Counterfactual Reasoning for Responsibility Attribution in Probabilistic Multi-Agent Systems

Chunyan Mu[1] , **Muhammad Najib**[2]
[1]University of Aberdeen
[2]Heriot-Watt University
{Chunyan.Mu@abdn.ac.uk, M.Najib@hw.ac.uk}

Abstract

Responsibility allocation—determining the extent to which agents causally contribute to outcomes—is a fundamental challenge in the design and analysis of multi-agent systems. In this work, we model such systems as concurrent stochastic multi-player games and introduce a notion of retrospective (backward) counterfactual responsibility, which quantifies an agent's causal contribution to outcomes resulting from a given strategy profile. To allocate responsibility among agents, we utilise the Shapley value and formally show that this method satisfies key desirable properties, including fairness and consistency. Building on this foundation, we propose a formal framework that supports both verification and strategic reasoning in responsibility-aware multi-agent systems. Furthermore, by adopting Nash equilibrium as the solution concept, we demonstrate how to compute stable strategy profiles in which agents trade off responsibility against expected reward.

1 Introduction

The growing emphasis on *trustworthy AI* [Chatila *et al.*, 2021; Li *et al.*, 2023] necessitates a paradigm shift, particularly regarding *responsibility* and *accountability* [Dignum, 2020; Kaur *et al.*, 2022; Belle *et al.*, 2024]. To enable responsible behaviour in agents, they must be endowed with an awareness of responsibility. This motivated the notion of *responsibility-aware agents* [Yazdanpanah *et al.*, 2023; Kobayashi *et al.*, 2023], which are assumed to consider not only their own rewards/payoffs but also the broader implications of their actions. In MAS setting, where agents interact and may collaborate, the problem of responsibility attribution becomes increasingly important. When multiple agents jointly bring about an outcome, how can we fairly attribute responsibility to each agent, taking into account their respective capability? To illustrate this challenge, consider the following example.

Example 1. *Consider two autonomous vehicles, A_1 and A_2, approaching an uncontrolled junction under snowy conditions. A_1 is travelling from east to west, A_2 from south to north. Each vehicle can either brake or not brake. If neither brakes, they will crash into each other. If both successfully brake (and stop), they will avoid a collision. However, due to the slippery road conditions, there is a 0.2 chance that A_1 will not be able to stop, and a 0.6 chance that A_2 will not be able to stop. Each vehicle aims to minimise its travel time to reach its destination. However, a crash would result in a significant cost.*

Suppose the outcome is a collision resulting from the agents' chosen actions (or strategies), *which vehicle should bear greater responsibility for the collision?* Naturally, this depends on the strategies adopted during the interaction. However, responsibility may also hinge on the agents' *capabilities* [Gerstenberg *et al.*, 2011]. To illustrate, suppose that A_2 can guarantee stopping with certainty. If both agents adopt the same mixed strategy that includes a non-zero probability of not braking, and a collision occurs, it seems reasonable to assign greater responsibility to A_2, as it had the capability to avoid the collision entirely. Moreover, there is an inherent trade-off between minimising travel time and avoiding collisions. This creates a strategic decision-making problem in which agents must weigh their utility (e.g., travel time) against potential responsibility penalties for a collision. Given these considerations, some important questions arise: *Given the strategy profiles and capabilities, how should responsibility be allocated?*, *If the agents are aware of responsibility attribution, how should they choose their strategies?*

In this paper, we aim to answer such questions. First, we study responsibility attribution through the lens of *counterfactual reasoning*, which evaluates an agent's contribution by asking how the outcome would change if the agent had acted differently. We follow well-established accounts of causation in philosophy and AI, where counterfactuals capture "*what would have happened otherwise*" [Chockler and Halpern, 2004; Lewis, 2013; Pearl, 2009]. We then examine strategic reasoning for responsibility-aware agents—agents that weigh not only their expected rewards but also the potential responsibility they may incur. This perspective provides a framework for addressing the earlier question of (strategic) *forward reasoning*: how such agents should choose their strategies.

Contribution Our main *contribution* is an approach for responsibility attribution in probabilistic multi-agent systems.

Firstly, we introduce a formal framework based on the *Shapley value* [Shapley, 1953] and show that the resulting allocations satisfy desirable properties such as *fairness* (responsibility reflects contribution) and *consistency* (an agent's share does not decrease when its influence increase). Although we do not claim this to be a definitive account of responsibility, the framework is interpretable, and based on a concept used widely across disciplines. Our focus is not on formalising *moral* responsibility, an extremely complex area with no clear consensus, nevertheless, we believe that our approach can contribute toward such a formalisation.

Secondly, we study strategic reasoning in the setting where agents employ *randomised memoryless strategies* (aka *behavioural* [Cristau *et al.*, 2010]), i.e., functions that map from state to a probability distribution over actions. Following the tradition of formal methods in MAS, we extend the logic PATL with *cumulative reward* (similar to rPATL [Kwiatkowska *et al.*, 2019]) and introduce *degree of responsibility operators*. We show that model checking this logic remains in PSPACE, thus no harder than rPATL [Kwiatkowska *et al.*, 2021]. Finally, we present a method for computing Nash equilibrium strategy profiles where agents' utility functions incorporate both reward and responsibility allocation. We show that when agents' temporal objectives are specified as either *reachability* or *safety* conditions, these equilibria can also be computed in PSPACE. Due to space constraints, extended version of the paper containing full proofs is provided in [Mu and Najib, 2026].

Related work Research on responsibility and attribution spans a large body of work; here, we highlight some relevant studies and key distinctions. Responsibility attribution is closely tied to causality and counterfactual reasoning. Foundational work by Halpern and Pearl [Halpern and Pearl, 2005] introduced structural causal models, which underpin formal definitions of causality and counterfactual dependence. Building on this, [Chockler and Halpern, 2004] proposed a quantitative notion of responsibility, which later extended to a notion of *blame attribution* [Halpern and Kleiman-Weiner, 2018]. While conceptually related to our approach, their framework relies on directed acyclic graphs and is primarily suited for reasoning about beliefs, intentions, and epistemic states, whereas our work focuses on temporally extended properties in strategic settings.

The Shapley value has also been applied for responsibility allocation. [Friedenberg and Halpern, 2019] extends the structural-model approach [Halpern and Kleiman-Weiner, 2018] by incorporating the Shapley value to distribute blame, and demonstrating that the allocation scheme satisfies desirable properties. As previously discussed, their model differs from ours and is suited for different purposes. Similarly, [Yazdanpanah *et al.*, 2019] apply Shapley-based responsibility allocation in concurrent epistemic games, enabling epistemic modeling but under deterministic assumptions, in contrast to our probabilistic setting. [Baier *et al.*, 2021a] studies backward and forward responsibility in *extensive-form games*, while [Baier *et al.*, 2021b] employs parametric Markov chains (pMCs), which are closely related to our model. Memoryless strategy profiles naturally correspond to parameters assignements in a pMC, although the converse does not always hold[1]. Furthermore, [Baier *et al.*, 2021b] primarily focuses on attributing changes to *parameters*, with its main concern being how variations in temporal satisfaction degrees should be attributed to changes in parameter assignments. In contrast, our work assigns responsibility to individual players.

Closer to our work is [Mu *et al.*, 2025], which uses concurrent stochastic games to reason about *causal active* and *causal passive* responsibilities, building on [Parker *et al.*, 2023]. Their approach considers how strategies contribute to or prevent outcomes but does not employ counterfactual reasoning nor guarantee fair attribution. From temporal-logic perspective, [De Giacomo *et al.*, 2025] uses LTL_f to represent outcomes—similar to ours—but focuses on a single agent from a first-person perspective, whereas our approach adopts a third-person, multi-agent view. Other related works include Naumov and Tao's logical frameworks for blameworthiness and "seeing-to-it" responsibility [Naumov and Tao, 2019; Naumov and Tao, 2021], as well as STIT-based accounts [Ciuni and Mastop, 2009; De Lima *et al.*, 2010; Baltag *et al.*, 2021; Abarca and Broersen, 2022]. These differ from ours in both their underlying models and their conceptions of responsibility. In particular, STIT approaches define responsibility via an agent's modal ability to "see to it that" an outcome occurs, whereas our framework evaluates responsibility through counterfactual dependence—what would happen were an agent to act differently—and seeks a fair quantitative attribution across agents.

