Proceedings of the 23rd International Conference on Principles of Knowledge Representation and Reasoning
KR meets Machine Learning and Explanation

Semantic Foundations of Neuro-Symbolic Multi-Agent Systems

Julian Gutierrez
University of Sussex
J.Gutierrez@sussex.ac.uk

Abstract

Neuro-symbolic AI aims to integrate learning-based and symbolic reasoning components within a unified framework. While most existing work focuses on single-agent settings and engineering architectures, formal foundations for neuro-symbolic multi-agent systems remain limited. In this paper, we introduce a game-theoretic formal model capturing the interaction between probabilistic neural evaluation and symbolic strategic reasoning in multi-agent environments. The framework extends logical models of strategic reasoning while embedding probabilistic propositional inference into a distributed setting. We establish basic properties of the model and provide a comprehensive complexity-theoretic analysis of optimally stable strategic behaviour. In particular, for one-shot games, we show that utility evaluation corresponds to weighted model counting and characterise the complexity of the main associated Nash equilibrium problems, ranging from FP and #P to PP and Σ_2^{PP}. We further study an iterated variant of the core model, showing PSPACE-completeness for the main decision problem in such a class of multi-player games. These results provide formal foundations for reasoning about neuro-symbolic multi-agent systems and clarify the computational limits of combining learning, uncertainty, and strategic interaction within the same formal reasoning framework.

1 Introduction

Neuro-symbolic AI seeks to integrate two traditionally distinct paradigms: symbolic reasoning and sub-symbolic learning (Bhuyan et al. 2024). Symbolic methods provide formal structure, compositional representations, and rigorous reasoning capabilities, whereas neural approaches enable data-driven learning and robustness in the presence of noise and uncertainty. The combination of these paradigms has recently gained relevance as a principled approach to constructing AI systems capable of acquiring knowledge from data and performing structured reasoning over that knowledge. From a foundational perspective, neuro-symbolic AI could bridge learning and reasoning within a unified mathematical framework. This suggests the need for further work into the integration of logic-based inference with neural computation, probabilistic reasoning, and differentiable models, emphasising the need for semantically well-founded models that capture the interaction between learned representations and symbolic reasoning methods (Besold et al. 2021; Dong et al. 2019).

Despite the rapid progress of neuro-symbolic research, the majority of existing models have been developed in a predominantly single-agent setting (Bhuyan et al. 2024). From a theoretical standpoint, this restriction limits the applicability of current frameworks, as many real-world domains inherently involve multiple autonomous decision-makers whose actions interact. Examples arise in distributed AI, cooperative and competitive multi-agent planning, autonomous robotics, and socio-technical systems, where reasoning must account not only for uncertainty and learning but also for strategic behaviour and interaction. While symbolic multi-agent reasoning and game-theoretic models are well-established in the literature, their integration with learning-based neuro-symbolic components remains comparatively underexplored. This gap highlights the need for formal models capable of capturing learning, uncertainty, and strategic interaction within a unified multi-agent setting, providing a foundation for analysing decision-making in complex neuro-symbolic environments.

A second fundamental challenge concerns the absence of general formal foundations for neuro-symbolic multi-agent systems. Many existing neuro-symbolic architectures combine learned predictors with symbolic decision layers in an application-specific and largely procedural manner, often without a precise mathematical semantics or complexity-theoretic characterisation. Consequently, core theoretical questions remain unresolved, for instance: How should learning-based predictions interact with symbolic reasoning in multi-agent environments? What guarantees can be provided about the behaviour of such systems? How computationally demanding is reasoning in the presence of both uncertainty and strategic interaction? Addressing these issues requires the development of a rigorous formal framework that unifies probabilistic learning, symbolic reasoning, and multi-agent decision-making, enabling principled analysis of correctness, expressiveness, and computational complexity in neuro-symbolic multi-agent systems.

In this paper, we address these issues and propose *Neuro-Symbolic Boolean Games* (NSBGs), a formal game-theoretic model for neuro-symbolic multi-agent systems. Our framework builds on Boolean games (Harrenstein et al. 2001), a well-established logical model of strategic reasoning that generalises distributed satisfiability. In NSBGs, each agent combines a symbolic goal expressed in propositional logic with a neural valuation mechanism that predicts uncertain truth values based on learned information. Utilities are defined as probabilities of goal satisfaction under these learned probabilistic valuations. This yields a unified and mathematically precise model capturing the interaction between learning-based perception and strategic control, allowing neuro-symbolic learning and reasoning in a rather general strategic and multi-agent systems framework.

Conceptually, NSBGs constitute a conservative extension of classical Boolean games and, at the same time, lift probabilistic propositional reasoning into a distributed and multi-agent framework. The proposed model clearly disentangles three fundamental components: symbolic objectives, strategic control, and neural-based evaluation. This modular structure enables a wide range of neuro-symbolic architectures to be captured within a single uniform formalism. As a result, the NSBG framework supports rigorous theoretical investigation while remaining expressive enough to model practical hybrid systems in which learning mechanisms and symbolic reasoning interact in a principled manner.

Our main technical contribution is the introduction of a new formal model for neuro-symbolic multi-agent AI systems together with a comprehensive complexity-theoretic analysis of its reasoning capabilities. We first develop the mathematical foundations of the framework and establish its basic properties, showing that utility evaluation corresponds to probabilistic inference and weighted model counting. This yields a refined complexity landscape ranging from FP and #P for exact evaluation to PP for threshold reasoning. We then analyse equilibrium computation in one-shot NSBGs, proving that checking whether a profile is a Nash equilibrium is $\mathrm{coNP}^{\mathrm{PP}}$-complete, while deciding the existence of an equilibrium is Σ_2^{PP}-complete. Extending the analysis to dynamic interaction, we investigate iterated variants of the model and show that equilibrium reasoning becomes PSPACE-complete in finite-horizon settings, whereas the fully general formulation is undecidable. Overall, these results delineate the computational boundaries of neuro-symbolic multi-agent reasoning and demonstrate how combining probabilistic evaluation with strategic interaction fundamentally increases computational complexity.

Beyond its technical depth, this paper provides foundational insight into the nature of reasoning in neuro-symbolic multi-agent systems. The proposed framework is, to the best of our knowledge, the first to offer a mathematically precise and complexity-theoretically complete account of strategic reasoning under learned probabilistic valuations. By unifying symbolic goals, probabilistic inference, and game-theoretic interaction within a single formal model, our results clarify the intrinsic sources of computational hardness in hybrid learning–reasoning systems. As such, this work establishes a new theoretical baseline for analysing multi-agent systems that combine data-driven perception with logical decision making. From an applied perspective, understanding these complexity boundaries is essential for designing scalable algorithms, identifying tractable fragments, and guiding the development of reliable neuro-symbolic architectures in domains such as distributed AI, autonomous decision making, and strategic reasoning under uncertainty.

Structure of the paper. Section 2 introduces Neuro-Symbolic Boolean Games (NSBGs), and Section 3 illustrates how different neuro-symbolic architectures can be captured by this model. Section 4 establishes core properties of the model, including probabilistic reasoning and utility evaluation. Section 5 studies the one-shot version of NSBGs and characterises the complexity of Nash equilibrium reasoning. Section 6 extends the analysis to some iterated variants of the model and investigates the resulting computational behaviour. Finally, Section 7 concludes with a discussion of related work and future research directions.

2 Neuro-Symbolic Multi-Agent Systems

Propositional variables and formulae. Let V be a finite set of propositional (Boolean) variables. A *(propositional) formula* over V is any expression generated by the grammar

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \psi) \mid (\varphi \vee \psi) \quad \text{where } p \in V.$$

We write $\mathsf{Prop}(V)$ for the set of all such formulae.

Satisfaction. A *valuation* is a function $v : V \to \{0,1\}$. The satisfaction relation $v \models \varphi$ is defined inductively:

- $v \models p$ iff $v(p) = 1$ for $p \in V$;
- $v \models \neg\varphi$ iff not $v \models \varphi$;
- $v \models (\varphi \wedge \psi)$ iff $v \models \varphi$ and $v \models \psi$;
- $v \models (\varphi \vee \psi)$ iff $v \models \varphi$ or $v \models \psi$.

