{6.1}$$

Proof. We show there is a set of primitive intervention sets that satisfy Property 6.1 in Theorem 3 for a set of interventions $\boldsymbol{\mathcal{I}}^i$ by constructing an example. We use the notation $\boldsymbol{\mathcal{P}}(\mathcal{I})$ to denote the primitive decomposition of $\mathcal{I}$ as shown in the proof of Theorem 2. Then, $\boldsymbol{\mathcal{P}}^i = \cup_{\mathcal{I} \in \boldsymbol{\mathcal{I}}^i} \boldsymbol{\mathcal{P}}(\mathcal{I})$ are the primitive interventions equivalent to each agent's visible interventions. Consider an arbitrary ordering of $\boldsymbol{\mathcal{P}}^i = \mathcal{P}^i_0, \ldots, \mathcal{P}^i_k$.

Let $\mathcal{Q}_k^i$ denote the inverse of $\mathcal{P}_k^i$. The construction is as follows.

$$\begin{aligned}
&\boldsymbol{\mathcal{P}}_0 = \boldsymbol{\mathcal{P}}^0\\
&\boldsymbol{\mathcal{P}}_j = \mathcal{P}_k^j \circ \cdots \circ \mathcal{P}_0^j \circ \mathcal{Q}_0^{j-1} \circ \cdots \circ \mathcal{Q}_k^{j-1} \text{ for } j \in 1, \ldots, N\\
&A_0 = \emptyset\\
&A_j = \{j\} \text{ for } j \in 1, \ldots, N
\end{aligned}$$

where $\boldsymbol{\mathcal{P}}^0$ are the fully pre-policy interventions visible to all agents. In this construction, A_j are singleton sets (except A_0) so we have $m = N$. For all i, let $j = i$. Then, we make an inductive argument. The base case $\boldsymbol{\mathcal{I}}^0(\mathcal{M}) = \boldsymbol{\mathcal{P}}^0(\mathcal{M}) = \boldsymbol{\mathcal{P}}_0(\mathcal{M})$ holds by definition of $\boldsymbol{\mathcal{P}}^0$. Assuming $\boldsymbol{\mathcal{I}}^{i-1}(\mathcal{M}) = (\boldsymbol{\mathcal{P}}_{j-1} \circ \boldsymbol{\mathcal{P}}_{j-2} \circ \cdots \circ \boldsymbol{\mathcal{P}}_0)(\mathcal{M})$, we have

$$\begin{aligned}
\boldsymbol{\mathcal{I}}^i(\mathcal{M}) &= \boldsymbol{\mathcal{P}}^i(\mathcal{M})\\
&= (\boldsymbol{\mathcal{P}}^i \circ \boldsymbol{\mathcal{Q}}^{i-1} \circ \boldsymbol{\mathcal{P}}^{i-1})(\mathcal{M})\\
&= (\boldsymbol{\mathcal{P}}_j \circ \boldsymbol{\mathcal{P}}^{i-1})(\mathcal{M})\\
&= (\boldsymbol{\mathcal{P}}_j \circ \boldsymbol{\mathcal{I}}^{i-1})(\mathcal{M})\\
&= (\boldsymbol{\mathcal{P}}_j \circ \boldsymbol{\mathcal{P}}_{j-1} \ldots \circ \boldsymbol{\mathcal{P}}_0)(\mathcal{M})
\end{aligned}$$

So, this assignment of primitive intervention sets satisfies Property 6.1. Algorithm 1 explicitly shows how this construction is used to calculate interventional queries. □

Taking this decomposition, we uniquely partition the agents into sets $A_0, \ldots, A_m$ according to the state of the game visible to them. The `decompose` function maps $\mathcal{I}$ to sets $\boldsymbol{\mathcal{P}}_0, \ldots, \boldsymbol{\mathcal{P}}_m$ satisfying property 6.1 in Theorem 3 and the corresponding partition of the agents $A_0, \ldots, A_m$. Then, Algorithm 1 solves the interventional query by iteratively calculating the $\mathcal{R}$-rational outcomes (e.g., NEs if $\mathcal{R} = \mathcal{R}^{\text{BR}}$), fixing the policies of agents who cannot observe future interventions, and applying interventions. This subsumes our earlier definition of pre-policy and post-policy interventional queries. Whenever the rationality assumptions have a solution existence guarantee (e.g., if $\mathcal{R} = \mathcal{R}^{\text{BR}}$, there is always at least one NE of the game), then Algorithm 1 successfully terminates. There are two special cases:

1. If $\boldsymbol{\mathcal{P}}_0 \cup \ldots \cup \boldsymbol{\mathcal{P}}_j = \emptyset$, the agent is not privy to any interventions and the interventional query is fully *post-policy* with respect to Agent i.

2. If $\boldsymbol{\mathcal{P}}_0 \cup \ldots \cup \boldsymbol{\mathcal{P}}_j = \mathcal{I}$, the agent is privy to all interventions and the interventional query is fully *pre-policy* with respect to Agent i.

Algorithm 1 Calculate the result of an arbitrary interventional query

Input: A causal game $\mathcal{M}$ with rationality relation $\mathcal{R}$, interventions $\boldsymbol{\mathcal{I}} = \boldsymbol{\mathcal{I}}^1 \cup \ldots \cup \boldsymbol{\mathcal{I}}^N$, and query $\phi(\boldsymbol{\pi})$.
$(\boldsymbol{\mathcal{P}}_0, \ldots, \boldsymbol{\mathcal{P}}_m), (A_0, \ldots, A_m) \leftarrow \textsc{decompose}(\boldsymbol{\mathcal{I}})$
$A' \leftarrow \emptyset$
for $j = 0, \ldots, m$ **do**
    $A \leftarrow (A_j \cup \ldots \cup A_N) \setminus A'$
    $\hat{\boldsymbol{\pi}} \leftarrow$ uniformly sample an $\mathcal{R}$-rational outcome
    for $i \in A_j$ **do**
        $\mathrm{do}(\Pi_{D^i} = \hat{\pi}_{D^i})$
    for $\mathcal{P} \in \boldsymbol{\mathcal{P}}_j$ **do**
        $\mathcal{P}(\mathcal{M})$
        if $\mathcal{P}$ acts on $V \in \boldsymbol{D}^i \cup \boldsymbol{\Pi}^i$ **then**
            $A' \leftarrow A' \cup \{i\}$
$\Pr^{\boldsymbol{\pi}}(\boldsymbol{v}) \leftarrow \prod_{V \in \boldsymbol{V}} \Pr^{\boldsymbol{\pi}}(v \mid \mathbf{pa}_V)$
return $\phi(\boldsymbol{\pi})$

Mechanism-level side effects: Object-level interventions can have unintuitive mechanism-level side effects. A side effect is a modification to the inter-mechanism edges in $\mathsf{m}\mathcal{G}$ and $\not\rightarrow$ denotes an edge removal. Proposition 20 formalises the side effects of an object-level intervention.

