STRATEGIC REASONING IN AUTOMATED MECHANISM DESIGN

We propose a quantitative and epistemic variant of **Strategy Logic** for enabling **automated verification of mechanisms**. We first show how to express the implementation of social choice functions. Second, we show how fundamental mechanism properties can be expressed as logical formulas, and thus evaluated by model checking. Finally, we prove that model checking can be done in polynomial space.

Authors: Bastien Maubert [1], Munyque Mittelmann [2], Aniello Murano [1], Laurent Perrussel [2]

Affiliations: [1] Università degli Studi di Napoli "Federico II"
[2] Université Toulouse 1 Capitole

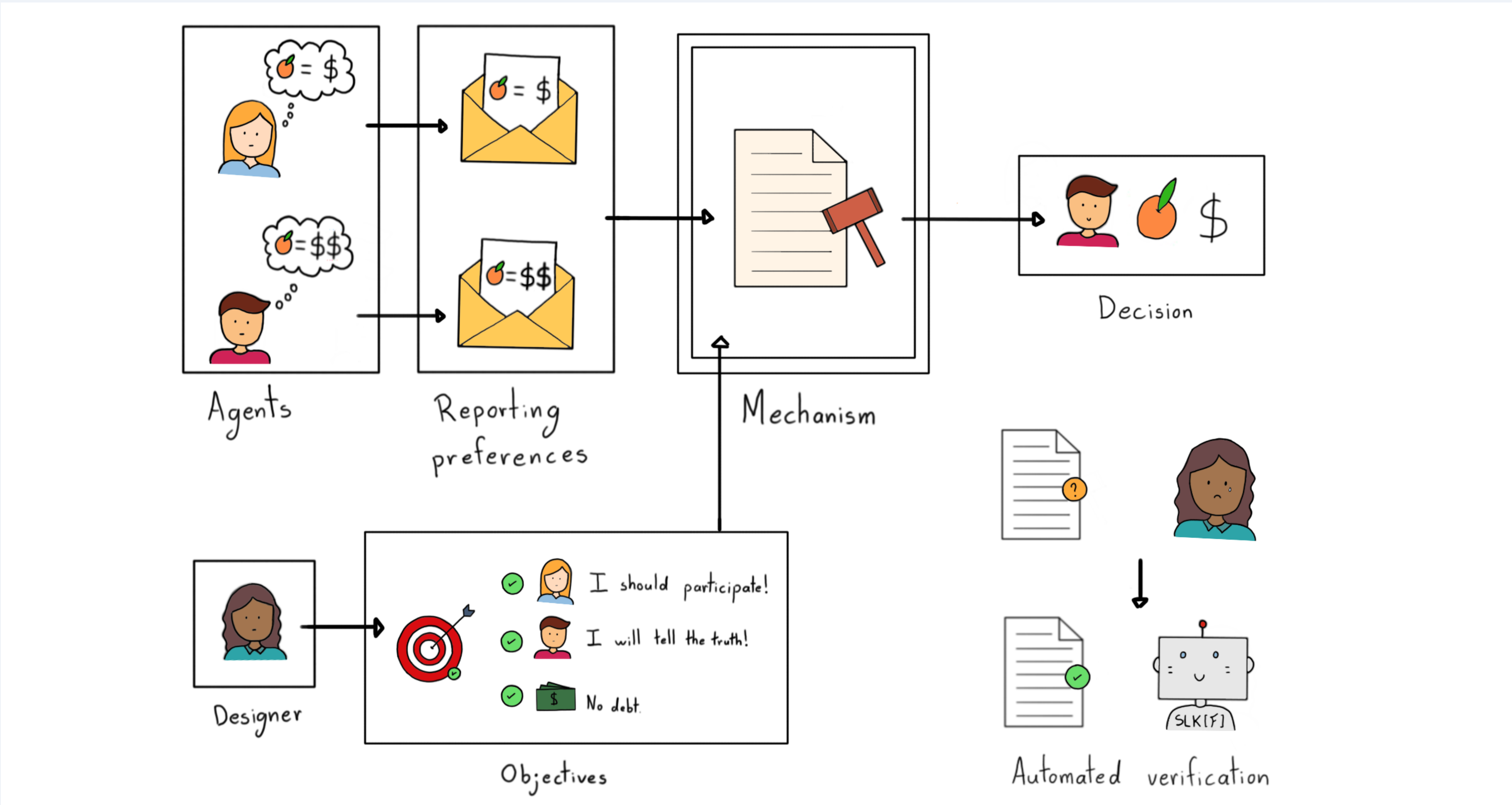
Contact: munyque.mittelmann@irit.fr

INTRODUCTION

A social choice function (SCF) is a mapping from agents' reported preferences to decisions over alternatives. A mechanism generalizes an SCF and maps strategies to alternatives. The mechanism designer aims at defining its rules in a way to incentivize a desirable behavior of the participants and ensure features of the outcome.

