Stackelberg-Pareto Synthesis and Verification

Véronique Bruyère
University of Mons - UMONS
Belgium

Workshop “Current Trends in Graph and Stochastic Games”
Maastricht - April 2022

Joint work with Jean-François Raskin and Clément Tamines
1 Reactive synthesis

2 Stackelberg non zero-sum games

3 Stackelberg-Pareto verification

4 Stackelberg-Pareto synthesis

5 Rational synthesis/verification
Reactive synthesis

Reactive systems

- **System** which constantly interacts with an uncontrollable **environment**
- It must satisfy some **property** against any behavior of the environment
- How to automatically design a correct **controller** for the system?

Modelization

- **Two-player zero-sum game** played on a finite directed **graph**
- **Property** = **objective** for the system
- **Synthesis** of a controller = construction of a winning strategy
Reactive synthesis

**Classical approach** with numerous results and several tools, see e.g.
- The book chapter “Graph Games and Reactive Synthesis” [BCJ18]

**Disadvantages**

Fully adversarial environment: **bold abstraction of reality**
- Assumes the only goal of the environment is to make the system fail
- Environment can be composed of one or several components, each with its own objective
More adequate models

**Stackelberg games**: non zero-sum games

- System: a specific player called the **leader**
- Environment: composed of the other players called **followers**
- The leader first **announces** his strategy and then the followers **respond** by playing **rationally** given that strategy
- The leader wants to satisfy his objective whatever the rational response of the followers

In the next slides

- **One follower**: presentation of the **new model** proposed in [BRT21] and the obtained results [BRT21, BRT22]
- **Several followers**: some results presented at the end of the talk
1. Reactive synthesis

2. Stackelberg non zero-sum games

3. Stackelberg-Pareto verification

4. Stackelberg-Pareto synthesis

5. Rational synthesis/verification
Stackelberg-Pareto games

Definitions

- **Game arena**: graph $G = (V, V_0, V_1, E, v_0)$ with $(V_0, V_1)$ a partition of $V$ and $v_0$ an initial vertex
- **Two players**: Player $i$ that controls vertices of $V_i$, $i = 0, 1$. Player 0 is the leader and Player 1 is the follower
- **Play**: infinite path starting from $v_0$
- **Objective** for Player $i$: subset $\Omega$ of plays. A play $\rho$ satisfies $\Omega$ if $\rho \in \Omega$

Example

- **Player 0**: circle vertices
- **Player 1**: square vertices
- **Objective $\Omega_0$** of Player 0: reach $\{v_6, v_7\}$
Stackelberg-Pareto games

Definitions

- **Stackelberg-Pareto game**: $\mathcal{G} = (G, \Omega_0, \Omega_1, \ldots, \Omega_t)$ with objective $\Omega_0$ for Player 0 and $t$ objectives $\Omega_1, \ldots, \Omega_t$ for Player 1
- **Strategy** $\sigma_0 : V^* \times V_0 \rightarrow V$ announces the choices of Player 0 after each history $hv$ with $v \in V_0$
- **Plays** $\sigma_0 = \{\text{plays } \rho \mid \rho \text{ consistent with } \sigma_0\}$
- **Payoff of** $\rho \in \text{Plays}_{\sigma_0}$ for Player 1: Boolean vector $\text{pay}(\rho) \in \{0, 1\}^t$

Example

- $\Omega_0$: reach $\{v_6, v_7\}$
- 3 objectives $\Omega_1, \Omega_2, \Omega_3$
- **Strategy** $\sigma_0$: choice of $v_3 \rightarrow v_7$ after history $v_0 v_2 v_3$
- **Plays** $\sigma_0 = \{v_0 v_1^{\omega}, v_0 v_2 v_3 v_7^{\omega}, v_0 v_2 v_4^{\omega}\}$
Stackelberg-Pareto games