Definition 35 (Reachability Path). *Let $D \in \boldsymbol{D}^i$. We write $\mathcal{R}(\mathsf{M}_V \to \Pi_D)$ to denote the set of paths that make M_V $\mathcal{R}$-relevant to Π_D. A reachability path is any path $p \in \mathcal{R}(\mathsf{M}_V \to \Pi_D)$. That is, a non-repeating sequence of nodes $V_0, ..., V_j \in \mathsf{m}_\perp \boldsymbol{V}$ of the independent mechanised graph $\mathsf{m}_\perp\mathcal{G}$ s.t $V_0 = \mathsf{M}_V$, and V_0 is $\mathcal{R}$-relevant to Π_D.*

Proposition 20 (Object-level intervention side effects). *An object-level intervention $\Pr^{\mathcal{I}}(x \mid \boldsymbol{pa}^*_X)$ has side effect $\mathsf{M}_V \not\rightarrow \Pi_D$ if, $\forall$ reachability paths $p \in \mathcal{R}(\mathsf{M}_V \to \Pi_D)$ we have $\exists W \in \boldsymbol{V}$ s.t.$(W \notin \boldsymbol{Pa}^*_X)$ and $((W \to X) \in p)$*

Proof. An intervention $\Pr^{\mathcal{I}}(x \mid \mathbf{pa}^*_X)$ has side effect $\mathsf{M}_V \not\to \Pi_D$ if in the intervened graph there are no reachability paths from M_V to Π_D. This means at least one causal arrow is broken in each such reachability path in the original graph. An object-level intervention on X breaks only the causal arrows $W \to X$ where $W \in \mathbf{Pa}_X$ but $W \notin \mathbf{Pa}^*_X$. □

That is to say, an intervention on X which severs at least one edge critical to each reachability path between M_V and Π_D through X, will delete the corresponding edge between those mechanisms in $\mathsf{m}\mathcal{G}$. Similarly, if an intervention creates at least one new reachability path, it will result in the addition of a new inter-mechanism edge.

Minimum intervention sets: Using these observations, we formalise the minimum set of interventions required to break a causal mechanism dependency $\mathsf{M}_V \to \Pi_D$. Since we are only interested in interventions that do not directly modify the target policy Π_D, and we know that reachability paths are calculated on the independent mechanised graph $\mathsf{m}_\perp\mathcal{G}$ which contains no edges of the form $\mathsf{M}_V \to \Pi_D$, we can restrict our attention to object-level interventions. Then, the minimum intervention set is the intersection across all reachability paths, of the variables with incoming edges that would break the dependency if removed.

Definition 36 (Minimum intervention set)**.** *The minimum set of objects to intervene on in order to break causal mechanism dependency* $\mathsf{M}_V \to \Pi_D$ *is:*

$$\bigcap_{p \in \mathcal{R}(\mathsf{M}_V \to \Pi_D)} \{X \in \boldsymbol{V} \mid \exists W \in \boldsymbol{V}.(W \to X) \in p\}$$

This metric measures how *robust* a causal mechanism dependency is to external intervention. The size of this set is the minimum number of object-level interventions required to ensure that, under every parameterization of the game, there is no incentive for target policy Π_D to depend on the mechanism variable M_V.

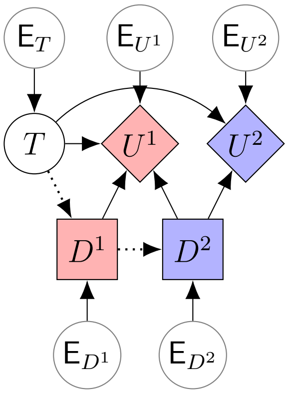
6.5 Counterfactuals

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 $\Pr^{\boldsymbol{\pi}}(V \mid \mathbf{Pa}_V)$, where $\mathsf{E}_V \in \mathbf{Pa}_V$.

Definition 37. *A (Markovian)* ***structural causal game (SCG)*** $\mathcal{M} = (\mathcal{G}, \boldsymbol{\theta})$ *is a causal game over exogenous and endogenous variables* $\mathsf{E} \cup \boldsymbol{V}$ *such that for* any *(deterministic) parameterisation of the decision variable CPDs* $\dot{\boldsymbol{\pi}}$*, the induced model with joint distribution* $\Pr^{\dot{\boldsymbol{\pi}}}(\boldsymbol{V}, \mathsf{E})$ *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 [44].

When we mechanise SCGs, although we introduce mechanism variables for the exogenous variables (as can be seen in Fig. 6.3b), 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 $\mathbf{Pa}^*_V$ [69]. With these caveats in place, both pre-policy and post-policy predictions and interventions may be defined in this model as described in Section 6.2 and Section 6.3 respectively.



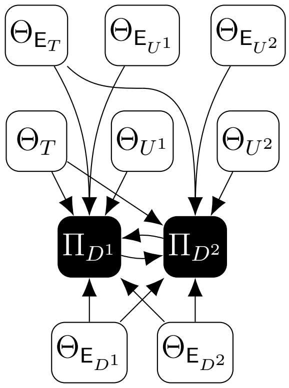


Figure 6.3: (a) A (Markovian) SCG representing Example 1. Note that we have included exogenous variables for U^1 and U^2, although as neither is stochastic, this is not strictly necessary. (b) The $\mathcal{R}^{\text{BR}}$-relevance graph of this game.

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.

6.5.1 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 of 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 [124]. 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 [31].

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 $\overline{\mathbf{Pa}}_D \coloneqq \mathbf{Pa}_D \setminus \{\mathsf{E}_D\}$.