2 Preliminaries

2.1 Concurrent stochastic game

Definition 1. *A concurrent stochastic multi-player game (CSG) is a tuple* $\mathcal{G} = (\mathsf{Ag}, S, s^0, (\mathsf{Act}_i)_{i\in\mathsf{Ag}}, \delta, \mathsf{Ap}, L)$ *where:*

- $\mathsf{Ag} = \{1, \dots, n\}$ *is a finite set of* agents*;*
- S *is a finite non-empty set of* states*;*
- $s^0 \in S$ *is the* initial state*;*
- Act_i *is a finite set of actions for* i*. With each agent* i *and state* $s \in S$*, we associate a non-empty set* $\mathsf{Act}_i(s)$ *of available actions that* i *can perform in* s*. Write* $\mathsf{Act}^{\mathsf{Ag}} = \mathsf{Act}_1 \times \cdots \times \mathsf{Act}_n$*.*
- $\Delta : S \times \mathsf{Act}^{\mathsf{Ag}} \to \mathsf{Dist}(S)$ *is a probabilistic transition function;*
- Ap *is a finite set of* atomic propositions*;*
- $L : S \to 2^{\mathsf{Ap}}$ *is the labelling function mapping each state to a set of atomic propositions drawn from* Ap*.*

Following [Kwiatkowska *et al.*, 2021], we augment CSGs with *reward structures* of the form $r = (r_s, r_a)$, where $r_s : S \to \mathbb{R}$ is the state reward and $r_a : \mathsf{Act}^{\mathsf{Ag}} \to \mathbb{R}$ is the action reward, and consider *cumulative rewards*, that is the sum of payoffs accumulated during the run until a specific point.

[1] Some parameter assignments cannot be expressed as individual strategies; see [Stan *et al.*, 2024, Remark 8]. We address this by imposing additional constraints in our parametric model (Section 6).

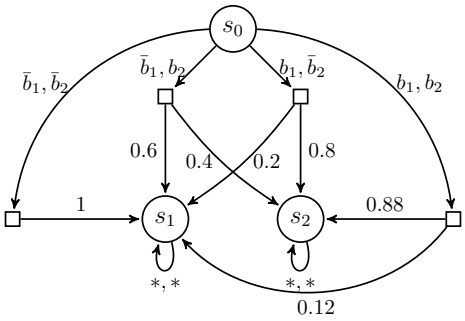

Figure 1: CSG model from the running example.

Example 2. *From the running example, we construct a model in Figure 1. With* $\mathsf{Ag} = \{1, 2\}$*, representing* A_1 *and* A_2*.* $\mathsf{Act}_i = \{b_i, \bar{b}_i\}$ *representing agent* $i \in \mathsf{Ag}$ *choosing to brake and not brake.* s_0 *is the initial state, and* $L(s_1) = \{crash\}, L(s_0) = \{init\}, L(s_2) = \{pass\}$*. Note that from* s_1 *and* s_2*, there are only self-loops.*

Definition 2. *A* path π *is a non-empty sequence* $s_0\vec{a}^0 s_1\vec{a}^1 \ldots$ *of states and joint actions, where* $\vec{a}^i \in \mathsf{Act}^{\mathsf{Ag}}$ *is the* i^{th} *joint action, and* $\Delta(s_i, \vec{a}^i, s_{i+1}) > 0$*. A* history ρ *is a finite path,* $\rho_s(i)$ *denotes the* i^{th} *state of* ρ*, and* $\rho_{\vec{a}}(i)$ *denotes the* i^{th} *joint action of* ρ*. In this case, we may write* $\rho_s(i) \xrightarrow{\rho_{\vec{a}}(i)} \rho_s(i+1)$*. Let* $\mathsf{Hist}_{\mathcal{G}}(s)$ *denote the set of histories of* $\mathcal{G}$ *starting from state* s*. We write* $\mathsf{Hist}_{\mathcal{G}}(s)^{\leq k}$ *to denote the set of histories up to and including* $\rho_s(k)$*.*

In this work, we assume that players have memoryless strategies. Informally, a memoryless strategy for player i prescribes, from each state s, the probability of each action $a \in \mathsf{Act}_i(s)$ being chosen.

Definition 3. *A* (memoryless) strategy *for* i *is a function from the set of states to a probability distribution over agent's set of actions* $\sigma_i : S \to \mathsf{Dist}(\mathsf{Act}_i)$*. A* strategy profile *is a tuple of* strategies *for a set of agents, it is denoted by* $\vec{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_n)$*. For a set (or coalition) of agents* $J \subseteq \mathsf{Ag}$*, write* $\vec{\sigma}_J$ *for* $(\sigma_i)_{i \in J}$*.*

Note that a strategy profile $\vec{\sigma}$ for a game $\mathcal{G}$ resolves non-determinism in the game arena. Denote by $\mathcal{G}_{\vec{\sigma}}$ the resulting game arena (essentially, a Markov chain) after applying $\vec{\sigma}$.

Definition 4. *A state* s *and a strategy profile* $\vec{\sigma}$ *induce a set of histories* $s_0\vec{a}^0 s_1\vec{a}^1 \ldots$*, with* $\mathbb{P}(\vec{a}^i) \cdot \Delta(s_i, \vec{a}^i, s_{i+1}) > 0$ *and* $s = s_0$*, denoted by* $\mathsf{Hist}_{\vec{\sigma}}(s)$*. For a history* $\rho \in \mathsf{Hist}_{\vec{\sigma}}(s)^{\leq k}$*, the probability of* $\rho = s_0 \xrightarrow{\vec{a}^0} s_1 \ldots \xrightarrow{\vec{a}^{k-1}} s_k$*, with* $s_0 = s$*, is given by:*

$$\mathbb{P}_{\vec{\sigma}}(\rho) \triangleq \prod_{j=0}^{k-1} \left[\left(\prod_{i=1}^{n} \sigma_i(s_j)(\vec{a}_i^j) \right) \cdot \Delta(s_j, \vec{a}^j, s_{j+1}) \right].$$

For a subset of agents $J \subseteq \mathsf{Ag}$ and strategies $\vec{\sigma}_J$, we say that a history ρ is *compatible* with $\vec{\sigma}_J$ if for every $k \in \mathbb{N}$, there exists a joint action $\vec{a}^k$ with $\sigma_i(\rho_s(k))(\vec{a}_i^k) > 0$ for each $i \in J$, such that $\Delta(s_k, \vec{a}^k, s_{k+1}) > 0$. We denote by $\mathsf{Hist}_{\vec{\sigma}_J}(s)$ the set of histories starting from s and compatible with $\vec{\sigma}_J$ that can be induced by memoryless strategies.

Definition 5. *The* payoff *function defined as a map from a set of histories to a real value* $\wp : \mathsf{Hist}_{\mathcal{G}}(s) \to \mathbb{R}^{|\mathsf{Ag}|}$*.* $\wp^i$ *denotes the payoff function of* $i \in \mathsf{Ag}$ *and is defined as:*

$$\wp^i(\rho) \triangleq \left(\sum_{j=0}^{t-1} (r_a^i(\vec{a}^j) + r_s^i(s_j)) \cdot \Delta(s_j, \vec{a}^j, s_{j+1}) \right)$$

where $\rho = s_0 \xrightarrow{\vec{a}^0} s_1 \xrightarrow{\vec{a}^1} \ldots \xrightarrow{\vec{a}^{t-1}} s_t \in \mathsf{Hist}_{\mathcal{G}}(s_0)$*.*

2.2 PATL

PATL [Chen and Lu, 2007] is a logic for reasoning about the strategic abilities of (coalitions of) agents in probabilistic systems. In this paper, we consider PATL with *bounded temporal properties*.