Rationality of Player 1

- Componentwise order $<$ on the payoffs $\text{pay}(\rho) \in \{0, 1\}^t$, $\forall \rho \in \text{Plays}_{\sigma_0}$
- Set $P_{\sigma_0}$ of Pareto-optimal payoffs of $\text{Plays}_{\sigma_0}$ w.r.t. $<$
- Player 1 only responds with plays $\rho \in \text{Plays}_{\sigma_0}$ with a Pareto-optimal payoff $\text{pay}(\rho) \in P_{\sigma_0}$
- Goal of Player 0: announce $\sigma_0$ such that $\Omega_0$ is satisfied by every such rational response

Example

- $\Omega_0$: reach $\{v_6, v_7\}$
- $\text{Plays}_{\sigma_0} =$
  $\{v_0 v_1^\omega, v_0 v_2 v_3 v_7^\omega, v_0 v_2 v_4^\omega\}$
- $P_{\sigma_0} =$
  $\{(0, 0, 1), (1, 1, 0), (1, 0, 0)\}$
<table>
<thead>
<tr>
<th></th>
<th>Reactive synthesis</th>
<th>Stackelberg games</th>
<th>Pareto verification</th>
<th>Pareto synthesis</th>
<th>Rational synthesis/verification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Reactive synthesis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Stackelberg non zero-sum games</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Stackelberg-Pareto verification</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Stackelberg-Pareto synthesis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Rational synthesis/verification</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Stackelberg-Pareto verification

Given a strategy $\sigma_0$ announced by Player 0, verify whether or not his goal is satisfied

Stackelberg-Pareto verification problem (SPV problem)

Given a Stackelberg-Pareto game $G = (G, \Omega_0, \Omega_1, \ldots, \Omega_t)$ where the strategy of $\sigma_0$ of Player 0 is fixed, decide whether every play in $\text{Plays}_{\sigma_0}$ with a Pareto-optimal payoff satisfies the objective of Player 0

Example

- $\Omega_0$: reach $\{v_6, v_7\}$
- $\text{Plays}_{\sigma_0} =$
  $\{v_0 v_1^\omega, v_0 v_2 v_3 v_7^\omega, v_0 v_2 v_4^\omega\}$
- $P_{\sigma_0} =$
  $\{(0, 0, 1), (1, 1, 0), (1, 0, 0)\}$
- No, $\Omega_0$ not always satisfied
Stackelberg-Pareto verification

Stackelberg-Pareto verification problem (SPV problem)

Given a Stackelberg-Pareto game $G = (G, \Omega_0, \Omega_1, \ldots, \Omega_t)$ where the strategy of $\sigma_0$ of Player 0 is fixed, decide whether every play in $Plays_{\sigma_0}$ with a Pareto-optimal payoff satisfies the objective of Player 0.

Theorem [BRT22]

The SPV problem is co-NP-complete for parity objectives, with a fixed-parameter algorithm (exponential in $t$).

Remarks

- **Parity**: a classical way to define $\omega$-regular objectives (reachability, safety, Büchi, co-Büchi, Streett, Rabin, Muller, LTL, etc).
- **Restriction to finite-memory strategies** $\sigma_0$, i.e., described by a finite automaton.
- **Fixed-parameter complexity**: in practice parameter $t$ is small.
Stackelberg-Pareto verification

Idea of the proof for co-NP membership

- Consider the complement of the SPV problem: does there exist a play in $\text{Plays}_{\sigma_0}$ with a Pareto-optimal payoff and not satisfying $\Omega_0$?

- Algorithm
  - non-deterministically guess a payoff $p \in \{0, 1\}^t$ (polynomial size)
  - check that there exists a play with payoff $p$ ($p$ is realizable)
  - check that there exists no play with a greater payoff ($p$ is Pareto-optimal)
  - check that there exists a play with payoff $p$ and not satisfying $\Omega_0$

- The last three checks can be done in polynomial time (using automaton)

- Therefore in co-NP
1 Reactive synthesis

2 Stackelberg non zero-sum games

3 Stackelberg-Pareto verification

4 Stackelberg-Pareto synthesis

5 Rational synthesis/verification
Problem

Stackelberg-Pareto Synthesis Problem (SPS problem)

Given a Stackelberg-Pareto game $G = (G, \Omega_0, \Omega_1, \ldots, \Omega_t)$, decide whether there exists a strategy $\sigma_0$ for Player 0 such that for every play $\rho \in \text{Plays}_{\sigma_0}$ with $\text{pay}(\rho) \in P_{\sigma_0}$, it holds that $\rho \in \Omega_0$.