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:

$$\begin{aligned}\Pr(\mathsf{E}_D^{\pi_D,\overline{\mathbf{pa}}_D} = d) &\coloneqq \pi_D(d \mid \overline{\mathbf{pa}}_D),\\ \dot{\pi}_D(D = d \mid \overline{\mathbf{pa}}_D, \mathsf{e}_D) &\coloneqq \delta(D, \mathsf{e}_D^{\pi_D,\overline{\mathbf{pa}}_D}),\end{aligned} \tag{6.2}$$

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 $\mathsf{E}_D^{\pi_D,\overline{\mathbf{pa}}_D}$ [125]. Proposition 21 below then follows immediately,[2] 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 21. *For distributions over E_D and D as governed by equations (6.2), 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 $\int_{dom(\mathsf{E}_D)} \dot{\pi}_D(d \mid \overline{\boldsymbol{pa}}_D, \mathsf{e}_D) \Pr(\mathsf{e}_D)\, d\mathsf{e}_D = \pi_D(d \mid \overline{\boldsymbol{pa}}_D)$.*

Proof. Recall from Section 6.5 that we define $\overline{\mathbf{Pa}}_D \coloneqq \mathbf{Pa}_D \setminus \{\mathsf{E}_D\}$. 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:

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

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

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 [125], we have that:

$$
\begin{aligned}
&\int_{dom(\mathsf{E}_D)} \dot{\pi}_D(d \mid \overline{\mathbf{pa}}_D, \mathsf{e}_D) \Pr(\mathsf{e}_D)\, d\mathsf{e}_D \\
=&\int_{dom(\mathsf{E}_D)} \delta(d,\mathsf{e}_D^{\pi_D,\overline{\mathbf{pa}}_D}) \Pr(\mathsf{e}_D)\, d\mathsf{e}_D \\
=&\sum_{\mathsf{e}_D^{\pi_D,\overline{\mathbf{pa}}_D} \in dom(\mathsf{E}_D^{\pi_D,\overline{\mathbf{pa}}_D})} \delta(d,\mathsf{e}_D^{\pi_D,\overline{\mathbf{pa}}_D}) \Pr(\mathsf{e}_D^{\pi_D,\overline{\mathbf{pa}}_D}) \cdot \int_{dom(\mathsf{E}_D^{-\pi_D,\overline{\mathbf{pa}}_D})} \Pr(\mathsf{e}_D^{-\pi_D,\overline{\mathbf{pa}}_D})\, d\mathsf{e}_D^{-\pi_D,\overline{\mathbf{pa}}_D} \\
=&\sum_{\mathsf{e}_D^{\pi_D,\overline{\mathbf{pa}}_D} \in dom(\mathsf{E}_D^{\pi_D,\overline{\mathbf{pa}}_D})} \delta(d,\mathsf{e}_D^{\pi_D,\overline{\mathbf{pa}}_D})\pi_D(\mathsf{e}_D^{\pi_D,\overline{\mathbf{pa}}_D} \mid \overline{\mathbf{pa}}_D) \cdot 1 \\
=&\pi_D(d \mid \overline{\mathbf{pa}}_D).
\end{aligned}
$$

□

6.5.2 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');

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 [39], 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 next subsection, we employ the simplicity principle above, which is also more in keeping with prior work[3], but

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

in the following subsection we provide an algorithm for computing $\mathbf{\Pi}(\mathcal{I})$ according to the closest possible world principle.

6.5.3 Defining Counterfactuals in Games

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

Definition 38. *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:*

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

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

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

is the set of ***actual-counterfactual rational outcomes****, and* $\mathbf{\Pi}(\mathcal{I})$ *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.,* $\mathbf{\Pi}(\mathcal{I}) = \mathbf{\Pi}$ *and* $\mathcal{R}(\mathsf{m}\mathcal{M}_{\mathcal{I}} \mid \boldsymbol{z}) = \Big\{ (\boldsymbol{\pi}, \boldsymbol{\pi}) : \boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z}) \Big\}$ *when* $\boldsymbol{Y} \subseteq \boldsymbol{V}$*.*

In the definition of $\Pr^{\mathcal{R}}(\boldsymbol{x}_{\mathcal{I}} \mid \boldsymbol{z})$, for every actual policy $\boldsymbol{\pi}$ and counterfactual policy $\boldsymbol{\pi}'$, we compute the product of three quantities. $\Pr^{\boldsymbol{\pi}}(\mathbf{e} \mid \boldsymbol{z})$ 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 $\Pr(\mathbf{e}^*)$.[4] 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 $\Pr^{\boldsymbol{\pi}'}(\boldsymbol{x}_{\mathcal{I}} \mid \mathbf{e}, \mathbf{e}^*)$. 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

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

$\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M} \mid \boldsymbol{z})$ 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 $\Pr^{\boldsymbol{\pi}}(u_g^1 \mid \neg g)$. 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 $\Pr(\mathbf{e})$ to $\Pr^{\boldsymbol{\pi}}(\mathbf{e} \mid \neg g)$. In this case, such an update amounts to changing $\Pr^{\boldsymbol{\pi}}(\mathsf{e}_T, \mathsf{e}_{D^1}, \mathsf{e}_{D^2})$ to $\Pr^{\boldsymbol{\pi}}(\mathsf{e}_T, \mathsf{e}_{D^1}, \mathsf{e}_{D^2} \mid \neg g)$, 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 $\Pr^{\mathcal{R}}(u_g^1 \mid \neg g) = \Big\{ \Pr^{\boldsymbol{\pi}}(u_g^1 \mid \neg g)\Big\}_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M} \mid \neg g)}$. We thus have $\Big\{\mathbb{E}_{\boldsymbol{\pi}}[U_g^1 \mid \neg g]\Big\}_{\boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M} \mid \neg g)} = \{\frac{7}{2}\}$.

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., $\mathcal{R}(\mathsf{m}\mathcal{M}_{\mathcal{I}} \mid \tilde{\pi}_{D^1}) = \Big\{(\boldsymbol{\pi}', \boldsymbol{\pi}') : \boldsymbol{\pi} \in \mathcal{R}(\mathsf{m}\mathcal{M}_{\mathcal{I}})\Big\}$. 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. *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 [126]. The 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.*

6.5.4 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 $\mathbf{\Pi}(\mathcal{I})$ as large as possible while propagating changes due to $\mathcal{I}$.

As a first attempt, we might set $\mathbf{\Pi}(\mathcal{I})$ to $\mathbf{\Pi} \setminus (\boldsymbol{Y} \cup \mathbf{Desc}_{\boldsymbol{Y}})$. 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 Π'_D does not change – and in which case there is no sense in which Π'_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.[5] 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 [26] – 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:

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

to denote the rational responses $c_j = (\pi_D)_{\Pi_D \in C_j}$. Then we may compute $\mathbf{\Pi}(\mathcal{I})$ using Algorithm 2.

Algorithm 2 functions by incrementally expanding the set $\mathbf{\Pi}(\mathcal{I})$ – which is at first initialised to $\mathbf{\Pi} \setminus (\boldsymbol{Y} \cup \mathbf{Desc}_{\boldsymbol{Y}})$ – 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

[5] 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.

Algorithm 2

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

intervention $\mathcal{I}$ is performed, i.e., if and only if $\Pi_D \in \boldsymbol{\Pi}(\mathcal{I})$. We may then use this set $\boldsymbol{\Pi}(\mathcal{I})$ in Definition 38 to define counterfactual queries in games under this principle.

6.6 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 [127], but we also explain some of the shortcomings of this approach relative to causal games.