Definition 6 (Syntax of PATL)**.** *The* state formulae *(ϕ) and* path formulae *(ψ) are defined as follows:*

$$\begin{array}{lcl} \phi & ::= & a \mid \neg\phi \mid \phi \wedge \phi \mid \langle A \rangle \mathbf{P}_{\bowtie p}[\psi] \\ \psi & ::= & \bigcirc\phi \mid (\phi \mathbf{U}_{\leq k} \phi) \end{array}$$

where $a \in \mathsf{Ap}$ *is an* atomic proposition*,* $A \subseteq \mathsf{Ag}$ *is a set of agents,* $\bowtie \in \{\leq, <, \geq, >\}$*,* $p \in [0, 1]$*,* $k \in \mathbb{N}$ *is a time bound, and* $\langle A \rangle$ *is the strategy quantifier.*

The until operator $\mathbf{U}$ allows one to derive the temporal modalities $\Diamond$ ("eventually") and $\Box$ ("always"): $\Diamond_{\leq k}\psi \triangleq \texttt{true}\ \mathbf{U}_{\leq k}\ \psi$ and $\Box_{\leq k}\psi \triangleq \neg\Diamond_{\leq k}(\neg\psi)$.

Definition 7 (Semantics of PATL)**.** *The satisfaction relation for a CSG* $\mathcal{G}$*, state* $s \in S$*, history* ρ*, atom* $a \in \mathsf{Ap}$*, and PATL formula is defined as follows:*

- $s \models a$ *iff* $a \in L(s)$*.*
- $s \models \neg\phi$ *iff* $s \not\models \phi$*.*
- $s \models \phi \wedge \phi'$ *iff* $s \models \phi$ *and* $s \models \phi'$*.*
- $s \models \langle A \rangle \mathbf{P}_{\bowtie p}[\psi]$ *iff* $\exists \vec{\sigma}_A . \forall \vec{\sigma}_{-A} . (\mathbb{P}(\{\rho \in \mathsf{Hist}_{(\vec{\sigma}_A, \vec{\sigma}_{-A})}(s) \mid \rho \models \psi\}) \bowtie p)$*.*
- $\rho \models \bigcirc\phi$ *iff* $\rho_s(1) \models \phi$*.*
- $\rho \models \phi \mathbf{U}_{\leq k} \phi'$ *iff there exists* $i \leq k$ *such that:* $\rho_s(i) \models \phi'$*, and* $\rho_s(j) \models \phi$ *for all* $j < i$*.*

We say that formula φ is *true* in a CGS $\mathcal{G}$ iff $s^0 \models \varphi$.

3 Formalising Counterfactual Responsibility Attribution

We build on the concept of *Necessary Element of a Sufficient Set* (NESS) test [Wright, 1985], which has been used in many game-theoretic frameworks (e.g., [Braham and van Hees, 2012; Baltag *et al.*, 2021; Parker *et al.*, 2023]), and adapt it to our setting.

Definition 8. *Given a CSG* $\mathcal{G}$*, we say that agent* i *bears backward Counterfactual Responsibility (bCR) in a strategy profile* $\vec{\sigma}$ *for outcome* φ*, specified in a PATL path formula, written as* $i \sim \mathsf{bCR}^{\vec{\sigma}}_{\varphi}$*, if the following conditions hold:* $\exists J \subseteq \mathsf{Ag} \setminus \{i\}$ *such that* $\exists \rho' \in \mathsf{Hist}_{\vec{\sigma}_J}(s^0) . \rho' \not\models \varphi$ *and* $\forall \rho'' \in \mathsf{Hist}_{\vec{\sigma}_{J \cup \{i\}}}(s^0) . \rho'' \models \varphi$*.*

In words, agent i is responsible for an outcome φ under strategy profile $\vec{\sigma}$, if it belongs to some coalition $J \cup \{i\}$ whose joint strategies are sufficient to guarantee φ, while the strategies of J (without i) are not.

In the definition above, bCR is a qualitative notion: i either bears responsibility or does not. Our aim is to extend this to a quantitative setting by determining "how much" responsibility agent i bears. Intuitively, this corresponds to measuring i's marginal contribution to increasing the likelihood/probability of satisfying φ. To this end, we allocate responsibility using the Shapley value. A challenge arises when measuring the marginal contribution of a coalition that does not contain all agents (i.e., non-grand coalition) since the strategies of agents outside that coalition are undefined. We address this by assuming that φ is an undesirable outcome that agents (retrospectively) would prefer to avoid. When computing computing marginal contributions, we assume that agents outside the coalition minimise the probability of satisfying φ. This imposes no moral duty on them; it simply provides a counterfactual baseline for non-grand coalitions, whose outside behaviour would otherwise be undefined.

Definition 9. *Given a bound*[2] $k \in \mathbb{N}$*, the degree of* i *bearing bCR in* $\vec{\sigma}$ *for* φ *is defined as:*

$$\mathbb{D}^i(\vec{\sigma}, \varphi) = \sum_{J \subseteq \mathsf{Ag} \setminus \{i\}} \frac{|J|!(|\mathsf{Ag}| - |J| - 1)!}{|\mathsf{Ag}|!} \cdot (v_{\vec{\sigma},\varphi}(J \cup \{i\}) - v_{\vec{\sigma},\varphi}(J))$$

For any $A \subseteq \mathsf{Ag}$*, the function* $v_{\vec{\sigma},\varphi}(A)$ *for the* k*-bounded histories is defined as:*

$$v_{\vec{\sigma},\varphi}(A) = \min_{\vec{\sigma}_{-A}} \mathsf{Prob}(\{\rho \mid \rho \in \mathsf{Hist}_{(\vec{\sigma}_A, \vec{\sigma}_{-A})}(s^0)^{\leq k} \wedge \rho \models \varphi\})$$

Example 3. *Continuing the running example, suppose given an outcome* $\varphi := \bigcirc$*crash (i.e., crash in the next time-step), and consider the strategy profile* $\vec{\sigma}$ *corresponding to each agent choosing not to brake with probability of* 1*. In this case, we obtain:*

$$v_{\vec{\sigma},\varphi}(\{1,2\}) = 1, \qquad v_{\vec{\sigma},\varphi}(\{1\}) = 0.6$$
$$v_{\vec{\sigma},\varphi}(\{2\}) = 0.2, \qquad v_{\vec{\sigma},\varphi}(\varnothing) = 0.12$$

Therefore:

$$\mathbb{D}^1(\vec{\sigma}, \varphi) = \frac{1}{2} \cdot (0.6 - 0.12) + \frac{1}{2} \cdot (1 - 0.2) = 0.64$$
$$\mathbb{D}^2(\vec{\sigma}, \varphi) = \frac{1}{2} \cdot (0.2 - 0.12) + \frac{1}{2} \cdot (1 - 0.6) = 0.24$$

Note that A_1 has a higher degree of responsibility than A_2, intuitively, because A_1 has a greater chance of being able to stop (i.e., 0.8, compared to 0.4 for A_2). For comparison, under the other pure profiles (b_1, b_2), $(b_1, \bar{b}_2)$, and $(\bar{b}_1, b_2)$, the corresponding responsibility pairs $(\mathbb{D}^1, \mathbb{D}^2)$ are $(0, 0)$, $(0, 0.08)$, and $(0.48, 0)$, respectively.

[2]This bound comes from the PATL path formula φ.

4 Desirable Properties

In this section, we demonstrate the desirable properties of our responsibility attribution: *efficiency*, *symmetry*, *null/dummy player*, *additivity*, and *monotonicity*.

To formalise this, we first introduce the notion of *attributable value*, which represents the total amount of responsibility available for distribution among agents. The attributable value for a strategy profile is the difference between the probability of φ occurring under the strategy profile and the minimum probability achievable if agents act optimally (together) to prevent φ.

To illustrate, consider an autonomous vehicle agent i approaching a T-junction. It can turn *left* or *right*. Turning left guarantees arriving late (probability 1), while turning right carries a 0.2 probability of lateness. Here, lateness is inevitable with at least a 0.2 probability. Suppose the vehicle adopts a strategy profile of turning left with certainty (probability 1). The attributable value is $1 - 0.2 = 0.8$, meaning the maximum responsibility assignable to i is 0.8.