Agents and control. Let $Ag = \{1, \dots, n\}$ be a set of agents. We partition (denoted by $\uplus$) the variable set V into

$$V \;=\; V_{\mathrm{ctrl}} \uplus V_{\mathrm{env}},$$

where V_{ctrl} are *controlled* variables and V_{env} are *environmental* (uncontrolled) variables. Moreover,

$$V_{\mathrm{ctrl}} \;=\; \biguplus_{i \in Ag} V_i,$$

where V_i is the set of variables controlled by agent i.

Control layer: Strategies and profiles. A *(pure) strategy* for agent i is an assignment

$$s_i : V_i \to \{0,1\}.$$

A *strategy profile* is an n-tuple $s = (s_1, \dots, s_n)$. Every profile induces a valuation on controlled variables, denoted as $v_s : V_{\mathrm{ctrl}} \to \{0,1\}$ and defined by

$$v_s(p) = s_i(p) \quad \text{for the unique } i \text{ such that } p \in V_i.$$

Environment parameters. Fix a feature space E and a probability distribution $\mathcal{D}$ over E. Elements $e \in E$ are feature vectors (possibly non-Boolean) representing environmental conditions and serving as inputs to neural perception mechanisms. These environmental parameters are treated as global and remain fixed throughout the model.

Probabilistic valuations (Bernoulli interpretation). A *probabilistic valuation* is a function

$$\hat{v} : V \to \{0,1\} \times [0,1],$$

where $\hat{v}(p) = (b_p, c_p)$ is interpreted as a Bernoulli belief:

$$\Pr(p = b_p) = c_p, \qquad \Pr(p \neq b_p) = 1 - c_p.$$

We assume independence across variables (product distribution). The value b_p represents the predicted Boolean value of p, while c_p denotes the confidence in that prediction.

Probability measures. Given a probabilistic valuation $\hat{v}$, let $\mu_{\hat{v}}$ be the induced product Bernoulli distribution over classical valuations:

$$\mu_{\hat{v}}(v) \;=\; \prod_{p \in V} \begin{cases} c_p & \text{if } v(p) = b_p, \\ 1 - c_p & \text{otherwise.} \end{cases}$$

Probability of satisfaction. For any $\varphi \in \mathsf{Prop}(V)$,

$$\Pr_{\hat{v}}(\varphi) \;=\; \sum_{v \models \varphi} \mu_{\hat{v}}(v).$$

Neural layer: Neural valuation mechanisms. First, let $S = \prod_{i \in Ag} \{0,1\}^{V_i}$ be the set of all pure strategy profiles, an $|Ag|$-tuple of players' choices over $V_{\text{ctrl}} = V_1 \uplus \ldots \uplus V_n$. Each agent i is equipped with two neural mechanisms:

$$\mathcal{N}_i^{\text{In}} : E \to (\{0,1\} \times [0,1])^{V_{\text{env}}},$$
$$\mathcal{N}_i^{\text{Out}} : S \to (\{0,1\} \times [0,1])^{V_{\text{ctrl}}}.$$

- $\mathcal{N}_i^{\text{In}}$ models *perception/sensing*: given a feature vector $e \in E$, this sub-component predicts probabilistic truth values for environmental variables.
- $\mathcal{N}_i^{\text{Out}}$ models *execution/actuation*: given a strategy profile s, this sub-component produces the probabilistic realised values of controlled variables.

Neural proxy of the symbolic world. For each agent i, profile $s \in S$, and feature vector $e \in E$, the neural mechanisms induce a probabilistic valuation

$$\hat{v}_i(s,e) : V \to \{0,1\} \times [0,1]$$

defined by

$$\hat{v}_i(s,e)(p) = \begin{cases} \mathcal{N}_i^{\text{Out}}(s)(p) & \text{if } p \in V_{\text{ctrl}}, \\ \mathcal{N}_i^{\text{In}}(e)(p) & \text{if } p \in V_{\text{env}}. \end{cases}$$

This valuation represents agent i's neural proxy of the symbolic world under profile s and environmental conditions e. We write

$$\hat{v}_i(s,e)(p) = (b_{i,s,e,p},\, c_{i,s,e,p})$$

for the predicted value and confidence associated with variable p. A *family of neural valuation mechanisms* is

$$\mathcal{N} = \left(\mathcal{N}_i^{\text{In}}, \mathcal{N}_i^{\text{Out}}\right)_{i \in Ag}.$$

Neuro-Symbolic Boolean Game (NSBG). An NSBG (parameterised by a fixed environment $(E, \mathcal{D})$) is a tuple

$$\mathcal{G} = \langle Ag, V, \Gamma, \mathcal{N} \rangle$$

where:

- Ag is a finite set of agents;
- $V = V_{\text{ctrl}} \uplus V_{\text{env}}$ with $V_{\text{ctrl}} = (V_i)_{i \in Ag}$;
- $\Gamma = (\gamma_i)_{i \in Ag}$ with each $\gamma_i \in \mathsf{Prop}(V)$;
- $\mathcal{N} = (\mathcal{N}_i^{\text{In}}, \mathcal{N}_i^{\text{Out}})_{i \in Ag}$.

Utility. For each agent i and strategy profile $s \in S$, define

$$u_i(s) = \mathbb{E}_{e \sim \mathcal{D}}\big[\Pr_{\hat{v}_i(s,e)}(\gamma_i) \big] \in [0,1].$$

Thus utility depends only on the strategy profile, while uncertainty over environmental features, due to probabilistic expectation, is averaged via the fixed distribution $\mathcal{D}$.

Nash equilibrium. A profile $s \in S$ is a *pure Nash equilibrium* if for every $i \in Ag$ and alternative strategy s'_i,

$$u_i(s) \;\geq\; u_i(s'),$$

with $s' = (s'_i, s_{-i}) = (s_0, \ldots, s_{i-1}, s'_i, s_{i+1}, \ldots, s_n)$. That is, no player i has an incentive to unilaterally deviate. Pure Nash equilibrium (PNE) is the simplest game-theoretic solution concept that we will consider to reason about neuro-symbolic multi-agent systems modelled as NSBGs.

3 Neuro-Symbolic Architectures

In this section we illustrate how different structural assumptions on the neural perception and execution mechanisms

$$\mathcal{N} = \left(\mathcal{N}_i^{\text{In}}, \mathcal{N}_i^{\text{Out}}\right)_{i \in Ag}$$

give rise to different neuro-symbolic architectures, each corresponding to a particular way in which agents perceive the environment and realise their controlled actions through learned components. In turn, these different neural features of otherwise equivalent AI systems, may induce game representations with drastically different stable points, that is, with different equilibria, where learning will be anchored.

Let us then illustrate a few different learning architectures, most of which are highly abstracted simplifications of problems that can be seen as relatively standard classification and decision-making under uncertainty scenarios.

Architecture 1: Shared neural perception. All agents share the same perception mechanism of the environment, while execution may remain agent-specific. Formally, for this modelling specification there exists a mapping

$$\mathcal{N}^{\text{In,sh}} : E \to (\{0,1\} \times [0,1])^{V_{\text{env}}}$$

such that for every $i \in Ag$ and every feature vector $e \in E$,

$$\mathcal{N}_i^{\text{In}}(e) = \mathcal{N}^{\text{In,sh}}(e).$$

This models scenarios where agents rely on a common sensing or predictive system (for instance, a shared diagnostic neural network). All agents therefore share the same probabilistic interpretation of environmental variables.

Example 1 (Shared neural perception)**.** *Consider a distributed safety-control scenario with two agents. Let*

$$V_1 = \{a_1\}, \qquad V_2 = \{a_2\}, \qquad V_{\text{env}} = \{t, f\},$$

where a_1, a_2 represent mitigation actions, t denotes that system temperature is safe, and f denotes that a fault is present.