6.6.1 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^i_j with an arbitrary probability distribution $P^*_j(V_j \mid \cap_{V_j \in J} \mathbf{Anc}_{V_j})$ or ${\sigma^i_j}^*(V_j \mid \cap_{V_j \in J} \mathbf{Anc}_{V_j})$, 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 $P^{\sigma}_{\mathcal{I}}$ 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.

6.6.2 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 [127]. 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 $\sigma^i_Z(z) = 1$, 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 [127], 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 [3, 128].

6.7 Conclusions

In this chapter, we generalised Pearl's causal hierarchy of models to the game-theoretic domain by introducing (structural) causal games. We then described 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 (the rationality relations introduced in Section 3.3) that agents use to play the game. In the next section, we will showcase some applications that have been built from this work.

7

Conclusions

Contents

7.1 Summary

In this thesis, we have introduced causal games as a unifying formalism for answering causal queries in games. We have also significantly generalised the theory behind MAIDs by introducing subgames and equilibrium refinements, demonstrating how to handle imperfect recall, proving existence conditions for different types of NE, and establishing crucial computational complexity results. Various research communities have long considered combinations of causal and game-theoretic reasoning. So we conclude with a brief summary of the relative advantages and disadvantages of causal games compared to other models, before offering some short final thoughts about promising future directions for building upon this work.

7.2 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 1.2, we expand briefly upon this claim by considering other causal and game-theoretic models in light of this thesis' contributions.

Causal Models

Standard causal models do not consider 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 [4], represented by the edges between mechanism variables in mechanised games. Alternative causal models that do include agents, such as CIDs [8, 30, 31], typically only consider a single agent.

To the best of our knowledge, the only exception to this rule is settable systems [46, 47]. However, 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 thesis. 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 [38, 39] to the relational setting [41, 42, 44]. However, the cyclic dependencies in causal games are of a specific form because 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, equilibrium refinements, and correlated equilibria, which are critical for game-theoretic reasoning and are not supported by other causal models. Such concepts have been historically underexplored in the context of graphical games compared to EFGs or strategic-form games.

Game-Theoretic Models

The most natural game-theoretic model to compare causal games with is EFGs, which are tree-based models instead of DAG-based. The most important benefit of DAG-based models is that they can be more readily used to support a wide range of causal queries natively. 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.[1] Mechanised games can define a wide range of pre- and post-policy queries in games. In contrast, as explained in Section 6.6, 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. For example, we saw in Chapter 5 that this is crucial for explicitly representing the assumptions made by behavioural, mixed, and correlated policies.

Moreover, a DAG-based representation of a game need never be larger 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.

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

In addition to their transparency, MAIDs allow us to exploit the conditional independencies between variables further using d-separation. Using $\mathcal{R}$-reachability, we may construct the $\mathcal{R}$-minimal mechanised 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 empirically in Section 4.5 and theoretically in Section 5.6), 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 [129].[2]

Finally, although 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, EFGs are likely to be a more suitable model. We hope further research on MAIDs and causal games will reduce this last difference.

7.3 Future Work

Our priority is to use causal games to help designers build safe AI systems. Existing work has already built upon the framework introduced in this thesis [19–22]. We hope that more will follow. The theoretical framework can still be extended further – especially to capture dynamic settings, fine-grained subjective beliefs, or optimisation. Given that we propose this thesis 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)

[2]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.

make them an attractive choice for researchers and practitioners interested in the intersection of causality and game theory.

Appendices

A
Appendices

Contents

A.1 Transformations between Game Representations

A.1.1 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 whether one wants to keep a record of the structure of the original MAID or if one wants to optimise the encoding by splitting 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 $\boldsymbol{S} = \boldsymbol{X} \cup \boldsymbol{D}$.[1] If instead, however, one wishes to minimise the complexity of

[1]We assume that $\mathbf{desc}_{\boldsymbol{U}} = \varnothing$.

computing equilibria, then one only needs to split on the MAID's decision variables and their parents, $\boldsymbol{S} = \mathbf{Fa}_{\boldsymbol{D}}$ [63]. 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 [26]. 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 $S_j \prec S_k$.[2] 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:

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

[2] 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^i_j for agent i at node V^i_j in the EFG is given by the domain of the corresponding decision variable $D = S_{V^i_j}$ in the MAID, i.e. $A^i_j \coloneqq dom(D)$.

- Define $\lambda : \boldsymbol{V} \times \boldsymbol{V} \to A$ for each decision node $V \in \boldsymbol{V}^1 \cup \cdots \cup \boldsymbol{V}^n$ 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 $\boldsymbol{V}^1 \cup \cdots \cup \boldsymbol{V}^n$ 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 $\boldsymbol{V}^i / \sim$ – the set of $\sim$ equivalence classes partitioning $\boldsymbol{V}^i$. In other words, two nodes $V, V' \in \boldsymbol{V}$ which correspond to the same decision variable S in $\mathcal{M}$ are in the same information set I^i_j for agent i if and only if their instantiation labels, restricted to their parents in $\mathcal{M}$, are the same.

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

A.1.2 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 $(\boldsymbol{V}, \mathscr{E})$ as T, where edges are directed from parents to children.

- For each of $\mathcal{E}$'s chance nodes $V \in \boldsymbol{V}^0$, 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 $V \in \boldsymbol{V}$ 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^i_j in $\mathcal{E}$, let $\rho^{\prec}(I^i_j)$ be the set of paths from R into the nodes of I^i_j. Next, define a function $\mu : \bigcup_{i \in N} I^i \to 2^{\boldsymbol{V}}$ that maps each information set I^i_j to the set of variables in $\mathbf{Anc}_{I^i_j}$ whose outgoing label is the same in every path in $\rho^{\prec}(I^i_j)$. Note that by Definition 21, 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 $V \in \boldsymbol{V}^0$, add the label $(v_j \mid \rho_{V_j}) : p$ where $p = P_j(v_j)$.

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

  - For each leaf variable $L \in \boldsymbol{L}$ 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^i_j into a single variable $D_j \in \boldsymbol{D}^i$, collecting the labels $(v \mid \rho_V[\mu(I^i_j)]) : \_$ for each $V \in I^i_j$ 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^i_j, remove all the incoming edges from variables not in $\mu(I^i_j)$.
- Merge each group of variables, collecting their incoming and outgoing labelled edges that belong to the same intervention set.
- Merge any utility variables U^i_j and U^i_k belonging to the same agent i that have the same sets of incoming edges and collect their labels.

