Credulous Acceptability, Poison Games & Modal Logic

Davide Grossi[1] and Simon Rey[2]

[1]Bernoulli Institute, University of Groningen, Groningen, The Netherlands
[2]Sorbonne Université, ENS Paris Saclay, Paris, France

The Poison Game

Rules A directed graph $\langle W, R\rangle$ is given. $\mathbb{P}$ select the first node $w_0 \in W$. After this initial choice, $\mathbb{O}$ selects w_1 a successor of w_0, $\mathbb{P}$ then selects a successor of w_1 and so on. $\mathbb{O}$ can choose any successor of the current node, but $\mathbb{P}$ can only select successors which have not yet been visited (*poisoned*) by $\mathbb{O}$.

Winning conditions $\mathbb{O}$ wins if and only if $\mathbb{P}$ ends up in a position with no moves available.

Semi-kernels

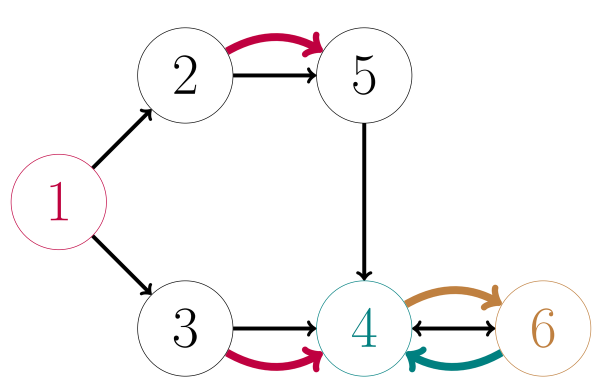
A semi-kernel in a directed graph $\langle W, R\rangle$ is set $X \subseteq W$ s.t.:

- $\nexists x, y \in X$ s.t. xRy
- $\forall x \in X$ if xRy then $\exists z \in X$ s.t. yRz.

If $\langle W, R\rangle$ is an attack graph, then X is an **admissible set** of the attack graph iff it is a semi-kernel of the reverse graph $\langle W, R^{-1}\rangle$.

Theorem [3] Let $\langle W, R\rangle$ be a finite directed graph. There exists a non-empty semi-kernel in $\langle W, R\rangle$ if and only if $\mathbb{P}$ has a winning strategy in the Poison Game for $\langle W, R\rangle$.

Example



In the graph $\{4\}$ and $\{6\}$ are two semi-kernels. There are three winning strategy for $\mathbb{P}$:

- choose 4 indefinitely: →
- choose 6 indefinitely: →
- choose 1 and then depending on the reply by $\mathbb{O}$ resort to one of the two earlier strategies: → .

More Information

Full paper at `https://arxiv.org/abs/1901.09180`.

Syntax and Semantics

The language $\mathcal{L}^{\mathfrak{p}}$ is defined by the following grammar in BNF:

$$\mathcal{L}^{\mathfrak{p}} : \varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \Diamond\varphi \mid \blacklozenge\varphi,$$

where $p \in \mathbf{P} \cup \{\mathfrak{p}\}$ with $\mathbf{P}$ a countable set of propositional atoms and $\mathfrak{p}$ a distinguished atom called *poison atom*.

This language is interpreted on Kripke models $\mathcal{M} = (W, R, V)$. A pointed model is a pair $(\mathcal{M}, w)$ with $w \in \mathcal{M}$.
The semantics for the $\blacklozenge$ modality is as follows (others are standard):

$$(\mathcal{M}, w) \models \blacklozenge\varphi \Leftrightarrow \exists v \in W, wRv, (\mathcal{M}_v^{\bullet}, v) \models \varphi.$$

where

$$\mathcal{M}_w^{\bullet} = (W, R, V)_w^{\bullet} = (W, R, V'),$$

with $V'(p) = V(p), \forall p \in \mathbf{P}$ and $V'(\mathfrak{p}) = V(\mathfrak{p}) \cup \{w\}$.

Story

1. The **Poison Game** is a game on graphs that characterizes the existence of **semi-kernels** in a directed graph.
2. Semi-kernels correspond to **admissible sets** of arguments in attack graphs as defined in abstract argumentation;
3. The existence of admissible sets (a.k.a., existence of credulously acceptable arguments) is thus characterized by the poison game.
4. Several argumentation theory notions (a.k.a. semantics) are known to be expressed by specific modal formulae [4], but **no modal formula is known to characterize the existence of credulously admissible arguments**.
5. Poison Modal Logic (PML) extends classical modal logic with one modality capturing the notion of 'poisoning' of a node in a graph.
6. The paper shows PML has the right expressivity to **characterize the existence of winning strategies** in the Poison Game (in bounded games), and therefore credulous admissibility.
7. We study the model theory and decidability of PML.