Assume both agents use a shared perception mechanism $\mathcal{N}^{\text{In,sh}}$ that processes sensor features $e \in E$ and predicts probabilistic values for t and f. For instance,

$$\Pr(t = 1) = 0.93, \qquad \Pr(f = 1) = 0.08.$$

Execution mechanisms $\mathcal{N}_i^{\text{Out}}$ may introduce noise in the realised action variables a_1, a_2. Let the agents' goals be

$$\gamma_1 = a_1 \wedge t \wedge \neg f, \qquad \gamma_2 = a_2 \wedge t.$$

Since perception is shared, both agents evaluate their goals under the same uncertain view of the environment, while retaining independent control over their actions.

Architecture 2: Agent-specific neural perception. Each agent possesses its own perception mechanism, and therefore may hold different probabilistic beliefs about the environment. Moreover, agents perfectly realise their own actions, so execution of controlled variables is deterministic. Formally, for every agent i, profile s, and $p \in V_i$,

$$\mathcal{N}_i^{\text{Out}}(s)(p) = (v_s(p), 1).$$

This architecture models agents equipped with private predictive models trained on different data sources, while perfectly accounting for their own actions, that is, assuming that own actuators are deployed using noise-free channels.

Example 2 (Agent-specific neural perception)**.** *Consider competing firms deciding whether to launch a product. Let*

$$V_1 = \{p_1\}, \qquad V_2 = \{p_2\}, \qquad V_{\text{env}} = \{d, c\},$$

where d denotes high market demand and c a stable supply chain. Agents have independent perception mechanisms $\mathcal{N}_1^{\text{In}}$ and $\mathcal{N}_2^{\text{In}}$ derived from private predictive models. For example, suppose that given the learnt data, we have:

$$\Pr_{\mathcal{N}_1^{\text{In}}}(d = 1) = 0.7, \qquad \Pr_{\mathcal{N}_1^{\text{In}}}(c = 1) = 0.4,$$

while agent 2 may hold different beliefs. Execution of p_1 and p_2 is deterministic. Moreover, let the goals of the agents be

$$\gamma_1 = (p_1 \wedge d \wedge c) \ \vee \ (\neg p_2 \wedge d),$$
$$\gamma_2 = (p_2 \wedge d \wedge c) \ \vee \ (\neg p_1 \wedge d).$$

This example shows how heterogeneous learned perceptions induce distinct probabilistic interpretations of the same symbolic world, while agents retain perfect knowledge of their own actions, thus combining different epistemic issues.

Architecture 3: Shared perception with self-certainty. In this mixed architecture, agents share a common perception mechanism while execution of their controlled variables is deterministic. Thus, for all $i \in Ag$, $s \in S$, and $p \in V$,

$$\hat{v}_i(s, e)(p) = \begin{cases} (v_s(p), 1) & \text{if } p \in V_i, \\ \mathcal{N}^{\text{In,sh}}(e)(p) & \text{if } p \in V_{\text{env}}, \\ \mathcal{N}_j^{\text{Out}}(s)(p) & \text{if } p \in V_j, \ j \neq i. \end{cases}$$

This corresponds to a shared learned world model combined with perfect self-observation of one's own actions. It is common in distributed AI systems where sensing is centralised but actuation is locally controlled.

It is worth noting that the three architectures above illustrate how different structural assumptions on perception and execution lead to different strategic models. In particular, changing the neural architecture may alter the utilities induced by identical symbolic goals and therefore modify the set of equilibria of the same underlying game. The modular structure of the NSBG framework allows many additional architectures to be defined by specifying suitable perception and execution mechanisms, enabling the formal study of a wide spectrum of neuro-symbolic multi-agent systems.

4 Basic Properties

For the complexity statements in this section we assume that every confidence value appearing in any probabilistic valuation is a rational number in $[0, 1]$ encoded in binary, *e.g.* as a pair of binary integers (a, b) with $b > 0$ representing a/b. (Equivalently, any standard exact binary encoding of rationals in $[0, 1]$ may be used.) All propositional formulas are given in general form (so, not necessarily in CNF/DNF).

Classical Boolean games. We begin by considering two simpler cases, which will highlight a number of complexity lower bounds for NSBGs. First, classical Boolean games arise as the special case of NSBGs in which $V_{\text{env}} = \emptyset$ and the neural execution mechanisms are deterministic, that is,

$$\mathcal{N}_i^{\text{Out}}(s)(p) = (v_s(p), 1) \quad \text{for all } i \in Ag, \ s \in S, \ p \in V_{\text{ctrl}}.$$

In this case, utilities become deterministic and coincide with the classical Boolean-game utilities. On the other hand, if, contrarily to the case of Boolean games, we assume that $V_{\text{ctrl}} = \emptyset$, then each agent in the game has a singleton strategy set, say $S = \{s^\star\}$, and strategic interaction disappears. Utilities reduce to $u_i(s^\star) = \mathbb{E}_{e \sim \mathcal{D}}\big[\Pr_{\hat{v}_i(s^\star, e)}(\gamma_i)\big]$, where the probabilistic valuation $\hat{v}_i(s^\star, e)$ depends only on the neural perception mechanism $\mathcal{N}_i^{\text{In}}$. As such, NSBGs become probabilistic propositional models in which utility evaluation depends only on the probabilistic perception layer and the distribution $\mathcal{D}$, which we analyse next.

Utility evaluation and weighted model counting. We now analyse the computational complexity of evaluating the *inner probabilistic satisfaction* that underlies utilities in NSBGs. Fix a probabilistic valuation $\hat{v} : V \to \{0, 1\} \times [0, 1]$ with $\hat{v}(p) = (b_p, c_p)$, and let $\mu_{\hat{v}}$ be the induced product Bernoulli distribution. (Equivalently, this corresponds to fixing a feature vector e and the resulting neural valuation $\hat{v}_i(s, e)$.) Then, for any sampled e from the environment, we consider two computational problems, useful later:

- U-VAL (Exact satisfaction probability).
  Input: $(V, \varphi, \hat{v})$.
  Question: What is the exact rational value $\Pr_{\hat{v}}(\varphi)$?
- U-THRESH (Threshold satisfaction probability).
  Input: $(V, \varphi, \hat{v}, \tau)$, with $\tau \in [0, 1] \cap \mathbb{Q}$ given in binary.
  Question: Is $\Pr_{\hat{v}}(\varphi) \geq \tau$?

To study the complexity of answering these questions in a weighted model counting setting, we will write each confidence c_p as a reduced binary-encoded rational $c_p = A_p/B_p$ with integers $0 \leq A_p \leq B_p$ and $B_p > 0$. Define integer weights for each variable $p \in V$ by

$$w_p(x) = \begin{cases} A_p & \text{if } x = b_p, \\ B_p - A_p & \text{if } x \neq b_p, \end{cases}$$

so that the weight depends on whether the assignment agrees with the predicted bit b_p. Concretely, for each assignment $v : V \to \{0,1\}$ define the integer weight

$$W(v) = \prod_{p \in V} \begin{cases} A_p & \text{if } v(p) = b_p, \\ B_p - A_p & \text{if } v(p) \neq b_p, \end{cases}$$

and the common denominator $D = \prod_{p \in V} B_p$. Then

$$\mu_{\hat{v}}(v) = \frac{W(v)}{D} \quad \text{and hence} \quad \Pr_{\hat{v}}(\varphi) = \frac{1}{D} \sum_{v \models \varphi} W(v).$$

The following observation is standard in weighted model counting (see, *e.g.*, (Roth 1996)), but we include a short proof here for completeness in our NSBG setting.

Lemma 1. *Let $F(V, \varphi, \hat{v}) = \sum_{v \models \varphi} W(v)$. Then, $F \in \#P$.*

Proof. Recall that $\#P$ (Valiant 1979) is the class of functions $f : \{0,1\}^* \to \mathbb{N}$ such that there exists a nondeterministic polynomial-time Turing machine M for which $f(x)$ equals the number of accepting computation paths of M on input x. Then, to prove containment, consider an NP Turing machine M that on input $(V, \varphi, \hat{v})$ proceeds as follows.

1. Nondeterministically guess a valuation $v : V \to \{0,1\}$ (one bit per variable).
2. Check in polynomial time whether $v \models \varphi$ (formula evaluation is linear in $|\varphi|$).
3. If $v \not\models \varphi$, then reject on all branches; otherwise create $W(v)$ accepting computation paths.