Definition 10. *Given a strategy profile* $\vec{\sigma}$*, an outcome* φ*, and a bound* $k \in \mathbb{N}$*, the attributable value of* φ *w.r.t.* $\vec{\sigma}$ *is defined as:* $\Upsilon(\vec{\sigma}, \varphi) = v_{\vec{\sigma},\varphi}(\mathsf{Ag}) - \min_{\vec{\sigma}} \mathsf{Prob}(\{\rho \mid \rho \in \mathsf{Hist}_{\vec{\sigma}}(s^0)^{\leq k} \wedge \rho \models \varphi\})$.

By Definition 9, we have that:

$$v_{\vec{\sigma},\varphi}(\varnothing) = \min_{\vec{\sigma}_{-\varnothing}} \mathsf{Prob}(\{\rho \mid \rho \in \mathsf{Hist}_{\vec{\sigma}}(s^0)^{\leq k} \wedge \rho \models \varphi\}$$

and since $\mathsf{Ag} \setminus \varnothing = \mathsf{Ag}$, therefore $\vec{\sigma}_{\mathsf{Ag}} = \vec{\sigma}$, and we immediately obtain the following lemma.

Lemma 1. *For a strategy profile* $\vec{\sigma}$ *and outcome* φ*, it holds:*

$$\Upsilon(\vec{\sigma}, \varphi) = v_{\vec{\sigma},\varphi}(\mathsf{Ag}) - v_{\vec{\sigma},\varphi}(\varnothing)$$

Example 4. *Consider Example 3, the attributable value is:*

$$\Upsilon(\vec{\sigma}, \varphi) = v_{\vec{\sigma},\varphi}(\mathsf{Ag}) - v_{\vec{\sigma},\varphi}(\varnothing) = 1 - 0.12 = 0.88$$

Fairness The first three properties relate to *fairness*. *Efficiency* guarantees that the sum of assigned responsibility values equals the attributable value, ensuring no unassigned responsibility remains. Conversely, it also ensures that the distributed responsibility does not exceed the attributable value, preventing any agent from bearing excess responsibility. *Symmetry* ensures that two agents with identical contributions to an outcome receive equal degrees of responsibility. The *dummy/null player* property ensures that agents who do not contribute to an outcome receive zero responsibility.

Proposition 1 (Efficiency)**.** *For a strategy profile* $\vec{\sigma}$ *and outcome* φ*, it holds that:* $\sum_{i \in \mathsf{Ag}} \mathbb{D}^i(\vec{\sigma}, \varphi) = \Upsilon(\vec{\sigma}, \varphi)$.

Example 5. *Consider again Example 3 and 4, note that:*

$$\sum_{i \in \mathsf{Ag}} \mathbb{D}^i(\vec{\sigma}, \varphi) = 0.64 + 0.24 = \Upsilon(\vec{\sigma}, \varphi) = 0.88$$

Proposition 2 (Symmetry)**.** *For two agents* $i, j \in \mathsf{Ag}$*, if:*

$$\forall J \subseteq \mathsf{Ag} \setminus \{i, j\}, v_{\vec{\sigma},\varphi}(J \cup \{i\}) = v_{\vec{\sigma},\varphi}(J \cup \{j\})$$

then:

$$\mathbb{D}^i(\vec{\sigma}, \varphi) = \mathbb{D}^j(\vec{\sigma}, \varphi)$$

Proposition 3 (Null Player)**.** *If for all* $J \subseteq \mathsf{Ag} \setminus \{i\} : v_{\vec{\sigma},\varphi}(J \cup \{i\}) = v_{\vec{\sigma},\varphi}(J)$*, then* $\mathbb{D}^i(\vec{\sigma}, \varphi) = 0$.

Consistency Additivity property states that for two *independent* games, the Shapley value of the combined game is equal to the sum of individual values. However, this raises the question of how to combine the games and to what extent they are truly independent[3], particularly, in our setting. One could define this by considering two separate CSGs, resulting in a "combined" responsibility degree that is simply the sum of the two games. However, we find this approach uninteresting and not particularly useful. Instead, we consider additivity with respect to disjunctions of outcomes $\varphi_1 \vee \varphi_2$. Observe that the function $\mathbb{D}^i(\cdot,\cdot)$ is generally not additive. To see this, consider $\varphi_1 := \varphi$ and $\varphi_2 := \neg\varphi$. Clearly, for any $\vec{\sigma}$, $\mathbb{D}^i(\vec{\sigma}, \varphi_1 \vee \varphi_2) = 0$, while it may be the case that $\mathbb{D}^i(\vec{\sigma}, \varphi_1) > 0$ or $\mathbb{D}^i(\vec{\sigma}, \varphi_2) > 0$. However, we can show that *subadditivity* holds, and for a special case of outcomes we obtain (exact) additivity. *Monotonicity* intuitively states that for two outcomes φ_1 and φ_2, where φ_1 is "included" in φ_2, the degree of responsibility for φ_2 would likely increase. Additivity and monotonicity relate to consistency in the sense that if we expand the outcome (under certain conditions), then responsibility should increase or at least not decrease.

Proposition 4 (Sub-additivity)**.** *Given $\vec{\sigma}$, for any two outcome φ_1 and φ_2, and any $i \in \mathsf{Ag}$:*

$$\mathbb{D}^i(\vec{\sigma}, \varphi_1 \vee \varphi_2) \leq \mathbb{D}^i(\vec{\sigma}, \varphi_1) + \mathbb{D}^i(\vec{\sigma}, \varphi_2)$$

Example 6. *Continue the running example, suppose given outcomes $\varphi_1 = \bigcirc$ crash and $\varphi_2 = \bigcirc$ pass, consider $\vec{\sigma}$ corresponding to each agent choosing* to brake *with probability of 1, let $\vee$ denote $\varphi_1 \vee \varphi_2$ for short. In this case, we have:*

$$v_{\vec{\sigma},\varphi_1}(\{1,2\}) = v_{\vec{\sigma},\varphi_1}(\{1\}) = v_{\vec{\sigma},\varphi_1}(\{2\}) = v_{\vec{\sigma},\varphi_1}(\varnothing) = 0.12,$$
$$v_{\vec{\sigma},\varphi_2}(\{1,2\}) = 0.88, \quad v_{\vec{\sigma},\varphi_2}(\{1\}) = 0.8,$$
$$v_{\vec{\sigma},\varphi_2}(\{2\}) = 0.4, \quad v_{\vec{\sigma},\varphi_2}(\varnothing) = 0,$$
$$v_{\vec{\sigma},\vee}(\{1,2\}) = v_{\vec{\sigma},\vee}(\{1\}) = v_{\vec{\sigma},\vee}(\{2\}) = v_{\vec{\sigma},\vee}(\varnothing) = 0.12$$

Thus:

$$\mathbb{D}^1(\vec{\sigma}, \varphi_1) = \mathbb{D}^2(\vec{\sigma}, \varphi_1) = 0,$$
$$\mathbb{D}^1(\vec{\sigma}, \varphi_2) = 0.64,\ \mathbb{D}^2(\vec{\sigma}, \varphi_2) = 0.24,$$
$$\mathbb{D}^1(\vec{\sigma}, \varphi_1 \vee \varphi_2) = \mathbb{D}^2(\vec{\sigma}, \varphi_1 \vee \varphi_2) = 0$$

Clearly, for $i = \{1,2\}$: $\mathbb{D}^i(\vec{\sigma}, \varphi_1 \vee \varphi_2) \leq \mathbb{D}^i(\vec{\sigma}, \varphi_1) + \mathbb{D}^i(\vec{\sigma}, \varphi_2)$.

To obtain exact additivity, we define the following conditions.

