

# Formally Reasoning about Strategies in Mechanisms

Munyuque Mittelmann<sup>1</sup>

<sup>1</sup>Università degli Studi di Napoli Federico II, Italy

## Abstract

In this paper, we present recently published work that proposes a new approach for Mechanism Design, the problem of designing games for aggregating preferences into social decisions. This approach is based on formal methods and logics for strategic reasoning in Multi-Agent Systems.

## Keywords

Strategic Reasoning, Formal Methods, Multi-Agent Systems, Mechanism Design

## 1. Introduction

When considering settings with multiple agents, two main possible situations arise. The first is when all agents can be assumed to share a common purpose. For instance, teams of cooperating agents have been considered in a wide range of applications including operations for search and rescue, surveillance, and space exploration [1]. However, a second scenario happens when some (or all) agents in the environment have their own, possibly conflicting, goals and preferences [2]. This is the case, for instance, when considering voting systems: agents may not share the same preference over the candidates. Moreover, some agents may even try to manipulate the results of an election through bribing and coercion to convince others to change their votes.

Much progress has been done in the development of logics for strategic reasoning in Multi-Agent Systems, including the formulation of Alternating-Time Temporal Logic (ATL) [3], Strategy Logic (SL) [4, 5] and its extensions to imperfect information [6, 7, 8], and quantitative semantics (denoted  $SL[\mathcal{F}]$ ) [9]. On the other side, research in economics has been concerned with the design of games (aka *mechanisms*) for aggregating individual preferences, while motivating agents to behave well (e.g., being truthful) and ensuring the quality of social decisions (e.g., allocative efficiency) [10]. The real-world applications of designing and verifying mechanisms for social choice are manifold, including the classic problems on auctions, markets, and government policies. Furthermore, new and trending applications include fair division protocols [11], secure voting [12], and even the formulation of human rights for economical development [13]. The problem of formally reasoning about mechanisms is, however, nontrivial: it requires considering quantitative information (e.g., utilities and payments), non-determinism, incomplete (e.g., Bayesian) and imperfect information about the participant's preferences, and complex strategic concepts (such as strategy dominance and equilibria).

---


OVERLAY 2022: 4th Workshop on Artificial Intelligence and Formal Verification, Logic, Automata, and Synthesis, December 28th, 2022, Udine, Italy

✉ [munyuque.mittelmann@ut-capitole.fr](mailto:munyuque.mittelmann@ut-capitole.fr) (M. Mittelmann)

🌐 <https://sites.google.com/view/mittelmann> (M. Mittelmann)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

In this paper, we present recently published work [14, 15, 16] that proposes a new approach for Mechanism Design (MD) based on the logics developed for strategic reasoning in MAS. By enhancing the connection between different research communities, we aim to facilitate the adoption of techniques from each field (for instance, the synthesis of game arenas from the algorithms for solving satisfiability). The logical formalization of MD may provide an inherently different view of social choice problems from the classical characterizations.

We start by presenting the base logic on which this approach of mechanism design is built, namely the Quantitative Strategy Logic,  $SL[\mathcal{F}]$ . Next, we summarize some recent results in extending and applying such logic for reasoning about mechanisms. Finally, we conclude with a short discussion and directions for future work.

## 2. Quantitative Strategy Logic

Let us first present  $SL[\mathcal{F}]$  [9] syntax and semantics.

**Definition 1.** The syntax of  $SL[\mathcal{F}]$  is defined by the grammar

$$\varphi ::= p \mid \exists s. \varphi \mid (i, s)\varphi \mid f(\varphi, \dots, \varphi) \mid \mathbf{X}\varphi \mid \varphi \mathbf{U}\varphi$$

where  $p \in AP$  is an atomic proposition,  $s \in \text{Var}$  is a strategy variable,  $i \in \mathbb{N}$  is an agent, and  $f \in \mathcal{F}$  is a function in  $[-1, 1]$ .

The intuitive reading is:  $\exists s. \varphi$  means that there exists a strategy such that  $\varphi$  holds;  $(i, s)\varphi$  means that when strategy  $s$  is assigned to agent  $i$ ,  $\varphi$  holds;  $\mathbf{X}$  and  $\mathbf{U}$  are the usual temporal operators “next” and “until”. The meaning of  $f(\varphi_1, \dots, \varphi_n)$  depends on the function  $f$ . We use  $\top$ ,  $\vee$ , and  $\neg$  to denote, resp., function 1, function  $x, y \mapsto \max(x, y)$  and function  $x \mapsto -x$ .