It remains to show how to generate exactly $W(v)$ accepting paths in polynomial time given v. Observe that $W(v) = \prod_{p \in V} w_p(v(p))$ is the product of $|V|$ binary-encoded integers, each of size at most polynomial in the input length; hence $\log W(v)$ is polynomial in the input size, and therefore $W(v)$ has at most exponential magnitude. Let $k = \lceil \log_2(W(v) + 1) \rceil$, which is polynomial in the input length. The machine branches on k fresh nondeterministic bits, producing 2^k computation paths, and interprets the bits as an integer $r \in \{0, \dots, 2^k - 1\}$. It accepts if and only if $r < W(v)$, which can be checked by binary comparison in polynomial time. Exactly $W(v)$ of the 2^k paths accept. Therefore the total number of accepting paths of M on input $(V, \varphi, \hat{v})$ equals $\sum_{v \models \varphi} W(v) = F(V, \varphi, \hat{v})$. Hence, $F \in \#P$. $\square$

The lemma above can be used to prove the following simple result, which will be useful when reasoning about utilities and satisfaction of goals in NSBGs.

Proposition 2. U-VAL *is* $\mathrm{FP}^{\#P}$*-complete.*

Proof. Membership. By Lemma 1, the numerator $F(V, \varphi, \hat{v}) = \sum_{v \models \varphi} W(v)$ is a $\#P$ function, and hence can be computed by one call to a $\#P$ oracle. The denominator $D = \prod_{p \in V} B_p$ is computable in deterministic polynomial time from the binary encoding of $\hat{v}$. Therefore the exact value $\Pr_{\hat{v}}(\varphi) = \frac{F}{D}$ can be computed in $\mathrm{FP}^{\#P}$ as a rational number written in binary.

Hardness. We reduce $\#$SAT to U-VAL. Let ψ be a CNF formula over variables V. Define a probabilistic valuation $\hat{v}$ by setting $\hat{v}(p) = (1, 1/2)$ for every $p \in V$. Then, $\mu_{\hat{v}}$ is the uniform distribution over all $2^{|V|}$ valuations, and therefore we have the following equality:

$$\Pr_{\hat{v}}(\psi) = \frac{\#\mathrm{SAT}(\psi)}{2^{|V|}}.$$

Hence, given the exact value of $\Pr_{\hat{v}}(\psi)$, one can compute $\#\mathrm{SAT}(\psi)$ by multiplying by $2^{|V|}$, which is polynomial-time computable from the input size. This gives a polynomial-time Turing reduction $\#\mathrm{SAT} \leq_T$ U-VAL, establishing $\mathrm{FP}^{\#P}$-hardness. $\square$

We now move onto study U-THRESH, and show that this problem is complete for PP.

Proposition 3. U-THRESH *is PP-complete.*

Proof. Hardness. We reduce from MAJSAT, which is PP-complete (Gill 1977): given a propositional formula ψ over variables V, decide whether at least half of all valuations satisfy ψ. Given ψ, define a probabilistic valuation $\hat{v}$ by setting $\hat{v}(p) = (1, 1/2)$ for every $p \in V$, and let $\tau = 1/2$. Then, $\mu_{\hat{v}}$ is the uniform distribution over all $2^{|V|}$ valuations, and equivalence between the following two expression holds,

$$\Pr_{\hat{v}}(\psi) = \frac{\#\mathrm{SAT}(\psi)}{2^{|V|}} \geq \frac{1}{2} \quad \Longleftrightarrow \quad \#\mathrm{SAT}(\psi) \geq 2^{|V|-1},$$

which is the MAJSAT predicate. Thus, as a consequence, MAJSAT $\leq_m$ U-THRESH, proving PP-hardness.

Membership. Let the input be $(V, \varphi, \hat{v}, \tau)$, where $\tau = A/B$ is a rational with $A, B \in \mathbb{N}$ and $B > 0$. Let $F = \sum_{v \models \varphi} W(v)$ and $D = \prod_{p \in V} B_p$, so that $\Pr_{\hat{v}}(\varphi) = F/D$. Then $\Pr_{\hat{v}}(\varphi) \geq \tau \iff F \cdot B \geq D \cdot A$. By Lemma 1, F is a $\#P$ function. Since B is polynomial-time computable from the input, $F \cdot B$ is also $\#P$-definable (by duplicating each accepting branch B times). The threshold $T = D \cdot A$ is computable in deterministic polynomial time. Therefore, the problem reduces to testing if a $\#P$ function is at least a given polynomial-time computable integer, which is decidable in PP. Hence, U-THRESH $\in$ PP. $\square$

With these results in place, note that each utility query $u_i(s) = \mathbb{E}_{e \sim \mathcal{D}}\big[\Pr_{\hat{v}_i(s,e)}(\gamma_i)\big]$ reduces to evaluating probabilistic satisfaction under a family of probabilistic valuations induced by the neural mechanisms in the game. Concretely, we have the following result.

Corollary 4. *Fix an NSBG $\mathcal{G}$, an agent $i \in Ag$, and a profile $s \in S$. Deciding whether $u_i(s) \geq \tau$ is PP-complete.*

Proof. Membership. By definition,

$$u_i(s) = \mathbb{E}_{e\sim\mathcal{D}}\big[\Pr_{\hat{v}_i(s,e)}(\gamma_i)\big] = \sum_{e\in E}\mathcal{D}(e)\cdot \Pr_{\hat{v}_i(s,e)}(\gamma_i),$$

recalling that $\mathcal{D}$ is the probability distribution over E. For each e, deciding whether $\Pr_{\hat{v}_i(s,e)}(\gamma_i) \geq \tau'$ for any rational threshold τ' is in PP (Proposition 3). Since $\mathcal{D}(e)$ and $|E|$ are part of the input and encoded in binary, the weighted sum above can be rewritten as a comparison between a $\#P$-definable integer and a polynomial-time computable threshold. Hence deciding $u_i(s) \geq \tau$ is in PP.

Hardness. We reduce from U-THRESH. Given an instance $(V, \varphi, \hat{v}, \tau)$, construct an NSBG with a single agent controlling no variables (so $S = \{s^\star\}$), and a singleton feature space $E = \{e^\star\}$ with $\mathcal{D}(e^\star) = 1$. Define the neural mechanism with $\hat{v}_1(s^\star, e^\star) = \hat{v}$ and set $\gamma_1 = \varphi$. Then, $u_1(s^\star) = \Pr_{\hat{v}}(\varphi)$, and deciding $u_1(s^\star) \geq \tau$ is U-THRESH. □

Let us now consider a decision problem that will be useful to reason about equilibria. Since beneficial deviations from a given profile are defined using *strict* improvement, we define the following strict utility-comparison problem:

- COMPARE-U$^>$: (Strict utility comparison).
  Input: NSBG $\mathcal{G}$, agent $i \in Ag$, profiles $s, s' \in S$.
  Question: Is $u_i(s) > u_i(s')$?

Proposition 5. COMPARE-U$^>$ *is PP-complete.*

Proof. Membership. For any agent i and profiles s, s', utilities are rational numbers

$$u_i(s) = \frac{F_s}{D_s}, \qquad u_i(s') = \frac{F_{s'}}{D_{s'}},$$

where $F_s, F_{s'}$ are $\#P$ functions (Lemma 1) and $D_s, D_{s'}$ are polynomial-time computable positive integers. Hence $u_i(s) > u_i(s') \iff F_s D_{s'} > F_{s'} D_s$. Define $\#P$ function

$$G = F_s D_{s'} - F_{s'} D_s.$$

Since multiplication by polynomial-time computable integers preserves $\#P$-definability, deciding whether $G > 0$ reduces to comparing a $\#P$ function with a polynomial-time computable threshold, which is decidable in PP. Thus, COMPARE-U$^>$ $\in$ PP.