Definition 11. *For two outcomes φ_1 and φ_2, we say that they are disjoint and avoidable, respectively, if*

i) for all $\vec{\sigma}$, we have $\mathsf{Prob}(\{\rho \mid \rho \in \mathsf{Hist}_{\vec{\sigma}}(s^0)^{\leq k} \wedge \rho \models \varphi_1 \wedge \varphi_2\}) = 0$

ii) there exists $\vec{\sigma}$, such that $\mathsf{Prob}(\{\rho \mid \rho \in \mathsf{Hist}_{\vec{\sigma}}(s^0)\rho \models (\varphi_1 \vee \varphi_2)\}) < 1$.

Proposition 5 (Exact Additivity)**.** *Let φ_1 and φ_2 be two outcome formulas such that they are disjoint and avoidable, then for any $i \in \mathsf{Ag}$:*

$$\mathbb{D}^i(\vec{\sigma}, \varphi_1 \vee \varphi_2) = \mathbb{D}^i(\vec{\sigma}, \varphi_1) + \mathbb{D}^i(\vec{\sigma}, \varphi_2)$$

[3]There have been many attempts to relax the additivity axiom, e.g. [Young, 1985; Hart and Mas-Colell, 1989; Chun, 1989; van den Brink, 2002].

Example 7. *Note that the scenario proposed in Example 6 does not satisfy Proposition 5, as it meets assumption i) but violates assumption ii).*

Proposition 6 (Monotonicity)**.** *If $\varphi_1 \Rightarrow \varphi_2$ and φ_2 is avoidable, then for any $i \in \mathsf{Ag}$ and a given $\vec{\sigma}$:*

$$\mathbb{D}^i(\vec{\sigma}, \varphi_1) \leq \mathbb{D}^i(\vec{\sigma}, \varphi_2)$$

Intuitively, responsibility should not decrease when the outcome becomes more inclusive - any trace satisfying φ_1 also satisfies φ_2, so agents influencing φ_1 also influencing φ_2, potentially to a greater extent.

Example 8. *Continue the running example, consider $\varphi_1 = \bigcirc$ crash and $\varphi_2 = \Diamond_{\leq 2}$ crash and $\vec{\sigma}$ corresponding to each agent choosing* not to brake *with probability of 1. Clearly, in this case, $\varphi_1 \Rightarrow \varphi_2$ and φ_2 is avoidable. We have: $\mathbb{D}^1(\vec{\sigma}, \varphi_1) = \mathbb{D}^1(\vec{\sigma}, \varphi_2) = 0.64$, $\mathbb{D}^2(\vec{\sigma}, \varphi_1) = \mathbb{D}^2(\vec{\sigma}, \varphi_2) = 0.24$. We have "=" specifically here as there is only a self-loop from s_1.*

5 Logic with CR attributions: PATL-SR

We introduce PATL-SR, a variant of PATL that incorporates quantified reward and responsibility attribution formulae as defined in Definition 9.

Definition 12. *The* syntax *of new formulae of* PATL-SR *is given as:*

$$\begin{array}{lll} \phi & ::= & a \mid \neg\phi \mid \phi \wedge \phi \mid \langle A\rangle \mathbf{P}_{\bowtie p}[\psi] \mid \langle A\rangle \mathbf{R}_{\bowtie q}[\psi] \mid \\ & & \langle A\rangle \mathbf{D}_{\bowtie d}[\mathsf{bCR}_i(\vec{\sigma}, \psi)] \\ \psi & ::= & \bigcirc\phi \mid \phi \mathbf{U}_{\leq k} \phi \end{array}$$

$q, d \in \mathbb{R}$ are reward and responsibility degree bounds, respectively.

The formula $\langle A\rangle \mathbf{R}_{\bowtie q}\ [\Diamond_{\leq k}\psi]$ expresses that the coalition A has a strategy such that the expected rewards of satisfying path formula ψ is $\bowtie q$ when the strategy is followed. The formula $\langle A\rangle \mathbf{D}_{\bowtie d}\ [\mathsf{bCR}_i(\vec{\sigma}, \psi)]$ expresses that, under strategy profile $\vec{\sigma}$, the backward counterfactual responsibility of i in ensuring ψ within coalition A is quantified by $\bowtie d$.

Definition 13. *Given a model $\mathcal{G}$, the semantics of the new* PATL-SR *operators are defined as follows; the full formal definitions are provided in the supplementary material.*

- $s \models \langle A\rangle \mathbf{R}_{\bowtie q}[\psi]$ *iff there exists a strategy profile of coalition A such that the expected accumulated reward of paths from state s which are consistent with the strategy and satisfy ψ is $\bowtie q$, i.e., $\exists \vec{\sigma}_A . \mathbb{E}^{\vec{\sigma}_A}_{\mathcal{G}}\big(s, \mathsf{rew}(r, \psi)\big) \bowtie q$, where:* $\mathsf{rew}(r, \psi)(\rho) =$

$$\begin{cases} \sum_{i=0}^{k} (r_a(\rho_a(i)) + r_s(\rho_s(i))) & \text{if } \rho \models \psi; \\ 0 & \text{otherwise} \end{cases}$$

  where k denotes the time bound of ψ.

- $s \models \langle A\rangle \mathbf{D}_{\bowtie d}[\mathsf{bCR}_i(\vec{\sigma}, \psi)]$ *iff* $\mathbb{D}^i(\vec{\sigma}_A, \psi) \bowtie d$, *where:* $\mathbb{D}^i(\vec{\sigma}_A, \psi) = \sum_{J \subseteq A \setminus \{i\}} \frac{|J|!(|A|-|J|-1)!}{|A|!} \cdot (v_{\vec{\sigma}_A,\psi}(J \cup \{i\}) - v_{\vec{\sigma}_A,\psi}(J))$.

Example 9. *Consider our running example with the outcome that "a crash takes place within 2 time steps", which can be expressed by: $\psi = \langle\{1,2\}\rangle \Diamond_{\leq 2}$ crash. Suppose* 1 *and* 2 *both choose to brake under $\vec{\sigma}$, thus: $v_{\vec{\sigma},\psi}(\{1,2\}) = v_{\vec{\sigma},\psi}(\{1\}) = v_{\vec{\sigma},\psi}(\{2\}) = v_{\vec{\sigma},\psi}(\varnothing) = 0.12$, and thus $\mathbb{D}^1(\vec{\sigma}, \psi) = \mathbb{D}^2(\vec{\sigma}, \psi) = 0$. So: $\mathbf{D}_{\leq 0}[\mathsf{bCR}_i(\vec{\sigma}, \psi)]$ is* `true`.

The model checking of PATL-SR works in a similar way to rPATL proposed in [Kwiatkowska *et al.*, 2019]. The main difference is that we have the new $\mathbf{D}_{\bowtie d}[\mathsf{bCR}_i(\vec{\sigma}, \psi)]$ operator, which corresponds to computing the function $\mathbb{D}^i(\vec{\sigma}_A, \psi)$ in Definition 13. Computing such function can be done in polynomial space.

Lemma 2. *Given a CSG $\mathcal{G}$, a coalition A, strategy profile $\vec{\sigma}$, and a path formula ψ, computing $\mathbb{D}^i(\vec{\sigma}_A, \psi)$ can be done in polynomial space.*

Theorem 1. *Model checking* PATL-SR *is in PSPACE.*

6 CR-aware Strategic Decision-Making

This section studies how agents choose strategies when they are aware of counterfactual responsibility attributions. We define a utility function that combines an agent's payoff with its degree of responsibility, and use *Nash equilibrium* (NE) to characterise stable strategy profiles under this utility. To retain decidability, we restrict attention to memoryless strategies and reachability/safety objectives [Ummels and Wojtczak, 2011].

First, we introduce a parametric model. Observe that the set of all memoryless strategies for player i from state s can be encoded by a set of variables $V^i_s = \{x_a : a \in \mathsf{Act}_i(s)\}$. Intuitively, the value of $x_a \in V^i_s$ corresponds to the probability of action a being chosen by player i in state s. Let $V^i = \bigcup_{s \in S} V^i_s$. A memoryless (mixed) strategy for i in $\mathcal{G}$ thus corresponds to an *evaluation* of such a set of variables represented by a function $C^i : V^i \to \mathbb{R}$.