- We then let $\boldsymbol{V} \coloneqq \boldsymbol{X} \cup \boldsymbol{D} \cup \boldsymbol{U}$ 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 4.4.2). This is dealt with for non-decision variables 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 $\Pr(v \mid \mathbf{pa}_V) \coloneqq \sum_{(v \mid \mathbf{pa}_V):p \, \in \, l(V)} p$.
    * 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 $(d \mid \rho_D[\mu(I^i_j)]) : \_$.

A.2 Codebase

In this appendix, we briefly describe PyCID [60], 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 levels one and two of the causal hierarchy. For the contributions of this thesis, 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 [8] and reasoning patterns in MAIDs [25]. This makes the codebase well-suited as a testbed for future research and applications. We refer the reader to an existing tool paper [60] and our online codebase for further up-to-date details, including several tutorials.

A.2.1 Creating MAIDs

Listings A.1 and A.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 [130]. 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 A.1a and A.1b) 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 causal games. Therefore, both can be defined as `MAID` objects using PyCID. In the case of MAIDs, a level one model, all class methods except those involving causal interventions are permitted.

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, with multiple ways to define them. Listing A.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.

```python
import pycid

job_market = pycid.MAID(
    [
        ("T", "D1"),
        ("T", "U1"),
        ("T", "U2"),
        ("D1", "D2"),
        ("D1", "U1"),
        ("D2", "U1"),
        ("D2", "U2"),
    ],
    agent_decisions={
        1: ["D1"],
        2: ["D2"],
    },
    agent_utilities={
        1: ["U1"],
        2: ["U2"],
    },
)

job_market.draw()

prob = 1/2

job_market.add_cpds(
    T = pycid.bernoulli(prob), # T = 1 corresponds to T = h (hard-working)
    D1 = [0,1], # D1 = 1 corresponds to D1 = g (going to university)
    D2 = [0,1], # D2 = 1 corresponds to D2 = j (offering a job)
    U1 = lambda d1, d2, t: 5*d2 - t*d1 - 2*d1*(1-t),
    U2 = lambda d2, t: 3*t*d2 - 2*(1-t)*d2 - (1-d2)*t,
)
```

Listing A.1: An instantiation of Example 1 in PyCID.

Listing A.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 Chapter 4; 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 $D^1 = q$.

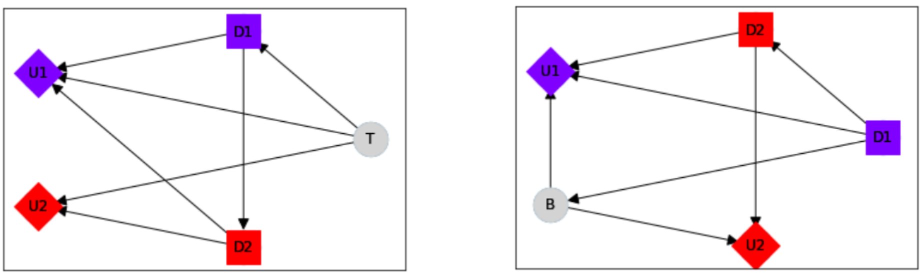
```python
warehouse_robots = pycid.MAID(
```

```
    [
        ("D1", "D2"),
        ("D1", "U1"),
        ("D1", "B"),
        ("B", "U2"),
        ("B", "U1"),
        ("D2", "U2"),
        ("D2", "U1"),
    ],
    agent_decisions={
        1: ["D1"],
        2: ["D2"],
    },
    agent_utilities={
        1: ["U1"],
        2: ["U2"],
    },
)

warehouse_robots.draw()

warehouse_robots.add_cpds(
    B = lambda d1 : {0: 1 if d1==0 else 2/3, 1: None}, # B = 1 corresponds to B = b (breaking something)
    D1 = [0,1], # D1 = 1 corresponds to D1 = q (moving quickly)
    D2 = [0,1], # D2 = 1 corresponds to D2 = p (patrolling)
    U1 = lambda d1, d2, b: (1 - 0.5*d2)*((1-d1)*2 + d1*5 - 3*b),
    U2 = lambda d2, b: 6*(1 - (1-d2)*b) - d2,
)
```

Listing A.2: An instantiation of Example 2 in PyCID.



(a) **(b)**

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

A.2.2 Computing Equilibria

Our open-source Python codebase, PyCID, finds all pure NEs and SPEs in a MAID directly and finds behavioural NEs and SPEs in N-agent games by employing a number of algorithms implemented in PyGambit [90]. We refer the interested reader to our online codebase `https://github.com/causalincentives/pycid` for up-to-date syntax showing how to compute NEs. All SPEs in a MAID are found by adapting Algorithm 6.2 from [26] 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 Proposition 11 in Section 5.4.1. In contrast to K&M's algorithm, our implementation finds almost all SPEs instead of just one. Specifically, at the time of writing, PyCID is guaranteed to compute all pure SPE in any MAID, all NE in MAIDs where the smallest proper subgame contains at most 2 agents, and at least 1 NE in arbitrary N-agent games.

References

[1] Peter Spirtes, Clark Glymour, and Richard Scheines. *Causation, Prediction, and Search*. New York, USA: Springer, 1993.

[2] Donald B Rubin. "Causal Inference Using Potential Outcomes". In: *Journal of the American Statistical Association* 100.469 (Mar. 2005), pp. 322–331.

[3] Judea Pearl. *Causality*. Cambridge, UK: Cambridge University Press, 2009.

[4] Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. *Elements of Causal Inference: Foundations and Learning Algorithms*. Cambridge, Massachusetts: The MIT Press, 2017.

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

[6] Noam Nisan et al., eds. *Algorithmic Game Theory*. Cambridge University Press, 2007.

[7] Yoav Shoham and Kevin Leyton-Brown. *Multiagent Systems: Algorithmic, Game-theoretic, and Logical Foundations*. Cambridge University Press, 2008.

[8] Tom Everitt et al. "Agent Incentives: A Causal Perspective". In: *Proceedings of the Thirty-fifth AAAI Conference on Artificial Intelligence*. 2021, pp. 11487–11495. URL: `https://ojs.AAAI.org/index.php/AAAI/article/view/17368`.

[9] Eric D. Langlois and Tom Everitt. "How RL Agents Behave When Their Actions Are Modified". In: *Proceedings of the Thirty-fifth AAAI Conference on Artificial Intelligence*. 2021, pp. 11586–11594. URL: `https://ojs.AAAI.org/index.php/AAAI/article/view/17378`.

[10] Ryan Carey and Tom Everitt. "Human control: definitions and algorithms". In: *Uncertainty in Artificial Intelligence*. PMLR. 2023, pp. 271–281.

[11] Chris van Merwijk, Ryan Carey, and Tom Everitt. "A Complete Criterion for Value of Information in Soluble Influence Diagrams". In: *Proceedings of the AAAI Conference on Artificial Intelligence* 36.9 (June 2022), pp. 10034–10041.