Expressivity

Let δ_n, with $n \in \mathbb{N}_{>0}$, be defined as: $\delta_1 = \Diamond\mathfrak{p}$; $\delta_{i+1} = \Diamond(\neg\mathfrak{p} \wedge \delta_i)$.

Fact Let $(\mathcal{M}, w) \in \mathfrak{M}^{\emptyset}$ with $\mathcal{M} = (W, R, V)$, then for $n \in \mathbb{N}_{>0}$ there exists $w \in W$ such that $(\mathcal{M}, w) \models \blacklozenge\delta_n$ if and only if there is a cycle of length $i \leq n$ in the frame (W, R). So PML is not bisimulation invariant nor enjoys the tree model property.

PML is a proper fragment of the memory logic known as $\mathcal{M}(\textcircled{r}, \textcircled{k})$ [1]. PML can be embedded into $\mathcal{H}(\downarrow)$.

Standard translation

P is the first order predicate for $p \in \mathbf{P}$ and $\mathfrak{P}$ is the one for $\mathfrak{p}$.

$$ST_x^N(p) = P(x), \forall p \in \mathbf{P}$$
$$ST_x^N(\neg\varphi) = \neg ST_x^N(\varphi)$$
$$ST_x^N(\varphi \wedge \psi) = ST_x^N(\varphi) \wedge ST_x^N(\psi)$$
$$ST_x^N(\Diamond\varphi) = \exists y \left(R(x, y) \wedge ST_y^N(\varphi)\right)$$
$$ST_x^N(\blacklozenge\varphi) = \exists y \left(R(x, y) \wedge ST_y^{N \cup \{y\}}(\varphi)\right)$$
$$ST_x^N(\mathfrak{p}) = \mathfrak{P}(x) \vee \bigvee_{y \in N} (y = x).$$

Main Theorems

- For $(\mathcal{M}, w) \in \mathfrak{M}$ and $\varphi \in \mathcal{L}^{\mathfrak{p}}$ a formula, we have: $(\mathcal{M}, w) \models \varphi$ iff $\mathcal{M} \models ST_x^{\emptyset}(\varphi)[x := w]$.
- For any two pointed models $(\mathcal{M}_1, w_1)$ and $(\mathcal{M}_2, w_2)$, if $(\mathcal{M}_1, w_1) \rightleftharpoons (\mathcal{M}_2, w_2)$ then $(\mathcal{M}_1, w_1) \rightsquigarrow (\mathcal{M}_2, w_2)$.
- For any two ω-saturated models $(\mathcal{M}_1, w_1)$ and $(\mathcal{M}_2, w_2)$, if $(\mathcal{M}_1, w_1) \rightsquigarrow (\mathcal{M}_2, w_2)$ then $(\mathcal{M}_1, w_1) \rightleftharpoons (\mathcal{M}_2, w_2)$.
- A FOL formula is equivalent to the translation of a $\mathcal{L}^{\mathfrak{p}}$ formula if and only if it is p-bisimulation invariant.
- The satisfiability problem for PML_3 is undecidable.
- PML does not have the Finite Model Property.

Where: $\rightsquigarrow$ is the modal equivalence and $\rightleftharpoons$ the poison bisimulation relation (p-bisimulation).

Winning positions in PML

Winning positions for $\mathbb{O}$ are defined by the following infinitary formula:

$$\mathfrak{p} \vee \blacklozenge\Box\mathfrak{p} \vee \blacklozenge\Box\blacklozenge\Box\mathfrak{p} \vee \cdots$$

Dually, winning positions for $\mathbb{P}$ are:

$$\neg\mathfrak{p} \wedge \Diamond\blacksquare\neg\mathfrak{p} \wedge \Diamond\blacksquare\Diamond\blacksquare\neg\mathfrak{p} \wedge \cdots.$$

Open questions

- Sabotage modal logic [2]
- Proof system for PML
- Fixed-point PML: μPML

References

[1] C. Areces, D. Figueira, S. Figueira, and S. Mera. The expressive power of memory logics. *Review of Symbolic Logic*, 4(2):290–218, 2011.

[2] G. Aucher, J. van Benthem, and D. Grossi. Modal logics of sabotage revisited. *Journal of Logic and Computation*, 28(2):269–303, 2017.

[3] P. Duchet and H. Meyniel. Kernels in directed graphs: a poison game. *Discrete mathematics*, 115(1-3):273–276, 1993.

[4] D. Grossi. On the logic of argumentation theory. In W. van der Hoek, G. Kaminka, Y. Lespérance, and S. Sen, editors, *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, pages 409–416. IFAAMAS, 2010.