We investigate a logic-based approach to **formal verification** of auction mechanisms in relation to the designer's objectives.

We propose a **quantitative and epistemic version of Strategy Logic** (SLK[$\mathscr{F}$]) and show how it is well fitted to capture the core properties of mechanisms.



BASICS

SLK[$\mathscr{F}$] is defined over Weighted Concurrent Game Structures (wCGS). A wCGS includes a **weight** function assigning a value in [-1, 1] to each proposition in a state.
Agents have imperfect information.
Strategies map states to actions and are consistent with the agents' observations.
Assignments map agents and variables to strategies.

SLK[$\mathscr{F}$] SYNTAX

$\varphi ::= p \mid f(\varphi, \dots, \varphi) \mid \exists s_a.\varphi \mid (a, s_a)\varphi \mid K_a\varphi \mid X\,\varphi \mid \varphi\,U\,\varphi$

where p is an atomic proposition, f is a function in $\mathscr{F}$, a is an agent and s_a is a strategy for agent a.

INTUITIVE READING

- $\exists s_a.\varphi$ means that there exists a strategy for agent a such that φ holds;
- $(a,s_a)\varphi$ means that when strategy s_a is assigned to agent a, φ holds;
- $K_a\varphi$ means "agent a knows that φ holds";
- X and U are the usual temporal operators "next" and "until";
- The meaning of $f(\varphi, \dots, \varphi)$ depends on the function f.

We use $\top, \vee$ and $\neg$ to denote, respectively, function 1, function $x,y \mapsto \max(x,y)$ and function $x \mapsto -x$. We also use the abreviations: $\varphi \wedge \varphi' := \neg(\neg\varphi \vee \neg\varphi')$, and $F\varphi := \top U \varphi$.

In the case where atomic propositions only take values in {−1, 1} and $\mathscr{F}$ consists of functions $\vee$ (disjunction) and $\neg$ (negation), SLK[$\mathscr{F}$] corresponds to Boolean-valued SLK.

SLK[$\mathscr{F}$] SEMANTICS

Given a wCGS and an assignment χ, the **satisfaction** of a formula φ in state v is a value in [−1, 1]:

- For *p*, its weight in v;
- For $\exists s_a.\varphi$, the maximal satisfaction value of φ in v over all possible strategies;
- For (a, s_a), the satisfaction value of φ in v when agent a is mapped to s_a;
- For $X\varphi$, the value of φ in the next state of v when following the strategies assigned by χ.

(...)

MODEL CHECKING

Assuming that functions in $\mathscr{F}$ can be computed in polynomial space, model checking SLK[$\mathscr{F}$] with memoryless agents is PSPACE-complete.

SOLUTION CONCEPTS IN SLK[$\mathscr{F}$]

A strategy profile s is a **Nash equilibrium** (NE) if no agent can improve her utility by a unilateral change of strategy:

$NE(s) := \bigwedge_{a\in Ag} \forall t.\, [\,(Ag_{-a}, s_{-a})(a, t)\, F\,(utility_a) \leqslant (Ag, s)\, F\,(utility_a)]$

A **dominant strategy equilibrium** (DSE) weakly maximizes each agent's utility, for all strategies of other agents:

$DSE(s) := \bigwedge_{a\in Ag} DS(s_a, a)$, where
$DS(s_a, a) := \forall t.[(a, t_a)(Ag_{-a}, t_{-a})\, F\,(utility_a) \leqslant (a, s_a)(Ag_{-a}, t_{-a})\, F\,(utility_a)]$

MECHANISM PROPERTIES

Individual rationality: agents can achieve at least zero utility: $IR := \bigwedge_{a\in Ag} 0 \leqslant utility_a$.

Strategyproofness: each agent a has an incentive to report their real preference θ_a: $SP := \bigwedge_{a\in Ag} DS(\theta_a, a)$.

We also consider budget balance, Pareto optimality, efficiency, and belief-based revenue benchmark.

A wCGS **E-implements** a SCF if they choose the same alternative (or *outcome*) in some E-equilibrium, for E ∈ {NE, DSE}.

MECHANISM VERIFICATION

We express a number of properties for mechanisms and prove that an SCF *f* has a property iff the corresponding formula holds at the E-equilibrium implementing *f*. Thus, verifying a mechanism **boils down to model checking** SLK[$\mathscr{F}$]-formulae.

FUTURE WORK

- Strategies with recall.
- Synthesis of mechanism specifications.
- Probabilistic SL for mechanism design.