[12] Sebastian Farquhar, Ryan Carey, and Tom Everitt. "Path-specific objectives for safer agent incentives". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 36. 9. 2022, pp. 9529–9538.

[13] Tom Everitt et al. "Reward tampering problems and solutions in reinforcement learning: A causal influence diagram perspective". In: *Synthese* 198.Suppl 27 (2021), pp. 6435–6467.

[14] Niki Kilbertus et al. "Avoiding Discrimination through Causal Reasoning". In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. 2017, pp. 656–666.

[15] Matt Kusner et al. “Counterfactual Fairness”. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. 2017, pp. 4069–4079.

[16] Junzhe Zhang and Elias Bareinboim. “Fairness in Decision-making - the Causal Explanation Formula”. In: *Proceedings of the Thirty-second AAAI Conference on Artificial Intelligence*. 2018, pp. 2037–2045. URL: https://www.AAAI.org/ocs/index.php/AAAI/AAAI18/paper/view/16949.

[17] Razieh Nabi and Ilya Shpitser. “Fair Inference on Outcomes”. In: *Proceedings of the Thirty-second AAAI Conference on Artificial Intelligence*. 2018, pp. 1931–1940. URL: https://www.AAAI.org/ocs/index.php/AAAI/AAAI18/paper/view/16683.

[18] Carolyn Ashurst et al. “Why Fair Labels Can Yield Unfair Predictions: Graphical Conditions for Introduced Unfairness”. In: *Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence*. 2022, pp. 9494–9503.

[19] Francis Rhys Ward et al. “Honesty is the best policy: defining and mitigating AI deception”. In: *Thirty-seventh Conference on Neural Information Processing Systems*. 2023.

[20] Zachary Kenton et al. “Discovering agents”. In: *Artificial Intelligence* 322 (2023), p. 103963.

[21] Jonathan Richens and Tom Everitt. “Robust agents learn causal world models”. In: *Proceedings of the Twelfth International Conference on Learning Representations* (2024).

[22] Francis Rhys Ward et al. “The reasons that agents act: Intention and instrumental goals”. In: *arXiv preprint arXiv:2402.07221* (2024).

[23] Joseph Y. Halpern and Max Kleiman-Weiner. “Towards Formal Definitions of Blameworthiness, Intention, and Moral Responsibility”. In: *Proceedings of the Thirty-second AAAI Conference on Artificial Intelligence*. 2018, pp. 1853–1860. URL: https://www.AAAI.org/ocs/index.php/AAAI/AAAI18/paper/view/16824.

[24] Meir Friedenberg and Joseph Y. Halpern. “Blameworthiness in Multi-agent Settings”. In: *Proceedings of the Thirty-third AAAI Conference on Artificial Intelligence*. 2019, pp. 525–532.

[25] Avi Pfeffer and Ya’akov Gal. “On the Reasoning Patterns of Agents in Games”. In: *Proceedings of the 22nd National Conference on Artificial Intelligence*. 2007, pp. 102–109.

[26] Daphne Koller and Brian Milch. “Multi-agent Influence Diagrams for Representing and Solving Games”. In: *Games and Economic Behavior* 45.1 (2003), pp. 181–221.

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

[28] Michael Spence. “Job Market Signaling”. In: *The Quarterly Journal of Economics* 87.3 (1973), p. 355.

[29] Michael Maschler, Shmuel Zamir, and Eilon Solan. *Game theory*. Cambridge University Press, 2020.

[30] Ronald A. Howard and James E. Matheson. "Influence Diagrams". In: *Decision Analysis* 2.3 (2005), pp. 127–143.

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

[32] David Heckerman and Ross Shachter. "A Decision-based View of Causality". In: *Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence*. 1994, pp. 302–310.

[33] Michael Kearns. "Graphical Games". In: *Algorithmic Game Theory*. Ed. by Noam Nisan et al. Cambridge University Press, 2007. Chap. 7, pp. 159–180.

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

[35] Prashant Doshi, Yifeng Zeng, and Qiongyu Chen. "Graphical Models for Interactive POMDPs: Representations and Solutions". In: *Autonomous Agents and Multi-agent Systems* 18.3 (Sept. 2008), pp. 376–416.

[36] Kyle Polich and Piotr Gmytrasiewicz. "Interactive Dynamic Influence Diagrams". In: *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems*. 2007, pp. 1–3.

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

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

[39] Stephan Bongers et al. "Foundations of Structural Causal Models with Cycles and Latent Variables". In: *The Annals of Statistics* 49.5 (2021), pp. 2885–2915.

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

[41] Lise Getoor et al. "Probabilistic Relational Models". In: *Introduction to Statistical Relational Learning*. Ed. by Lise Getoor and Ben Taskar. The MIT Press, 2007.

[42] Marc Maier, Katerina Marazopoulou, and David Jensen. "Reasoning about Independence in Probabilistic Models of Relational Data". In: *arXiv:1302.4381* (2013).

[43] Fabio G. Cozman. "Credal Networks". In: *Artificial Intelligence* 120.2 (2000), pp. 199–233.

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

[45] Mauricio Gonzalez Soto, Luis E Sucar, and Hugo Jair Escalante. "Causal games and causal nash equilibrium". In: *Res Comput Sci* 149 (2020), pp. 123–133.

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

[47] Halbert White, Haiqing Xu, and Karim Chalak. "Causal Discourse in a Game of Incomplete Information". In: *Journal of Econometrics* 182.1 (Sept. 2014), pp. 45–58.

[48] Mogens Nielsen, Gordon Plotkin, and Glynn Winskel. "Petri Nets, Event Structures and Domains". In: *Semantics of Concurrent Computation*. Springer, 1979, pp. 266–284.

[49] Silvain Rideau and Glynn Winskel. "Concurrent Strategies". In: *2011 IEEE 26th Annual Symposium on Logic in Computer Science*. 2011.

[50] Glynn Winskel. "Distributed Probabilistic and Quantum Strategies". In: *Electronic Notes in Theoretical Computer Science* 298 (Nov. 2013), pp. 403–425.

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

[52] Valeria Zahoransky et al. "Partial Order Games". In: *Games* 13.1 (Dec. 2021), p. 2.

[53] Jonathan Richens, Rory Beard, and Daniel H Thompson. "Counterfactual harm". In: *Advances in Neural Information Processing Systems* 35 (2022), pp. 36350–36365.

[54] Natasha Jaques et al. "Social Influence As Intrinsic Motivation for Multi-agent Deep Reinforcement Learning". In: *Proceedings of the 36th International Conference on Machine Learning*. 2019, pp. 3040–3049.