Definition 14. *Given $\mathcal{G} = (\mathsf{Ag}, S, s^0, (\mathsf{Act}_i)_{i \in \mathsf{Ag}}, \delta, \mathsf{Ap}, L)$, we construct the corresponding* parametric stochastic multi-agent system *(PSMAS) as a tuple $\mathcal{M}$=$(\mathsf{Ag}, S, s^0, \mathsf{Act}, V, \Delta, \mathsf{Ap}, L)$, where: $V^i = \{x_{i,1}, \ldots, x_{i,k}\} \subseteq V$ is a finite set of variables (*parameters*) over $\mathbb{R}^k$ for each agent i; $\Delta : S \times \mathsf{Act}^{\mathsf{Ag}} \times S \to \mathcal{F}_x$ is the* probabilistic transition function*, where $\mathcal{F}_x$ is the set of polynomials over V with rational coefficients which can be viewed as a parametric transition probability matrix that respects the distribution in δ.*

Observe that the resulting PSMAS $\mathcal{M}$ is polynomial in size wrt the game $\mathcal{G}$. Next, we introduce the notion of *admissible* evaluations in a given PSMAS. An evaluation C^i is admissible if: (1) $\Delta_{C^i}(s, \vec{a}, s) \in [0, 1]$ for all $s, s' \in S$ and $\vec{a} \in \mathsf{Act}^{\mathsf{Ag}}$, i.e., each joint action has a probability between 0 and 1; (2) $C^i(x) \in [0, 1]$ for all $i \in \mathsf{Ag}$ and $x \in V^i$, i.e., for each player, each action probability is between 0 and 1; (3) $\sum_{x \in V^i_s} C^i(x) = 1$ for all $i \in \mathsf{Ag}$ and $s \in S$, i.e., For each player and each state, the total of action probability values should equal to 1. Henceforth, we assume that the model takes the form of a PSMAS. When evaluating solutions, we further assume that such solutions are admissible.

Definition 15 (Responsibility-aware Utility Function)**.** *Given $\mathcal{M}$ and an outcome φ and strategy profile $\vec{\sigma}$, for an agent i, the utility function is defined as:*

$$u^{\varphi}_i(\vec{\sigma}) \triangleq \sum_{\rho \in \mathsf{Hist}_{\vec{\sigma}}(s^0)} \wp^i(\rho) - \lambda \cdot \mathbb{D}^i(\vec{\sigma}, \varphi) \qquad (1)$$

where $\wp^i(\rho)$ and $\mathbb{D}^i(\vec{\sigma}, \varphi)$ are the payoff function and bCR degree function defined in Definition 5 and 9 respectively, $\lambda \in \mathbb{R}$ is a coefficient.

Definition 16 (NE-COMPUTATION)**.** *Given a PSMAS $\mathcal{M}$, a* PATL-SR *path formula φ, and utility functions $(u^{\varphi}_i(\vec{\sigma}))_{i \in \mathsf{Ag}}$, the* NE-COMPUTATION *problem is to compute a NE strategy profile $\vec{\sigma}$ that satisfies φ.*

The high level process to address NE-COMPUTATION consists of two main steps: (i) *Satisfaction filtering:* identify the set of strategy profiles that satisfy φ. This is achieved through parametric model checking on $\mathcal{M}$, which yields rational valuation functions V^i for each agent $i \in \mathsf{Ag}$. These functions capture the probabilities of selecting actions $a \in \mathsf{Act}$ in each state s under mixed strategies. (ii) *Equilibrium filtering:* refine this set by isolating the strategies that constitute an NE. This involves formulating a system of equations characterising NE conditions and solving them within the strategy set obtained in (i).

Proposition 7 (Quasi-Concavity of Utility)**.** *The utility function u^{φ}_i is quasi-concave if the following conditions hold:*

1) *linearity of reward: the expected reward is linear in $\vec{\sigma}_i$;*
2) *linearity of responsibility contribution: the responsibility-valuation function $v_{\vec{\sigma},\varphi}(A)$ is linear in $\vec{\sigma}_i$ for all coalition $A \subseteq \mathsf{Ag}$;*
3) *fixed strategies of others: the strategy profile $\vec{\sigma}_{-i}$ of other agents is fixed.*

In our setting, quasi-concavity is evaluated with respect to an individual agent's mixed strategy. To ensure this is well-defined, we fix the strategies of all other agents. This is a standard approach in both game-theoretic best-response analysis and parametric model checking [Shoham and Leyton-Brown, 2009; Basset *et al.*, 2015]. When both reward and responsibility allocation are linear in the agent's strategy, the resulting utility function is quasi-concave, enabling NE computation via convex optimisation.

In a PSMAS, memoryless strategies are encoded symbolically by parameters denoting action probabilities. With the other agents' strategies fixed, quasi-concavity follows whenever the reward and responsibility terms are linear in the strategy parameters of the agent under consideration. We show that, for reachability and safety objectives in our setting, these terms satisfy the required conditions, yielding the following result.

Proposition 8. *In a game where the formula φ is a reachability or safety formula, the quasi-concavity property (Proposition 7) holds.*

In general, u^{φ}_i is quasi-concave in $\vec{\sigma}_i$ when the agent's preference for mixing strategies leads to at least as good an outcome as playing pure strategies, or formally, when the level sets $\vec{\sigma}_i \mid u^{\varphi}_i(\vec{\sigma}_i, \vec{\sigma}_{-i}) \geq c$ are convex for every $c \in \mathbb{R}$ [Osborne and Rubinstein, 1994]. This result enables the use of convex optimisation (best response over pure strategies) to compute Nash equilibria when the assumptions hold.

Computing bCR-aware utility. For a fixed outcome φ, computing $u^{\varphi}_i(\vec{\sigma})$ requires: i) calculating the expected rewards: $\mathbb{E}(\wp^i(\{\rho \in \mathsf{Hist}_{\vec{\sigma}_A}(s) \mid \rho \models_{\mathcal{M}} \psi\}))$, and ii) calculating the bCR degree: $\mathbb{D}^i(\vec{\sigma}_A, \psi)$. Finally, computing the

best response set $\mathbf{BR}_i(\vec{\sigma}_{\mathsf{Ag}\setminus i})$ corresponds to finding the set of strategies σ_i that maximise the utility as the NE.

Definition 17 (NE for bCR-aware Strategy)**.** *Given $\mathcal{M}$, for each agent i, and $\vec{\sigma}_{\mathsf{Ag}\setminus i}$, a strategy σ_i is a* best response *w.r.t. the utility function defined in (1), if it is the set:*

$$\mathbf{BR}_i(\vec{\sigma}_{\mathsf{Ag}\setminus i}) \triangleq \{\sigma_i \mid \max_{\sigma_i}(u_i^{\varphi}(\sigma_i, \vec{\sigma}_{\mathsf{Ag}\setminus i}))\}$$

A strategy profile $\vec{\sigma}$ is a mixed NE if it is a best response for every agent, meaning that $\vec{\sigma} \in \mathbf{BR}_i(\vec{\sigma}_{\mathsf{Ag}\setminus i})$ holds for all $i \in \mathsf{Ag}$ in $\mathcal{M}$.