Hardness. We reduce from U-THRESH, which is PP-complete. Let $(V, \varphi, \hat{v}, \tau)$ be an instance, where $\tau = A/B$ is a binary-encoded rational. Construct an NSBG with one agent and no controlled variables, so the strategy space initially contains a single profile $s^\star$. Define a singleton feature space $E = \{e^\star\}$ with $\mathcal{D}(e^\star) = 1$, and set the neural mechanism so that $\hat{v}_1(s^\star, e^\star) = \hat{v}$ and $\gamma_1 = \varphi$. Then, it follows that $u_1(s^\star) = \Pr_{\hat{v}}(\varphi)$.

We now introduce a second profile s' by adding one dummy controlled variable p. Define the neural mechanism so that, under s', the probabilistic valuation satisfies a fresh propositional variable q with probability exactly τ, independently of all other variables. Let the goal for this profile be $\gamma_1' = q$. Then $u_1(s') = \tau$. Therefore,

$$u_1(s^\star) > u_1(s') \iff \Pr_{\hat{v}}(\varphi) > \tau,$$

which is the predicate of U-THRESH. Since the construction is polynomial, then COMPARE-U$^>$ is PP-hard. □

5 Nash Equilibrium

The problems and results in Section 4 give us a number of lower bounds for some of the main questions we will study in further sections. More specifically, knowing (i) that NS-BGs conservatively extend classical Boolean games (by letting $V_{\text{env}} = \emptyset$ and $\mathcal{N}$ be a family of deterministic neural evaluation mechanisms), and (ii) that reasoning about utilities in an NSBG, even when $V_{\text{ctrl}} = \emptyset$, is a PP-complete problem, tells us that any attempt to compute (Nash) equilibria for these neuro-symbolic games will be at least as hard as these two simpler scenarios. In particular, due to PP-completeness, that any non-trivial problem for these games will be as hard as the complete polynomial hierarchy (PH).

Then, building on results obtained in the previous section, we now study the complexity of reasoning about pure Nash equilibria in Neuro-Symbolic Boolean Games (NSBGs). We assume the same encoding conventions as in the previous section: propositional formulas are given in general form, and confidence values in probabilistic valuations are rational numbers in $[0, 1]$ encoded in binary, which is standard.

Decision problems. Recall that, under our model, utilities are defined by $u_i(s) = \mathbb{E}_{e\sim\mathcal{D}}\big[\Pr_{\hat{v}_i(s,e)}(\gamma_i)\big]$, and utility comparison is PP-complete. Recall also that a profile s is a pure Nash equilibrium iff no agent has a strictly profitable deviation. This leads to the following decision problems:

- IS-PNE.
  Input: NSBG $\mathcal{G}$ and profile $s \in S$.
  Question: Is s a pure Nash equilibrium?
- EXISTS-PNE.
  Input: NSBG $\mathcal{G}$.
  Question: Does $\mathcal{G}$ have a pure Nash equilibrium?

5.1 PNE Checking

We first study the computational complexity of answering the first question above. To do that, we will make use of the following standard NP$^{\text{PP}}$-complete problem:

E-MAJSAT.
Input: Formula $\psi(X, Z)$ with disjoint variable sets.
Question: Does there exist x such that

$$\Pr_{z\sim\text{Unif}}[\psi(x, z)] > \tfrac{1}{2}?$$

Here, Unif denotes the *uniform probability distribution* over all Boolean valuations of the variables in Z, *i.e.*, each valuation $z : Z \to \{0, 1\}$ is chosen with probability $2^{-|Z|}$. Then, E-MAJSAT is NP$^{\text{PP}}$-complete (Arora and Barak 2009). We then can prove the first results of this section.

Theorem 6. IS-PNE *is coNP*$^{\text{PP}}$*-complete.*

Proof. Membership. A profile s is not a pure Nash equilibrium (PNE) if and only if there exists an agent i and a deviation s_i' such that $u_i(s_i', s_{-i}) > u_i(s)$. We can nondeterministically guess (i, s_i') and verify the strict inequality using COMPARE-U$^>$, which is in PP (Proposition 5). Hence, the complement of IS-PNE lies in NP$^{\text{PP}}$, and therefore it follows that IS-PNE $\in$ coNP$^{\text{PP}}$, as required.

Hardness. We reduce from the complement of E-MAJSAT. Given an instance $\psi(X, Z)$ of E-MAJSAT, construct an NSBG with a single agent controlling the variables in X together with one additional mode bit m. Thus,

$$V_{\text{ctrl}} = X \uplus \{m\}, \qquad V_{\text{env}} = Z \uplus \{q\},$$

where q is a fresh environmental variable.

Use a singleton feature space $E = \{e^\star\}$ with $\mathcal{D}(e^\star) = 1$, so utilities reduce to single probabilistic evaluations. The neural valuation mechanism is defined as follows.

- Controlled variables are deterministic:

$$\hat{v}_1(s, e^\star)(p) = (v_s(p), 1) \quad \text{for every } p \in V_{\text{ctrl}}.$$

- Each $z \in Z$ is uniform:

$$\hat{v}_1(s, e^\star)(z) = (1, 1/2).$$

- The fresh variable q is also uniform:

$$\hat{v}_1(s, e^\star)(q) = (1, 1/2).$$

The agent's goal is

$$\gamma_1 = (\neg m \wedge q) \;\vee\; (m \wedge \psi).$$

Consider the profile $s^\star$ in which $m = 0$ and all variables in X are set to 0. Since $m = 0$, the goal γ_1 reduces to q, and therefore

$$u_1(s^\star) = \Pr[q = 1] = \tfrac{1}{2}.$$

Now consider any unilateral deviation s'_1 of the agent.

If s'_1 still sets $m = 0$, then γ_1 again reduces to q, and the utility remains

$$u_1(s'_1) = \tfrac{1}{2}.$$

Thus no deviation with $m = 0$ is profitable.

If s'_1 sets $m = 1$, then the term $\neg m \wedge q$ becomes false and γ_1 reduces to ψ. Moreover, the same deviation also chooses an assignment x for the variables in X, while the variables in Z remain uniform. Hence

$$u_1(s'_1) = \Pr_{z \sim \text{Unif}}[\psi(x, z)].$$

Therefore, a profitable deviation from $s^\star$ exists if and only if there exists an assignment x to X such that

$$\Pr_{z \sim \text{Unif}}[\psi(x, z)] > \tfrac{1}{2},$$

which is exactly the E-MAJSAT predicate.

Consequently, $s^\star$ is a pure Nash equilibrium if and only if the given instance of E-MAJSAT is false. This gives a polynomial-time many-one reduction from the complement of E-MAJSAT to IS-PNE, proving coNP$^{\text{PP}}$-hardness. $\square$

5.2 PNE Existence

We now study the complexity of deciding whether a pure Nash equilibrium exists. To do so, we use the Σ_2^{PP}-complete problem $\exists\forall$-MAJSAT (Toda 1991; Arora and Barak 2009).

Theorem 7. EXISTS-PNE *is* Σ_2^{PP}*-complete.*

Proof. Membership. A pure Nash equilibrium exists iff

$$\exists s \; \forall i \in Ag \; \forall s'_i \quad u_i(s) \geq u_i(s'_i, s_{-i}).$$

For fixed s, i, s'_i, checking $u_i(s) \geq u_i(s'_i, s_{-i})$ is decidable in PP (Proposition 5). Therefore, the problem has the form

$$\exists s \; \forall (i, s'_i) \; R(s, i, s'_i),$$

where R is a PP predicate. Hence, EXISTS-PNE $\in \Sigma_2^{\text{PP}}$.

Hardness. We reduce from $\exists\forall$-MAJSAT, which is Σ_2^{PP}-complete, in a way similar to the reduction for IS-PNE. An instance is a propositional formula $\psi(X, Y, Z)$ with pairwise disjoint variable sets, and the question is whether

$$\exists x \; \forall y \quad \Pr_{z \sim \text{Unif}}[\psi(x, y, z)] \geq \tfrac{1}{2}.$$

Given such an instance, construct an NSBG with three agents. The controlled variables are distributed as follows:

$$V_1 = X \uplus \{a\}, \qquad V_2 = Y \uplus \{m\}, \qquad V_3 = \{b\}.$$

Here, m is a mode bit, as defined before, and a, b are auxiliary bits used to rule out spurious equilibria in which $m = 1$.