**Definition 2.** A *weighted concurrent game structure* (wCGS) is a tuple  $\mathcal{G} = (\mathcal{B}, V, v_\iota, \delta, \ell)$  where (i)  $\mathcal{B}$  is a finite set of *actions*; (ii)  $V$  is a finite set of *positions*; (iii)  $v_\iota \in V$  is an *initial position*; (iv)  $\delta : V \times \mathcal{B}^{\mathbb{N}} \rightarrow V$  is a *transition function*; (v)  $\ell : V \times AP \rightarrow [-1, 1]$  is a *weight function* assigning a value in  $[-1, 1]$  to each atomic proposition in each position.

In a position  $v \in V$ , each player  $i$  chooses an action  $a_i \in \mathcal{B}$ , and the game proceeds to position  $\delta(v, \mathbf{a})$  where  $\mathbf{a} = (a_i)_{i \in \mathbb{N}}$  is the *action profile* (i.e., a tuple with an action for each agent). We write  $\mathbf{o}$  for a tuple of objects  $(o_i)_{i \in \mathbb{N}}$ , one for each agent, and such tuples are called *profiles*. Given a profile  $\mathbf{o}$  and  $i \in \mathbb{N}$ , we let  $o_i$  be agent  $i$ 's component, and  $\mathbf{o}_{-i}$  is  $(o_r)_{r \neq i}$ . Similarly, we let  $\mathbb{N}_{-i} = \mathbb{N} \setminus \{i\}$ . A *play*  $\pi = v_1 v_2 \dots$  is an infinite sequence of positions such that for every  $i \geq 1$  there exists an action profile  $\mathbf{a}$  such that  $\delta(v_i, \mathbf{a}) = v_{i+1}$ . A *history*  $h$  is a finite prefix of a play. A *strategy* is a function  $\sigma : \text{Hist} \rightarrow \mathcal{B}$  that maps each history to an action. We let  $\text{Str}$  be the set of strategies. An *assignment*  $\mathcal{A} : \mathbb{N} \cup \text{Var} \rightarrow \text{Str}$  is a function from players and variables to strategies. For an assignment  $\mathcal{A}$ , an agent  $i$  and a strategy  $\sigma$  for  $i$ ,  $\mathcal{A}[a \mapsto \sigma]$  is the assignment that maps  $a$  to  $\sigma$  and is otherwise equal to  $\mathcal{A}$ , and  $\mathcal{A}[s \mapsto \sigma]$  is defined similarly, where  $s$  is a variable.

**Definition 3.** (Partial, see complete def. in [9]) Let  $\mathcal{G}$  be a wCGS, and  $\mathcal{A}$  an assignment. The satisfaction value  $\llbracket \varphi \rrbracket_{\mathcal{A}}^{\mathcal{G}}(h) \in [-1, 1]$  of an  $SL[\mathcal{F}]$  formula  $\varphi$  in a history  $h$  is defined as follows:

$$\llbracket p \rrbracket_{\mathcal{A}}^{\mathcal{G}}(h) = \ell(\text{last}(h), p)$$

$$\begin{aligned} \llbracket \exists s. \varphi \rrbracket_{\mathcal{A}}^{\mathcal{G}}(h) &= \max_{\sigma \in \text{Str}} \llbracket \varphi \rrbracket_{\mathcal{A}[s \rightarrow \sigma]}^{\mathcal{G}}(h) \\ \llbracket (i, s) \varphi \rrbracket_{\mathcal{A}}^{\mathcal{G}}(h) &= \llbracket \varphi \rrbracket_{\mathcal{A}[i \rightarrow \mathcal{A}(s)]}^{\mathcal{G}}(h) \end{aligned}$$

### 3. Reasoning about Mechanisms

We now briefly describe recent work at extending and applying  $\text{SL}[\mathcal{F}]$  to MD.

#### 3.1. Verification of Mechanisms under Imperfect Information

In [14] we propose an epistemic extension of  $\text{SL}[\mathcal{F}]$ , denoted  $\text{SLK}[\mathcal{F}]$ , for considering imperfect information about agents' preferences (or *types*). We show how  $\text{SLK}[\mathcal{F}]$  can express solution concepts (such as Nash equilibria) and fundamental properties from mechanisms, including budget-balance, individual rationality, strategyproofness and Pareto optimality. We prove that our encoding of such properties is correct w.r.t. the classical definitions from economics.

For instance, a strategy profile is a *Nash equilibrium* (NE) if no agent can increase her utility with a unilateral change of strategy [17]. Which is expressed with the  $\text{SLK}[\mathcal{F}]$ -formula:

$$\text{NE}(\mathbf{s}) =_{\text{def}} \bigwedge_{i \in \mathbb{N}} \forall t. [(\mathbf{N}_{-i}, \mathbf{s}_{-i})(i, t) \mathbf{F} \text{util}_i \leq (\mathbf{N}, \mathbf{s}) \mathbf{F} \text{util}_i]$$

where  $\mathbf{s} = (s_i)_{i \in \mathbb{N}}$  is a profile of strategy variables (for details, see [14]).