Example 10. *Revisiting our running example, consider $\psi = \bigcirc$ crash, A_1 and A_2 choose "braking" with probability x_1 and 'x_2 respectively under $\vec{\sigma}$. Suppose that the payoffs for A_1 and A_2 "braking" yields 3 and 1 respectively, while "not-braking" results in a payoff of 2 and 3 respectively, the payoffs of reaching s_1 (crash) for A_1 and A_2 is -3 and -1 respectively. The expected payoff can be computed as:*

$$\sum_{\rho \in \mathsf{Hist}_{\vec{\sigma}}(s^0)} \wp^1(\rho) = x_1 + 0.4x_2 - 1.6x_1x_2 - 1$$

$$\sum_{\rho \in \mathsf{Hist}_{\vec{\sigma}}(s^0)} \wp^2(\rho) = -1.6x_1 - 2x_2 + 1.6x_1x_2 + 2$$

$$v_{\vec{\sigma},\psi}(\{1,2\}) = v_{\vec{\sigma},\psi}(\{2\}) = 0.2x_1(1 - x_2)$$

$$v_{\vec{\sigma},\psi}(\{1\}) = v_{\vec{\sigma},\psi}(\varnothing) = 0.12x_1x_2$$

$$\mathbb{D}^1(\vec{\sigma}, \psi) = 0,\ \mathbb{D}^2(\vec{\sigma}, \psi) = 0.2x_1 - 0.32x_1x_2$$

Consider an example form of the polynomial function:

$$\sum_{\rho \in \mathsf{Hist}_{\vec{\sigma}}(s^0)} \wp^i(\rho) - 10 \cdot \mathbb{D}^i(\vec{\sigma}, \psi)$$

we obtain the utility function as:

$$u_1^{\psi}(\vec{\sigma}) = x_1 + 0.4x_2 - 1.6x_1x_2 - 1$$
$$u_2^{\psi}(\vec{\sigma}) = -3.6x_1 - 2x_2 + 4.8x_1x_2 + 2.$$

Parametric best response expressions. When performing parametric model checking against PATL-SR formulas to evaluate best responses, we obtain parametric expressions that encode optimal strategies. These expressions include parameters representing the probabilities of different actions for each agent. By varying these parameters, we can identify best-response strategies within the strategy profile space, i.e., action probabilities that maximise each agent's utility. This approach enables a systematic exploration of how changes in action probabilities affect both the overall performance and allocation of responsibility within the strategy profile.

Finding the stable responsibility-aware strategy. If the utility function u_i^{φ} is quasi-concave with respect to agent i's mixed strategies in a given model $\mathcal{M}$, then a mixed strategy profile $\vec{\sigma}$ is a NE if and only if every pure strategy played with positive probability in σ_i is a best response to the mixed strategies of the other agents [Osborne and Rubinstein, 1994]. This implies that, in a mixed NE, an agent randomises only among actions that yield the same expected utility, i.e., every action in the support of an agent's equilibrium mixed strategy must lead to an identical valuation: $u_i^{\varphi}(\vec{\sigma}_{-i}, \sigma_i) = u_i^{\varphi}(\vec{\sigma}_{-i}, \sigma_{i\prime})$. For any two actions $a, a' \in \mathsf{Act}$ corresponding to σ_i and $\sigma_{i\prime}$, provided they have nonzero probabilities. This property forms the basis for expressing the NE condition:

$$\begin{cases} u_i^{\varphi}(\vec{\sigma}_{-i}, \sigma_i) = u_i^{\varphi}(\vec{\sigma}_{-i}, \sigma_{i\prime}) \quad \forall \sigma_i, \sigma_{i\prime} \\ \sum_{j=1}^{m} \sigma_{ij} = 1 \quad \forall i \\ 0 \leq \sigma_{ij} \leq 1 \quad \forall i, j \end{cases} \tag{2}$$

Solving these equations allows us to identify the NE and determine the optimal strategy profile for the given objective. This leads to the following theorem, whose proof follows from a reduction to the existential theory of real numbers [Canny, 1988].

Theorem 2. *Computing an NE strategy satisfying* PATL-SR *formula φ can be done in PSPACE.*

Example 11. *Revisiting Example 10, by solving the NE condition equations, we have: $x_1 = 0.42$, $x_2 = 0.625$.*

7 Conclusion and Future Work

We introduce a formal definition of backward (counterfactual) responsibility allocation and provide a framework to model and reason about it. Our definition of responsibility allocation satisfies desirable properties such as fairness and consistency. Utilising this definition, we propose PATL-SR as a language to strategically reason about responsibility allocation, reward, and temporal objectives in probabilistic multi-agent systems. We demonstrate that the model checking problem is in PSPACE, thus no more complex than rPATL as proposed in [Kwiatkowska *et al.*, 2019]. Furthermore, we present an approach to calculate NE strategy profiles, where players' utility functions consider both reward and responsibility allocation. These strategy profiles correspond to stable behaviour in the game and offer a method to determine how agents should or would choose their strategies. This corresponds to game theory's *normative* (how agents should/ought to act) and *descriptive* (how agents would act) interpretations [Wooldridge, 2012]. Both interpretations are valuable for the analysis, verification, and design of responsibility-aware MASs.

As discussed in [Kwiatkowska *et al.*, 2018], bounded temporal properties may require memory. Therefore, a potential future direction is to generalise our approach to *finite-memory* agents. However, this would necessitate a different approach, as using a parametric model would no longer suffice. Another intriguing future direction is to incorporate responsibility allocation into *normative systems* [Ågotnes *et al.*, 2007], where norms can be introduced to achieve desirable outcomes from the designer's perspective [Perelli, 2019]. Additionally, when integrating learning components into the systems, such as in a *multi-agent reinforcement learning* setting, it is worth exploring how to incorporate responsibility allocation with rewards. This could involve designing *reward structures* or *reward machines* [Icarte *et al.*, 2022] to achieve desirable outcomes.

Acknowledgments

Chunyan Mu acknowledges the support provided by the University of Aberdeen through the research leave project on

"Responsibility-aware Trustworthy AI", and the National Decommissioning Centre (NDC) project on "Decision-Making for Oil & Gas Decommissioning: A Formal AI-Driven Approach".

References

[Abarca and Broersen, 2022] Aldo Iván Ramírez Abarca and Jan M Broersen. A stit logic of responsibility. In *AAMAS*, pages 1717–1719, 2022.

[Ågotnes *et al.*, 2007] Thomas Ågotnes, Wiebe van der Hoek, and Michael Wooldridge. Normative system games. In *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, pages 1–8, 2007.

[Baier *et al.*, 2021a] Christel Baier, Florian Funke, and Rupak Majumdar. A game-theoretic account of responsibility allocation. In *IJCAI*, pages 1773–1779. ijcai.org, 2021.

[Baier *et al.*, 2021b] Christel Baier, Florian Funke, and Rupak Majumdar. Responsibility attribution in parameterized markovian models. In *AAAI*, pages 11734–11743. AAAI Press, 2021.

[Baltag *et al.*, 2021] Alexandru Baltag, Ilaria Canavotto, and Sonja Smets. Causal agency and responsibility: a refinement of stit logic. *Logic in high definition: Trends in logical semantics*, pages 149–176, 2021.

[Basset *et al.*, 2015] Nicolas Basset, Marta Z. Kwiatkowska, Ufuk Topcu, and Clemens Wiltsche. Strategy synthesis for stochastic games with multiple long-run objectives. In Christel Baier and Cesare Tinelli, editors, *TACAS*, volume 9035 of *Lecture Notes in Computer Science*, pages 256–271. Springer, 2015.

[Belle *et al.*, 2024] Vaishak Belle, Hana Chockler, Shannon Vallor, Kush R Varshney, Joost Vennekens, and Sander Beckers. Trustworthiness and responsibility in ai-causality, learning, and verification (dagstuhl seminar 24121). *Dagstuhl Reports*, 14:75–91, 2024.

[Braham and van Hees, 2012] Matthew Braham and Martin van Hees. An anatomy of moral responsibility. *Mind*, 121:601–634, 2012.

[Canny, 1988] John Canny. Some algebraic and geometric computations in pspace. In *Proceedings of the twentieth annual ACM symposium on Theory of computing*, pages 460–467, 1988.

[Chatila *et al.*, 2021] Raja Chatila, Virginia Dignum, Michael Fisher, Fosca Giannotti, Katharina Morik, Stuart Russell, and Karen Yeung. Trustworthy ai. *Reflections on artificial intelligence for humanity*, pages 13–39, 2021.

[Chen and Lu, 2007] Taolue Chen and Jian Lu. Probabilistic alternating-time temporal logic and model checking algorithm. In *FSKD*, pages 35–39. IEEE Computer Society, 2007.