As in the hardness proof for IS-PNE, let $V_{\text{env}} = Z \uplus \{q\}$, where q is a fresh environmental variable, use a singleton feature space $E = \{e^\star\}$ with $\mathcal{D}(e^\star) = 1$, and define the neural valuation mechanisms so that controlled variables are deterministic and environmental variables are uniform.

The goals are defined by

$$\gamma_1 = \neg m \vee (a \leftrightarrow b),$$
$$\gamma_2 = (\neg m \wedge q) \;\vee\; (m \wedge \neg\psi),$$
$$\gamma_3 = \neg m \vee \neg(a \leftrightarrow b).$$

With these definitions in place, it follows that the constructed NSBG has a pure Nash equilibrium if and only if the given $\exists\forall$-MAJSAT instance is true. As a consequence, since the construction defined above is polynomial in the size of ψ, it follows that EXISTS-PNE is Σ_2^{PP}-hard. $\square$

The above results establish the complexity of the main Nash equilibrium decision problems for NSBGs under the general neural valuation model introduced in Section 2. While these complexity bounds hold uniformly across all admissible neural evaluation mechanisms, different neuro-symbolic architectures may induce different strategic structures, and therefore different equilibrium sets, even when goals and variables remain unchanged. The following examples illustrate this phenomenon, which highlight the impact that a concrete neural mechanism can have on the set of stable/optimal outcomes of a neuro-symbolic system.

Example 3 (Neural-architecture dependent equilibria)**.** *Consider a two-agent NSBG with* $V_1 = \{x\}$*,* $V_2 = \{y\}$*, and one environmental variable* e*. Let the goals be*

$$\gamma_1 = x \wedge e \wedge \neg y, \qquad \gamma_2 = y \wedge e \wedge \neg x.$$

Assume a simple environment $E = \{e^\star\}$ *and that every neural perception mechanism satisfies* $\Pr_{\hat{v}_i(s,e^\star)}(e = 1) = \frac{1}{2}$*, that is, the only environmental variable* e *is uniform. Based on this, let us analyse the set of equilibria that may arise depending on the type of neural mechanism in place.*

Shared neural perception. *Suppose all agents share the same neural proxy and that both x and y are perceived with confidence $\frac{1}{2}$ independently of the profile. Then all profiles yield the same expected utility for both agents, and* every *profile is a pure Nash equilibrium.*

Shared perception with self-certainty. *Suppose agents remain certain about their own controlled variable while retaining the same uncertainty about the opponent's variable. Then each agent strictly prefers setting its variable to* 1*, and the unique pure Nash equilibrium is* $(x = 1, y = 1)$.

Agent-specific perception. *If, on the other hand, agents hold asymmetric private beliefs about the opponent's realised variable, the equilibrium set may change again, with concrete settings where, for instance, strategy profiles such as* $(1, 0)$ *or* $(0, 1)$ *are in the set of pure Nash equilibria.*

6 Iterated Neuro-Symbolic Boolean Games

In this section, we extend one-shot NSBGs to repeated interaction. We consider a finite-time horizon game and show that equilibrium reasoning becomes PSPACE-complete. As before, we fix the environment parameters $(E, \mathcal{D})$ and assume that all confidence values are rationals in $[0, 1]$.

6.1 Finite-Horizon Iterated NSBGs

Time-unfolding of variables. Let $\mathcal{G} = \langle Ag, V, \Gamma, \mathcal{N} \rangle$ be a one-shot NSBG (parameterised by fixed $(E, \mathcal{D})$) as in Section 2, with $V = V_{\text{ctrl}} \uplus V_{\text{env}}$ and $V_{\text{ctrl}} = \biguplus_{i \in Ag} V_i$. Fix a horizon $T \in \mathbb{N}$ (given as part of the input). For each round $t \in \{0, \ldots, T-1\}$ we create a fresh copy of the variable set

$$\begin{aligned} V^t &= \{ p^t \mid p \in V \}, \\ V_i^t &= \{ p^t \mid p \in V_i \}, \\ V_{\text{env}}^t &= \{ p^t \mid p \in V_{\text{env}} \}. \end{aligned}$$

Intuitively, p^t refers to p at round t. We also lift goals and neural mechanisms to the time-unfolded setting.

Stage strategies and histories. At each round t, each agent i chooses a valuation

$$s_i^t : V_i^t \to \{0, 1\}.$$

A stage profile is $s^t = (s_i^t)_{i \in Ag}$. Let $S^t = \prod_{i \in Ag} \{0, 1\}^{V_i^t}$ be the set of all stage profiles. A history of length t is a sequence $h^t = (s^0, \ldots, s^{t-1})$ of past stage profiles.

Neural mechanisms at each round. In the iterated game, the environment produces a feature vector $e^t \sim \mathcal{D}$ at each round. Each agent retains the two neural mechanisms:

$$\begin{aligned} \mathcal{N}_{i,t}^{\text{In}} &: E \to (\{0, 1\} \times [0, 1])^{V_{\text{env}}^t}, \\ \mathcal{N}_{i,t}^{\text{Out}} &: S^t \to (\{0, 1\} \times [0, 1])^{V_{\text{ctrl}}^t}. \end{aligned}$$

Given $s^t \in S^t$ and $e^t \in E$, these induce a probabilistic valuation $\hat{v}_i^t(s^t, e^t) : V^t \to \{0, 1\} \times [0, 1]$ defined by

$$\hat{v}_i^t(s^t, e^t)(p^t) = \begin{cases} \mathcal{N}_{i,t}^{\text{Out}}(s^t)(p^t) & \text{if } p^t \in V_{\text{ctrl}}^t, \\ \mathcal{N}_{i,t}^{\text{In}}(e^t)(p^t) & \text{if } p^t \in V_{\text{env}}^t. \end{cases}$$

Stage utility and total utility. Let γ_i^t be the formula obtained from γ_i by replacing each variable p by p^t. Define the stage utility of agent i at round t by

$$u_i^t(s^t) = \mathbb{E}_{e^t \sim \mathcal{D}} \Big[\Pr_{\hat{v}_i^t(s^t, e^t)} (\gamma_i^t) \Big] \in [0, 1].$$

The total utility over the T rounds is additive:

$$U_i(h^T) = \sum_{t=0}^{T-1} u_i^t(s^t).$$

Strategies and subgame-perfect equilibrium. A (pure) strategy for agent i is a function σ_i mapping every history h^t to a stage action $\sigma_i(h^t) = s_i^t \in \{0, 1\}^{V_i^t}$. A strategy profile $\sigma = (\sigma_i)_{i \in Ag}$ induces a unique history $h^T(\sigma)$ and utilities $(U_i(h^T(\sigma)))_{i \in Ag}$. Given a history h^t, the continuation from t to $T - 1$ is a subgame. A strategy profile σ is a *subgame-perfect Nash equilibrium (SPNE)* if for every history h^t and every agent i, no unilateral deviation σ_i' yields strictly larger continuation utility. (Formally, σ restricts to a Nash equilibrium in every subgame.)

Main decision problem. In this section, we will consider the main decision question for problems posed as a game:

EXISTS-FH-SPNE.
Input: finite-horizon iterated NSBG $\mathcal{G}$ and horizon T.
Question: Does $\mathcal{G}$ have an SPNE?

Because of the same reason discussed before, under PNE, these games are not guaranteed to have an equilibrium.

Complexity. We show that existence of SPNE is PSPACE-complete, already for two agents and deterministic neural mechanisms (probabilities 0 or 1 only).

Theorem 8. EXISTS-FH-SPNE *is PSPACE-complete.*

Proof. Membership. The extensive-form game induced by a horizon T has depth T and branching exponential in the number of controlled variables per stage. We can explore the game tree by depth-first search while storing only the current history, hence using polynomial space in the input size. At any node (history) we must check whether some agent has a profitable one-step deviation followed by an optimal continuation. This reduces to comparing continuation utilities; by Section 4 such utility comparisons can be implemented via PP oracles (ultimately, weighted model counting / thresholding). Since PSPACE is closed under polynomially many PP-oracle calls, the problem lies in PSPACE.