[55] Micah Carroll et al. "Characterizing Manipulation from AI Systems". In: *arXiv preprint arXiv:2303.09387* (2023).

[56] Panagiotis Toulis and David C. Parkes. "Long-term Causal Effects Via Behavioral Game Theory". In: *Advances in Neural Information Processing Systems 29*. 2016, pp. 2604–2612. URL: `https://proceedings.neurips.cc/paper/2016/hash/af4732711661056eadbf798ba191272a-Abstract.html`.

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

[58] John von Neumann. "Zur Theorie Der Gesellschaftsspiele". In: *Mathematische Annalen* 100.1 (Dec. 1928), pp. 295–320.

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

[60] James Fox et al. "PyCID: A Python Library for Causal Influence Diagrams". In: *Proceedings of the 20th Python in Science Conference*. 2021, pp. 43–51.

[61] Robert J Aumann. "Subjectivity and correlation in randomized strategies". In: *Journal of mathematical Economics* 1.1 (1974), pp. 67–96.

[62] Bernhard Von Stengel and Françoise Forges. “Extensive-form correlated equilibrium: Definition and computational complexity”. In: *Mathematics of Operations Research* 33.4 (2008), pp. 1002–1022.

[63] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1988.

[64] Judea Pearl. “Causal inference in statistics: An overview”. In: (2009).

[65] Thomas Verma and Judea Pearl. “Causal Networks: Semantics and Expressiveness”. In: *Proceedings of the Fourth Annual Conference on Uncertainty in Artificial Intelligence*. 1990, pp. 69–78.

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

[67] Dan Geiger, Thomas Verma, and Judea Pearl. “Identifying Independence in Bayesian Networks”. In: *Networks* 20.5 (Aug. 1990), pp. 507–534.

[68] Francis Rhys Ward et al. “Defining deception in structural causal games (extended abstract)”. In: *Proceedings of the 22nd International Conference on Autonomous Agents and Multiagent Systems*. AAMAS ’23. International Foundation for Autonomous Agents and Multiagent Systems, 2023.

[69] Juan Correa and Elias Bareinboim. “A Calculus for Stochastic Interventions: Causal Effect Identification and Surrogate Experiments”. In: *Proceedings of the Thirty-fourth AAAI Conference on Artificial Intelligence*. 2020, pp. 10093–10100.

[70] Frederick Eberhardt and Richard Scheines. “Interventions and Causal Inference”. In: *Philosophy of Science* 74.5 (Dec. 2007), pp. 981–995. URL: `https://www.journals.uchicago.edu/doi/abs/10.1086/525638`.

[71] Kevin B Korb et al. “Varieties of Causal Intervention”. In: *PRICAI 2004: Trends in Artificial Intelligence*. Ed. by Chengqi Zhang, Hans W. Guesgen, and Wai-Kiang Yeap. Springer Berlin Heidelberg, 2004, pp. 322–331.

[72] Allen C Miller III et al. *Development of Automated Aids for Decision Analysis*. Tech. rep. 1976.

[73] Lewis Hammond et al. “Reasoning about causality in games”. In: *Artificial Intelligence* 320 (July 2023), p. 103919.

[74] Manuel Gómez. “Real-world applications of influence diagrams”. In: *Advances in Bayesian networks*. Springer, 2004, pp. 161–180.

[75] Uffe B Kjaerulff and Anders L Madsen. “Bayesian networks and influence diagrams”. In: *Springer Science+ Business Media* 200 (2008), p. 114.

[76] Tom Everitt et al. “Agent incentives: A causal perspective”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. 13. 2021, pp. 11487–11495.

[77] Owain Evans et al. “Truthful AI: Developing and governing AI that does not lie”. In: *arXiv preprint arXiv:2110.06674* (2021).

[78] Jonathan Richens, Rory Beard, and Daniel H Thompson. “Counterfactual harm”. In: *Advances in Neural Information Processing Systems* 35 (2022), pp. 36350–36365.

[79] Francis Rhys Ward, Francesca Toni, and Francesco Belardinelli. “On Agent Incentives to Manipulate Human Feedback in Multi-Agent Reward Learning Scenarios.” In: *AAMAS*. 2022, pp. 1759–1761.

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

[81] Ya’akov Gal and Avi Pfeffer. “Networks of Influence Diagrams: A Formalism for Representing Agents’ Beliefs and Decision-Making Processes”. In: *CoRR* abs/1401.3426 (2014). arXiv: 1401.3426. URL: http://arxiv.org/abs/1401.3426.

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

[83] Herbet A. Simon. *Models of Man. Social and Rational. Mathematical Essays on Rational Human Behavior in a Social Setting*. New York, USA: John Wiley & Sons, 1957.

[84] Constantinos Daskalakis, Aranyak Mehta, and Christos Papadimitriou. “A Note on Approximate Nash Equilibria”. In: *International Workshop on Internet and Network Economics*. Springer. 2006, pp. 297–306.

[85] Thomas D. Nielsen and Finn Verner Jensen. “Welldefined Decision Scenarios”. In: *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence*. 1999, pp. 502–511. URL: https://dslpitt.org/uai/displayArticleDetails.jsp?mmnu=1%5C&smnu=2%5C&article%5C_id=203%5C&proceeding%5C_id=15.

[86] Daphne Koller and Brian Milch. “Multi-agent Influence Diagrams for Representing and Solving Games”. In: *Proceedings of the 17th International Joint Conference on Artificial Intelligence*. 2001, pp. 1027–1034.

[87] Lewis Hammond et al. “Equilibrium Refinements for Multi-agent Influence Diagrams: Theory and Practice”. In: *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems*. 2021, pp. 574–582.

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

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

[90] Richard D. McKelvey, Andrew M. McLennan, and Theodore L. Turocy. *Gambit: Software Tools for Game Theory*. Version 0.2006. 01.20. 2016.

[91] Constantinos Daskalakis, Paul W Goldberg, and Christos H Papadimitriou. “The Complexity of Computing a Nash Equilibrium”. In: *Siam Journal on Computing* 39.1 (2009), pp. 195–259.

[92] Apiruk Detwarasiti and Ross D Shachter. “Influence diagrams for team decision analysis”. In: *Decision Analysis* 2.4 (2005), pp. 207–228.

[93] Michele Piccione and Ariel Rubinstein. “The Absent-minded Driver’s Paradox: Synthesis and Responses”. In: *Games and Economic Behavior* 20.1 (1997), pp. 121–130.

[94] Kevin Waugh et al. "A Practical Use of Imperfect Recall." In: *SARA*. 2009.

[95] Daphne Koller and Nir Friedman. *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

[96] Christos Papadimitriou. *Computational Complexity*. Addison Wesley, 1994.