[Chockler and Halpern, 2004] Hana Chockler and Joseph Y. Halpern. Responsibility and blame: A structural-model approach. *J. Artif. Intell. Res.*, 22:93–115, 2004.

[Chun, 1989] Youngsub Chun. A new axiomatization of the shapley value. *Games and Economic Behavior*, 1(2):119–130, 1989.

[Ciuni and Mastop, 2009] Roberto Ciuni and Rosja Mastop. Attributing distributed responsibility in stit logic. In *Logic, Rationality, and Interaction: Second International Workshop, LORI 2009, Chongqing, China, October 8-11, 2009. Proceedings 2*, pages 66–75. Springer, 2009.

[Cristau *et al.*, 2010] Julien Cristau, Claire David, and Florian Horn. How do we remember the past in randomised strategies? *Proceedings of GandALF 2010*, 25, 2010.

[De Giacomo *et al.*, 2025] Giuseppe De Giacomo, Emiliano Lorini, Timothy Parker, and Gianmarco Parretti. Responsibility anticipation and attribution in ltlf. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence, IJCAI*, 2025.

[De Lima *et al.*, 2010] Tiago De Lima, Lambér Royakkers, and Frank Dignum. A logic for reasoning about responsibility. *Logic Journal of the IGPL*, 18(1):99–117, 2010.

[Dignum, 2020] Virginia Dignum. Responsibility and artificial intelligence. *The oxford handbook of ethics of AI*, 4698:215, 2020.

[Friedenberg and Halpern, 2019] Meir Friedenberg and Joseph Y. Halpern. Blameworthiness in multi-agent settings. In *AAAI*, pages 525–532, 2019.

[Gerstenberg *et al.*, 2011] Tobias Gerstenberg, Anastasia Ejova, and David Lagnado. Blame the skilled. In *Proceedings of the annual meeting of the cognitive science society*, volume 33, 2011.

[Halpern and Kleiman-Weiner, 2018] Joseph Y. Halpern and Max Kleiman-Weiner. Towards formal definitions of blameworthiness, intention, and moral responsibility. In *AAAI*, pages 1853–1860, 2018.

[Halpern and Pearl, 2005] Joseph Y Halpern and Judea Pearl. Causes and explanations: A structural-model approach. part i: Causes. *The British journal for the philosophy of science*, 2005.

[Hart and Mas-Colell, 1989] Sergiu Hart and Andreu Mas-Colell. Potential, value, and consistency. *Econometrica: Journal of the Econometric Society*, pages 589–614, 1989.

[Icarte *et al.*, 2022] Rodrigo Toro Icarte, Toryn Q Klassen, Richard Valenzano, and Sheila A McIlraith. Reward machines: Exploiting reward function structure in reinforcement learning. *Journal of Artificial Intelligence Research*, 73:173–208, 2022.

[Kaur *et al.*, 2022] Davinder Kaur, Suleyman Uslu, Kaley J Rittichier, and Arjan Durresi. Trustworthy artificial intelligence: a review. *ACM computing surveys (CSUR)*, 55(2):1–38, 2022.

[Kobayashi *et al.*, 2023] Tsutomu Kobayashi, Martin Bondu, and Fuyuki Ishikawa. Formal modelling of safety architecture for responsibility-aware autonomous vehicle via event-b refinement. In *International Symposium on Formal Methods*, pages 533–549. Springer, 2023.

[Kwiatkowska *et al.*, 2018] Marta Kwiatkowska, Gethin Norman, David Parker, and Gabriel Santos. Automated verification of concurrent stochastic games. In *Quantitative Evaluation of Systems: 15th International Conference, QEST 2018, Beijing, China, September 4-7, 2018, Proceedings 15*, pages 223–239. Springer, 2018.

[Kwiatkowska *et al.*, 2019] Marta Kwiatkowska, Gethin Norman, David Parker, and Gabriel Santos. Equilibria-based probabilistic model checking for concurrent stochastic games. In *International Symposium on Formal Methods*, pages 298–315. Springer, 2019.

[Kwiatkowska *et al.*, 2021] Marta Kwiatkowska, Gethin Norman, David Parker, and Gabriel Santos. Automatic verification of concurrent stochastic systems. *Formal Methods in System Design*, 58(1):188–250, 2021.

[Lewis, 2013] David Lewis. *Counterfactuals*. John Wiley & Sons, 2013.

[Li *et al.*, 2023] Bo Li, Peng Qi, Bo Liu, Shuai Di, Jingen Liu, Jiquan Pei, Jinfeng Yi, and Bowen Zhou. Trustworthy ai: From principles to practices. *ACM Computing Surveys*, 55(9):1–46, 2023.

[Mu and Najib, 2026] Chunyan Mu and Muhammad Najib. Counterfactual reasoning for causal responsibility attribution in probabilistic multi-agent systems. *CoRR*, abs/2605.13077, 2026.

[Mu *et al.*, 2025] Chunyan Mu, Muhammad Najib, and Nir Oren. Responsibility-aware strategic reasoning in probabilistic multi-agent systems. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 23258–23266, 2025.

[Naumov and Tao, 2019] Pavel Naumov and Jia Tao. Blameworthiness in strategic games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 3011–3018, 2019.

[Naumov and Tao, 2021] Pavel Naumov and Jia Tao. Two forms of responsibility in strategic games. In *IJCAI*, pages 1989–1995. ijcai.org, 2021.

[Osborne and Rubinstein, 1994] Martin J. Osborne and Ariel Rubinstein. *A course in game theory*. The MIT Press, 1994.

[Parker *et al.*, 2023] Timothy Parker, Umberto Grandi, and Emiliano Lorini. Anticipating responsibility in multiagent planning. In *ECAI*, volume 372, pages 1859–1866, 2023.

[Pearl, 2009] Judea Pearl. *Causality*. Cambridge university press, 2009.

[Perelli, 2019] Giuseppe Perelli. Enforcing equilibria in multi-agent systems. In *AAMAS*, volume 1, pages 188–196. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), 2019.

[Shapley, 1953] Lloyd S. Shapley. A value for n-person games. In Harold W. Kuhn and Albert W. Tucker, editors, *Contributions to the Theory of Games II*, pages 307–317. Princeton University Press, Princeton, 1953.

[Shoham and Leyton-Brown, 2009] Yoav Shoham and Kevin Leyton-Brown. *Multiagent Systems - Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press, 2009.

[Stan *et al.*, 2024] Daniel Stan, Muhammad Najib, Anthony Widjaja Lin, and Parosh Aziz Abdulla. Concurrent stochastic lossy channel games. In *32nd EACSL Annual Conference on Computer Science Logic (CSL 2024)*, pages 46–1. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2024.

[Ummels and Wojtczak, 2011] Michael Ummels and Dominik Wojtczak. The complexity of nash equilibria in stochastic multiplayer games. *Logical Methods in Computer Science*, 7, 2011.

[van den Brink, 2002] René van den Brink. An axiomatization of the shapley value using a fairness property. *International Journal of Game Theory*, 30:309–319, 2002.

[Wooldridge, 2012] Michael Wooldridge. Does game theory work? *IEEE Intelligent Systems*, 27(6):76–80, 2012.

[Wright, 1985] Richard W Wright. Causation in tort law. *Calif. L. Rev.*, 73:1735, 1985.

[Yazdanpanah *et al.*, 2019] Vahid Yazdanpanah, Mehdi Dastani, Wojciech Jamroga, Natasha Alechina, and Brian Logan. Strategic responsibility under imperfect information. In *AAMAS*, pages 592–600, 2019.

[Yazdanpanah *et al.*, 2023] Vahid Yazdanpanah, Enrico H Gerding, Sebastian Stein, Mehdi Dastani, Catholijn M Jonker, Timothy J Norman, and Sarvapali D Ramchurn. Reasoning about responsibility in autonomous systems: challenges and opportunities. *Ai & Society*, 38(4):1453–1464, 2023.

[Young, 1985] Peyton H Young. Monotonic solutions of cooperative games. *International Journal of Game Theory*, 14(2):65–72, 1985.