Hardness. We reduce from QBF, which is a canonical logic-based PSPACE-complete problem. Let

$$\Phi = \exists x_0 \forall x_1 \exists x_2 \cdots Q_{n-1} x_{n-1} \ \psi(x_0, \ldots, x_{n-1})$$

be a quantified Boolean formula, where ψ is propositional.

We build a two-agent finite-horizon iterated NSBG with horizon $T = n$. Intuitively, round t fixes the value of x_t, and at the final round a terminal payoff evaluates ψ on the *entire* assignment $(x_0^0, \ldots, x_{n-1}^{n-1})$ chosen across the n rounds.

Variables and control. For each $t \in \{0, \ldots, n-1\}$, introduce a single controlled variable x_t^t. If $Q_t = \exists$ then $x_t^t \in V_1^t$ (agent 1 controls it at round t); if $Q_t = \forall$ then $x_t^t \in V_2^t$ (agent 2 controls it at round t). No other controlled variables exist. Moreover, let $V_{\text{env}}^t = \emptyset$ for all t. Thus each round is concurrent, but only one agent has a meaningful choice (that is, the other agent has an empty control set at that round).

Neural mechanisms. Let $E = \{e^\star\}$ and $\mathcal{D}(e^\star) = 1$. Define neural mechanisms deterministically so that realised controlled values match the chosen values: for each round t, each agent i, and each stage profile s^t,

$$\mathcal{N}_{i,t}^{\text{Out}}(s^t)(x_t^t) = (s^t(x_t^t), 1).$$

(Since $V_{\text{env}}^t = \emptyset$, $\mathcal{N}_{i,t}^{\text{In}}$ is irrelevant.)

Goals and payoffs. We enforce that *only the final round yields utility*, and that it encodes ψ under the assignment built across rounds. Let Ψ be the formula obtained from ψ by replacing each x_t with the time-stamped variable x_t^t.

Define stage goals:

$$\gamma_1^t := \bot, \quad \gamma_2^t := \bot \qquad \text{for all } t < n-1,$$
$$\gamma_1^{n-1} := \Psi, \quad \gamma_2^{n-1} := \neg\Psi.$$

Since neural valuations are deterministic, we have for any terminal history $h^n = (s^0, \ldots, s^{n-1})$ that $u_1^{n-1}(s^{n-1})$ is:

$$\Pr(\Psi) = \begin{cases} 1 & \text{if the induced assignment satisfies } \Psi, \\ 0 & \text{otherwise,} \end{cases}$$

while $u_2^{n-1}(s^{n-1}) = 1 - u_1^{n-1}(s^{n-1})$, and $u_i^t = 0$ for all $t < n-1$. Hence, total utilities under these definitions satisfy $U_1(h^n) = u_1^{n-1}$ and $U_2(h^n) = u_2^{n-1}$.

Correctness. Consider the standard QBF game semantics: players alternate choosing truth values for variables according to quantifiers, and $\exists$ wins if and only if the resulting assignment satisfies ψ. By construction, agent 1 chooses exactly the existential variables and agent 2 chooses exactly the universal variables (one per round), and the terminal payoff is 1 for agent 1 if and only if ψ is satisfied.

If Φ is true, then agent 1 has a winning strategy in the QBF game, hence a strategy in the iterated NSBG guaranteeing $U_1 = 1$ against all strategies of agent 2. Fix such a strategy for agent 1 and any strategy for agent 2; in every subgame, agent 1 cannot improve beyond 1 and agent 2 cannot force $U_2 = 1$, so the induced profile is a SPNE.

If Φ is false, then agent 2 has a winning strategy in the QBF game, hence can force $\neg\psi$, yielding $U_1 = 0$ and $U_2 = 1$. In that case, any purported SPNE would allow a profitable deviation for the losing player at some history (the standard backward-induction property of finite perfect-information games), contradicting subgame perfection. Therefore, a SPNE exists iff Φ is true.

Since QBF is reduced to a zero-sum sub-case of NSBGs in which winning strategies exist, no deviations at any stage are possible. Thus, QBF reduces to EXISTS-FH-SPNE, proving PSPACE-hardness of the iterated game. $\square$

6.2 On Expressiveness and Complexity

Finite-horizon games do not reduce to the stage game. Note that although utilities add across rounds, strategies may condition on the entire history and so optimality must hold in every subgame. Consequently, round-t choices must anticipate future strategic responses. The unfolding above makes this explicit: the terminal utility depends on the *sequence* of choices across rounds, so the problem cannot be reduced to solving the one-shot stage game and multiplying by T.

The membership/verification problem. We also note that the complexity of IS-FH-SPNE will depend on how strategy profiles are represented, whether explicitly or succinctly. We do not explore this problem further, but simply note that if using a succinct representation the problem may well be PSPACE-hard due to the possibility to encode QBF, similar to how it is done for EXISTS-FH-SPNE. But, the relevant complexity for this problem lies in how such strategies are ultimately implemented (in practice).

Infinite-horizon games and undecidability. We did not considered infinite-horizon games either. However, given several known results on probabilistic (stochastic) games of infinite duration, it is likely that most games of this kind are undecidable (Ummels and Wojtczak 2011), unless restrictions are imposed on the type of strategies under consideration, or the type of goals against which the games are defined – often requiring some kind of discounted utility function. Under most reasonable conditions, and surely in the general case, these games are expected to be undecidable.

A comprehensive complexity landscape. With the PSPACE result of this section, we obtain a comprehensive picture of where reasoning with learning components sits when coupled with symbolic reasoning, as we can see:

$$coNP \subseteq \Sigma_2^P \subseteq PH \subseteq PP \subseteq \Sigma_2^{PP} \subseteq CH \subseteq PSPACE.$$

7 Concluding Remarks and Related Work

Neuro-symbolic multi-agent systems. Neuro-symbolic approaches have been studied in areas such as knowledge representation, program synthesis, reasoning over learned representations, and hybrid learning–reasoning architectures (d'Avila Garcez et al. 2015; Raedt et al. 2016; Marra et al. 2024; Badreddine et al. 2022). However, most existing work focuses on *single-agent* settings. In contrast, many real-world systems are inherently distributed and involve multiple autonomous agents interacting strategically under uncertainty. Indeed, some strands of work have explored learning and reasoning in multi-agent systems, including probabilistic reasoning in multi-agent environments, learning in games, and neural approaches to multi-agent coordination (Shoham and Leyton-Brown 2009; Busoniu, Babuska, and Schutter 2008; Albrecht, Christianos, and Schäfer 2024).

As such, neuro-symbolic techniques have also been applied to multi-agent planning, reinforcement learning, and symbolic policy extraction, although typically without a unified formal game-theoretic semantics. From an implementation perspective, recent systems integrate probabilistic inference, logic programming, and neural learning (Raedt, Kimmig, and Toivonen 2007; Manhaeve et al. 2018; Hitzler and Sarker 2021), yet they generally lack a strategic foundation capturing equilibrium reasoning. Our work differs by providing a formal computational model in which neural perception, symbolic goals, and strategic behaviour coexist. This enables rigorous complexity analysis and formal reasoning about equilibria in systems that combine learning and symbolic reasoning, bridging a gap between neuro-symbolic distributed AI and algorithmic game theory.

Probabilistic reasoning and extended Boolean games. Probabilistic reasoning over propositional logic has been studied in AI and theoretical computer science, particularly in the context of Bayesian networks, probabilistic databases, and weighted model counting (Roth 1996; Sang, Beame, and Kautz 2005; Chavira and Darwiche 2008). Our results build on these foundations by showing that utility evaluation in neuro-symbolic Boolean games corresponds to weighted model counting under product distributions.