[97] Robert J Aumann, Sergiu Hart, and Motty Perry. "The absent-minded driver". In: *Games and Economic Behavior* 20.1 (1997), pp. 102–116.

[98] Shizuo Kakutani. "A generalization of Brouwer's fixed point theorem". In: *Duke Mathematical Journal* 8.3 (Sept. 1941).

[99] Mamoru Kaneko and J Jude Kline. "Behavior strategies, mixed strategies and perfect recall". In: *International Journal of Game Theory* 24.2 (1995), pp. 127–145.

[100] Philipp C. Wichardt. "Existence of Nash Equilibria in Finite Extensive Form Games with Imperfect Recall: A Counterexample". In: *Games and Economic Behavior* 63.1 (2008), pp. 366–369. URL: `https://www.sciencedirect.com/science/article/pii/S0899825607001534`.

[101] In-Koo Cho and David M Kreps. "Signaling games and stable equilibria". In: *The Quarterly Journal of Economics* 102.2 (1987), pp. 179–221.

[102] Sergiu Hart and David Schmeidler. "Existence of correlated equilibria". In: *Mathematics of Operations Research* 14.1 (1989), pp. 18–25.

[103] Wan Huang and Bernhard von Stengel. "Computing an extensive-form correlated equilibrium in polynomial time". In: *International Workshop on Internet and Network Economics*. Springer. 2008, pp. 506–513.

[104] Hans L Bodlaender. "A linear time algorithm for finding tree-decompositions of small treewidth". In: *Proceedings of the twenty-fifth annual ACM symposium on Theory of computing*. 1993, pp. 226–234.

[105] Michael L Littman, Stephen M Majercik, and Toniann Pitassi. "Stochastic boolean satisfiability". In: *Journal of Automated Reasoning* 27.3 (2001), pp. 251–296.

[106] Valeria Zahoransky et al. "Partial order games". In: *Games* 13.1 (2021), p. 2.

[107] Denis Deratani Mauá, Cassio P de Campos, and Marco Zaffalon. "Solving limited memory influence diagrams". In: *Journal of Artificial Intelligence Research* 44 (2012), pp. 97–140.

[108] Denis Deratani Mauá and Fabio Gagliardi Cozman. "Fast local search methods for solving limited memory influence diagrams". In: *International Journal of Approximate Reasoning* 68 (2016), pp. 230–245.

[109] Steffen L Lauritzen and Dennis Nilsson. "Representing and solving decision problems with limited information". In: *Management Science* 47.9 (2001), pp. 1235–1251.

[110] Cassio P de Campos and Qiang Ji. "Strategy selection in influence diagrams using imprecise probabilities". In: *Proceedings of the Twenty-Fourth Conference on Uncertainty in Artificial Intelligence*. 2008, pp. 121–128.

[111] Hans L Bodlaender, Frank van den Eijkhof, and Linda C van der Gaag. "On the complexity of the MPA problem in probabilistic networks". In: *ECAI*. 2002, pp. 675–679.

[112] Solomon Eyal Shimony. "Finding MAPs for belief networks is NP-hard". In: *Artificial intelligence* 68.2 (1994), pp. 399–410.

[113] Dan Roth. "On the hardness of approximate reasoning". In: *Artificial Intelligence* 82.1-2 (1996), pp. 273–302.

[114] Johan Henri Petrus Kwisthout et al. *The computational complexity of probabilistic networks*. Utrecht University, 2009.

[115] Johan Kwisthout and Linda C van der Gaag. "The computational complexity of sensitivity analysis and parameter tuning". In: *Proceedings of the Twenty-Fourth Conference on Uncertainty in Artificial Intelligence*. 2008, pp. 349–356.

[116] James D Park and Adnan Darwiche. "Complexity results and approximation strategies for MAP explanations". In: *Journal of Artificial Intelligence Research* 21 (2004), pp. 101–133.

[117] Cassio Polpo De Campos and Fabio Gagliardi Cozman. "The inferential complexity of Bayesian and credal networks". In: *IJCAI*. Vol. 5. Citeseer. 2005, pp. 1313–1318.

[118] Lloyd S Shapley. "Stochastic games". In: *Proceedings of the national academy of sciences* 39.10 (1953), pp. 1095–1100.

[119] Luca De Alfaro and Rupak Majumdar. "Quantitative Solution of Omega-Regular Games". In: *Proceedings of the thirty-third annual ACM symposium on Theory of computing*. 2001, pp. 675–683.

[120] Eric Maskin and Jean Tirole. "Markov perfect equilibrium: I. Observable actions". In: *Journal of Economic Theory* 100.2 (2001), pp. 191–219.

[121] Krishnendu Chatterjee, Thomas A. Henzinger, and Nir Piterman. "Strategy Logic". In: *Information and Computation* 208.6 (2010). Special Issue: 18th International Conference on Concurrency Theory, pp. 677–693.

[122] Michael Wooldridge et al. "Rational Verification: From Model Checking to Equilibrium Checking". In: *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*. 2016, pp. 4184–4190.

[123] Marta Kwiatkowska et al. "Multi-player Equilibria Verification for Concurrent Stochastic Games". In: *Proceedings of the 17th International Conference on Quantitative Evaluation of SysTems*. Springer, 2020.

[124] Alexander Balke and Judea Pearl. "Probabilistic Evaluation of Counterfactual Queries". In: *Proceedings of the Twelfth AAAI National Conference on Artificial Intelligence*. 1994, pp. 230–237.

[125] Heinz Bauer. *Probability Theory*. Vol. 23. De Gruyter Studies in Mathematics. Translated by: Robert B. Burckel. De Gruyter, 1996.

[126] Chen Avin, Ilya Shpitser, and Judea Pearl. "Identifiability of Path-specific Effects". In: *Proceedings of the 19th International Joint Conference on Artificial Intelligence*. 2005, pp. 357–363.

[127] Tim Genewein et al. “Algorithms for Causal Reasoning in Probability Trees”. In: *arXiv:2010.12237* (Oct. 2020). arXiv: `http://arXiv.org/abs/2010.12237v1 [cs.AI]`.

[128] A. P. Dawid. “Causal Inference without Counterfactuals”. In: *Journal of the American Statistical Association* 95.450 (June 2000), pp. 407–424.

[129] Craig Boutilier et al. “Context-specific Independence in Bayesian Networks”. In: *Proceedings of the Twelfth International Conference on Uncertainty in Artificial Intelligence*. 1996, pp. 115–123.

[130] Ankur Ankan and Abinash Panda. “Pgmpy: Probabilistic Graphical Models Using Python”. In: *Proceedings of the 14th Python in Science Conference*. 2015, pp. 6–11.