Similarly, individual rationality (IR) is a property of mechanisms in which each agent can ensure to always get nonnegative utility [18], which is encoded with our approach by evaluating the formula  $\text{IR} =_{\text{def}} \bigwedge_{i \in \mathbb{N}} 0 \leq \text{util}_i$  in strategic equilibrium.

We then consider epistemic aspects and show how, thanks to the epistemic operators in  $\text{SLK}[\mathcal{F}]$ , we can express properties relating agents' revenues with their knowledge about other agents' preferences. Verifying that a mechanism satisfies a property then consists in model checking an  $\text{SLK}[\mathcal{F}]$  formula, which we show can be done in  $\text{PSPACE}$  for memoryless strategies.

#### 3.2. Mechanisms with Natural Strategies

A Natural Strategy [19, 20] is a list of condition-action pairs, denoting the action the agent will take in case the condition is met. In [15] we have extended  $\text{SL}[\mathcal{F}]$  to handle this type of strategies. The resulting language, denoted  $\text{NatSL}[\mathcal{F}]$ , provides a tool for reasoning about mechanisms in relation to strategies with bounded complexity. We demonstrated the usefulness of our approach by modeling and evaluating strategies for repeated keyword auctions. In terms of technical results, we proved that the model checking problem for  $\text{NatSL}[\mathcal{F}]$  is  $\text{PSPACE}$ -complete, that is, no harder than model checking for the much less expressive language of quantitative LTL ( $\text{LTL}[\mathcal{F}]$ ). We also showed that  $\text{NatSL}[\mathcal{F}]$  has incomparable distinguishing and expressive power to  $\text{SL}[\mathcal{F}]$ . This means that the characterizations based on simple bounded strategies offer an inherently different view of mechanisms from characterizations using combinatorial strategies of arbitrary complexity.

### 3.3. Mechanism Synthesis

In [16], we rephrased the problem of designing new mechanisms in terms of synthesis from  $SL[\mathcal{F}]$ -specifications. Such specifications may include not only game rules but also requirements over the strategic behavior of participants and quality of the outcome. While mechanism synthesis from  $SL[\mathcal{F}]$  formulas is undecidable, we solve it in two cases: when the number of actions is bounded, and when agents play in turn. These two restrictions still preserve enough expressiveness to model relevant scenarios of mechanism synthesis, which we illustrate with examples based on auctions. The quantitative semantics of  $SL[\mathcal{F}]$  makes it possible to synthesize mechanisms that approximate a specification (satisfy it as much as possible), or maximize some value (such as social welfare in an equilibrium), which is not possible with SL.

## 4. Discussion

In this paper, we discuss a novel approach for Automated Mechanism Design in which mechanisms can be automatically (i) verified with respect to epistemic and quantitative properties (ii) verified with respect to natural strategies, and (iii) generated (or *synthesized*) from partial or complete specifications in a rich logical language. An advantage of our approach lies in the high expressivity and generality of logics for strategic reasoning. Moreover, by relying on precise semantics, formal methods provide tools for rigorously analyzing the correctness of systems, which is important to improve trust in mechanisms (fully or partially) created by machines. The ability of SL to naturally express key strategic concepts such as Nash Equilibria, and the possibility to extend it with quantitative aspects and epistemic operators, as we do with  $SLK[\mathcal{F}]$ , make it a perfect candidate to become a standard logic for verifying mechanisms.

The present setting is enough to capture many kinds of mechanisms where memoryless strategies are sufficient to represent the agents' behavior, such as one-shot or English auctions. However, when participating in sequential auctions, agents could gather information from other agents' behavior and act based on what happened in previous steps of the game [21]. For such situations we plan to study further the model-checking problem for  $SLK[\mathcal{F}]$  with perfect recall. In the qualitative setting already, imperfect information yields undecidability, but known decidable cases exist [7, 6]. We will investigate them in the quantitative case. Yet another interesting direction for future work is to study the complexity of synthesizing from  $SL[\mathcal{F}]$ -fragments, inspired from the SL-fragments One-Goal SL [22, 23] and Simple-Goal SL [24], for instance. These fragments are usually computationally easier than full SL, and we can hope that similar results can be established in the quantitative setting.

## Acknowledgments

I am very grateful to my co-authors Laurent Perrussel, Aniello Murano, Bastien Maubert, Francesco Belardinelli, Wojtek Jamroga, Vadim Malvone. This research is supported by the ANR project AGAPE ANR-18-CE23-0013 and the PRIN project RIPER (No. 20203FFYLK).