Boolean games (Harrenstein et al. 2001; Bonzon et al. 2006), on the other hand, have been extended to incorporate uncertainty, preferences, and possibility reasoning (van Ditmarsch and Simon 2024; Clercq et al. 2018; Popovici and Negreanu 2015). In particular, such Boolean games and variants have been studied to model uncertainty in outcomes or control without explicit probabilistic features, let alone neural or learning components included. Other works have explored epistemic and belief-based extensions, including reasoning under incomplete information (Ågotnes et al. 2013). Our model differs in that uncertainty arises from (probabilistic) neural valuation mechanisms rather than epistemic beliefs or imperfect information about what it is observable. Moreover, we study computational properties of the resulting neuro-symbolic games and show that reasoning about utilities already yields PP and $\#P$ complexity, while equilibrium problems reach Σ_2^{PP}. These results demonstrate that probabilistic reasoning and strategic interaction combine in a non-trivial way, generalising both Boolean games and probabilistic propositional reasoning.

Reasoning under uncertainty and formal models. Formal models for reasoning under uncertainty include probabilistic logic programming, Markov logic networks, and stochastic planning (Richardson and Domingos 2006; Raedt, Kimmig, and Toivonen 2007; Koller and Friedman 2009). These approaches provide tools for modelling uncertain knowledge and probabilistic inference, but typically do not incorporate strategic interaction among multiple decision-makers. Conversely, stochastic and Bayesian games capture strategic behaviour under uncertainty but lack explicit symbolic reasoning over logical goals. Our framework combines these perspectives by integrating probabilistic symbolic reasoning with strategic game-theoretic analysis. This enables formal study of how learned probabilistic beliefs influence equilibrium behaviour, providing a unified view that connects probabilistic reasoning, symbolic logic, and multi-agent strategic interaction.

Acknowledgements

I am grateful to Sebastian Amaya and Nox Cowie for their constructive feedback. I would also like to thank the KR 2026 reviewers of this paper for their valuable comments.

AI Declaration

The author has not employed any Generative AI tools.

References

Ågotnes, T.; Harrenstein, P.; van der Hoek, W.; and Wooldridge, M. J. 2013. Boolean games with epistemic goals. In Grossi, D.; Roy, O.; and Huang, H., eds., *Logic, Rationality, and Interaction - 4th International Workshop, LORI 2013, Hangzhou, China, October 9-12, 2013, Proceedings*, volume 8196 of *Lecture Notes in Computer Science*, 1–14. Springer.

Albrecht, S. V.; Christianos, F.; and Schäfer, L. 2024. *Multi-Agent Reinforcement Learning: Foundations and Modern Approaches*. MIT Press.

Arora, S., and Barak, B. 2009. *Computational Complexity: A Modern Approach*. Cambridge University Press.

Badreddine, S.; d'Avila Garcez, A. S.; Serafini, L.; and Spranger, M. 2022. Logic tensor networks. *Artif. Intell.* 303:103649.

Besold, T. R.; d'Avila Garcez, A. S.; Bader, S.; Bowman, H.; Domingos, P. M.; Hitzler, P.; Kühnberger, K.; Lamb, L. C.; Lima, P. M. V.; de Penning, L.; Pinkas, G.; Poon, H.; and Zaverucha, G. 2021. Neural-symbolic learning and reasoning: A survey and interpretation. In Hitzler, P., and Sarker, M. K., eds., *Neuro-Symbolic Artificial Intelligence: The State of the Art*, Frontiers in Artificial Intelligence and Applications. IOS Press. 1–51.

Bhuyan, B. P.; Ramdane-Cherif, A.; Tomar, R.; and Singh, T. P. 2024. Neuro-symbolic artificial intelligence: a survey. *Neural Comput. Appl.* 36(21):12809–12844.

Bonzon, E.; Lagasquie-Schiex, M.; Lang, J.; and Zanuttini, B. 2006. Boolean games revisited. In Brewka, G.; Coradeschi, S.; Perini, A.; and Traverso, P., eds., *ECAI 2006, 17th European Conference on Artificial Intelligence, August 29 - September 1, 2006, Riva del Garda, Italy, Including Prestigious Applications of Intelligent Systems (PAIS 2006), Proceedings*, Frontiers in Artificial Intelligence and Applications, 265–269. IOS Press.

Busoniu, L.; Babuska, R.; and Schutter, B. D. 2008. A comprehensive survey of multiagent reinforcement learning. *IEEE Trans. Syst. Man Cybern. Part C* 38(2):156–172.

Chavira, M., and Darwiche, A. 2008. On probabilistic inference by weighted model counting. *Artif. Intell.* 172(6-7):772–799.

Clercq, S. D.; Schockaert, S.; Nowé, A.; and Cock, M. D. 2018. Modelling incomplete information in boolean games using possibilistic logic. *Int. J. Approx. Reason.* 93:1–23.

d'Avila Garcez, A. S.; Besold, T. R.; Raedt, L. D.; Földiák, P.; Hitzler, P.; Icard, T.; Kühnberger, K.; Lamb, L. C.; Miikkulainen, R.; and Silver, D. L. 2015. Neural-symbolic learning and reasoning: Contributions and challenges. In

2015 AAAI Spring Symposia, Stanford University, Palo Alto, California, USA, March 22-25, 2015. AAAI Press.

Dong, H.; Mao, J.; Lin, T.; Wang, C.; Li, L.; and Zhou, D. 2019. Neural logic machines. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net.

Gill, J. 1977. Computational complexity of probabilistic turing machines. *SIAM Journal on Computing* 6(4):675–695.

Harrenstein, P.; van der Hoek, W.; Meyer, J.-J.; and Witteveen, C. 2001. Boolean games. In *TARK*.

Hitzler, P., and Sarker, M. K., eds. 2021. *Neuro-Symbolic Artificial Intelligence: The State of the Art*, volume 342 of *Frontiers in Artificial Intelligence and Applications*. IOS Press.

Koller, D., and Friedman, N. 2009. *Probabilistic Graphical Models - Principles and Techniques*. MIT Press.

Manhaeve, R.; Dumancic, S.; Kimmig, A.; Demeester, T.; and Raedt, L. D. 2018. Deepproblog: Neural probabilistic logic programming. In Bengio, S.; Wallach, H. M.; Larochelle, H.; Grauman, K.; Cesa-Bianchi, N.; and Garnett, R., eds., *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, December 3-8, 2018, Montréal, Canada*, 3753–3763.

Marra, G.; Dumancic, S.; Manhaeve, R.; and Raedt, L. D. 2024. From statistical relational to neurosymbolic artificial intelligence: A survey. *Artif. Intell.* 328:104062.

Popovici, M., and Negreanu, L. 2015. Hiding and revealing information in boolean games. In *20th International Conference on Control Systems and Computer Science, CSCS 2015, Bucharest, Romania, May 27-29, 2015*, 889–893. IEEE.

Raedt, L. D.; Kersting, K.; Natarajan, S.; and Poole, D. 2016. *Statistical Relational Artificial Intelligence: Logic, Probability, and Computation*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers.

Raedt, L. D.; Kimmig, A.; and Toivonen, H. 2007. Problog: A probabilistic prolog and its application in link discovery. In Veloso, M. M., ed., *IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007*, 2462–2467.

Richardson, M., and Domingos, P. M. 2006. Markov logic networks. *Mach. Learn.* 62(1-2):107–136.

Roth, D. 1996. On the hardness of approximate reasoning. *Artificial Intelligence* 82(1–2):273–302.

Sang, T.; Beame, P.; and Kautz, H. A. 2005. Performing bayesian inference by weighted model counting. In Veloso, M. M., and Kambhampati, S., eds., *Proceedings, The Twentieth National Conference on Artificial Intelligence and the Seventeenth Innovative Applications of Artificial Intelligence Conference, July 9-13, 2005, Pittsburgh, Pennsylvania, USA*, 475–482. AAAI Press / The MIT Press.

Shoham, Y., and Leyton-Brown, K. 2009. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press.

Toda, S. 1991. PP is as hard as the polynomial-time hierarchy. *SIAM Journal on Computing* 20(5):865–877.

Ummels, M., and Wojtczak, D. 2011. The complexity of Nash equilibria in stochastic multiplayer games. *Log. Methods Comput. Sci.* 7(3).

Valiant, L. G. 1979. The complexity of computing the permanent. *Theoretical Computer Science* 8:189–201.

van Ditmarsch, H., and Simon, S. 2024. Boolean observation games. *J. Artif. Intell. Res.* 79:307–357.