## References

- [1] Y. Wang, E. Garcia, D. Casbeer, F. Zhang (Eds.), *Cooperative Control of Multi-Agent Systems*, Wiley, 2017. URL: <https://doi.org/10.1002/9781119266235>. doi:10.1002/9781119266235.
- [2] W. van der Hoek, M. Wooldridge, Chapter 24 multi-agent systems, in: *Handbook of Knowledge Representation*, Elsevier, 2008, pp. 887–928. URL: [https://doi.org/10.1016/s1574-6526\(07\)03024-6](https://doi.org/10.1016/s1574-6526(07)03024-6). doi:10.1016/s1574-6526(07)03024-6.
- [3] R. Alur, T. A. Henzinger, O. Kupferman, Alternating-time temporal logic, *Journal of the ACM* 49 (2002) 672–713.
- [4] K. Chatterjee, T. A. Henzinger, N. Piterman, Strategy logic, *Information and Computation* 208 (2010) 677–693.
- [5] F. Mogavero, A. Murano, G. Perelli, M. Y. Vardi, Reasoning about strategies: On the model-checking problem, *ACM Transactions on Computational Logic (TOCL)* 15 (2014) 1–47.
- [6] F. Belardinelli, A. Lomuscio, A. Murano, S. Rubin, Verification of multi-agent systems with public actions against strategy logic, *Artificial Intelligence* 285 (2020).
- [7] R. Berthon, B. Maubert, A. Murano, S. Rubin, M. Vardi, Strategy logic with imperfect information, *ACM Transactions on Computational Logic* 22 (2021).
- [8] P. Cermák, A. Lomuscio, F. Mogavero, A. Murano, Practical verification of multi-agent systems against SLK specifications, *Information and Computation* 261 (2018) 588–614.
- [9] P. Bouyer, O. Kupferman, N. Markey, B. Maubert, A. Murano, G. Perelli, Reasoning about quality and fuzziness of strategic behaviours, in: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI 2019)*, 2019.
- [10] T. Sandholm, Automated mechanism design: A new application area for search algorithms, in: *Principles and Practice of Constraint Programming – CP 2003*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2003, pp. 19–36.
- [11] J. Lang, J. Rothe, Fair division of indivisible goods, in: *Economics and Computation*, Springer, 2016, pp. 493–550.
- [12] W. Jamroga, M. Tabatabaei, Preventing coercion in e-voting: Be open and commit, in: *Proc. Elect. Voting-16*, 2016.
- [13] D. Mohan, Social choice theory and its application in a human rights based approach to development, *Journal of governance & regulation* (2017) 7–13.
- [14] B. Maubert, M. Mittelmann, A. Murano, L. Perrussel, Strategic reasoning in automated mechanism design, in: *Proc. of the Int. Conference on Principles of Knowledge Representation and Reasoning (KR 2021)*, 2021, pp. 487–496.
- [15] F. Belardinelli, W. Jamroga, V. Malvone, M. Mittelmann, A. Murano, L. Perrussel, Reasoning about human-friendly strategies in repeated keyword auctions, in: *Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2022)*, ACM, 2022, pp. 1602–1604.
- [16] M. Mittelmann, B. Maubert, A. Murano, L. Perrussel, Automated synthesis of mechanisms, in: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI 2022)*, 2022.
- [17] D. C. Parkes, L. H. Ungar, Iterative combinatorial auctions: Achieving economic and

- computational efficiency, University of Pennsylvania Philadelphia, PA, 2001.
- [18] N. Nisan, T. Roughgarden, É. Tardos, V. Vazirani, *Algorithmic Game Theory*, Cambridge University Press, 2007.
  - [19] W. Jamroga, V. Malvone, A. Murano, Natural strategic ability, *Artificial Intelligence* 277 (2019) 103170.
  - [20] W. Jamroga, V. Malvone, A. Murano, Natural strategic ability under imperfect information, in: *Proceedings of the International Conference on Autonomous Agents and MultiAgent Systems (AAMAS 2019)*, International Foundation for Autonomous Agents and Multiagent Systems, Richland, 2019, p. 962–970.
  - [21] T. D. Jeitschko, Learning in sequential auctions, *Southern Economic Journal* (1998) 98–112.
  - [22] F. Mogavero, A. Murano, G. Perelli, M. Y. Vardi, Reasoning about strategies: on the satisfiability problem, *Logical Methods in Computer Science* 13 (2017). URL: [https://doi.org/10.23638/LMCS-13\(1:9\)2017](https://doi.org/10.23638/LMCS-13(1:9)2017). doi:10.23638/LMCS-13(1:9)2017.
  - [23] P. Cermák, A. Lomuscio, A. Murano, Verifying and synthesising multi-agent systems against one-goal strategy logic specifications, in: *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2015)*, 2015.
  - [24] F. Belardinelli, W. Jamroga, D. Kurpiewski, V. Malvone, A. Murano, Strategy logic with simple goals: Tractable reasoning about strategies, in: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI 2019)*, 2019.