Example

- Yes, such a strategy $\sigma_0$ exists:
  - after $v_0 v_2 v_3$: $v_3 \rightarrow v_5$
  - after $v_0 v_2 v_3 v_5 v_3$: $v_3 \rightarrow v_7$

- $\text{Plays}_{\sigma_0} =$
  $$\{v_0 v_1^\omega, v_0 v_2 v_3 v_5 v_6^\omega, v_0 v_2 v_3 v_5 v_3 v_7^\omega, v_0 v_2 v_4^\omega\}$$

- $P_{\sigma_0} =$
  $$\{(0, 0, 1), (0, 1, 1), (1, 1, 0), (1, 0, 0)\}$$
Results

Theorem [BRT21]
The SPS problem is \textit{NEXPTIME-complete} for \textit{parity objectives}, with a fixed-parameter algorithm (double exponential in $t$ and exponentiel in the highest priorities).

Remark
- For \textit{reachability objectives}, the SPS problem is \textit{NEXPTIME-complete} and becomes \textit{NP-complete} on tree arenas.
NEXPTIME-membership

Idea of the proof for NEXPTIME-membership

- If Player 0 has a solution $\sigma_0$ to the SPS problem, then he has a finite-memory one with an exponential size

- Algorithm
  - non-deterministically guess a strategy $\sigma_0$ (with exponential size)
  - check that it is a solution in exponential time (using automaton)

Constructing a finite-memory strategy

- Given a solution $\sigma_0$, take one play $\rho_i$ (witness) for each Pareto-optimal payoff $p_i \in P_{\sigma_0}$
NEXPTIME-membership

Constructing a finite-memory strategy

- Given a solution $\sigma_0$, take one play $\rho_i$ (witness) for each Pareto-optimal payoff $p_i \in P_{\sigma_0}$
- Modify $\sigma_0$ into $\hat{\sigma}_0$ on deviations from the witnesses: punish by imposing $\Omega_0$ or a not Pareto-optimal payoff
- Modify $\hat{\sigma}_0$ into $\tilde{\sigma}_0$: decompose each $\rho_i$ into at most exponentially many parts and compact it as $c\rho_i$
NP-hardness for reachability objectives on tree arenas

Idea of the proof: NP-hardness is shown using the set cover problem

Given

- $C = \{e_1, e_2, \ldots, e_n\}$ of $n$ elements
- $m$ subsets $S_1, S_2, \ldots, S_m$ such that $S_i \subseteq C$
- an integer $k \leq m$

Find $k$ indexes $i_1, i_2, \ldots, i_k$ such that $C = \bigcup_{j=1}^{k} S_{i_j}$.

Devise a Stackelberg-Pareto game such that Player 0 has a solution to the SPS problem $\iff$ solution to the set cover problem
NP-hardness for reachability objectives on tree arenas

\[ C = \{e_1, e_2, e_3\}, \quad S_1 = \{e_1, e_3\}, \quad S_2 = \{e_2\}, \quad S_3 = \{e_1, e_2\}, \quad k = 2 \]

- Every play in \( G_1 \) is consistent with any strategy of Player 0 and does not satisfy \( \Omega_0 \)
- Hence in a solution, payoffs from \( G_1 \) cannot be Pareto-optimal and must be \(<\) than some payoff in \( G_2 \)
1. Reactive synthesis
2. Stackelberg non zero-sum games
3. Stackelberg-Pareto verification
4. Stackelberg-Pareto synthesis
5. Rational synthesis/verification
Another model - Several followers

Recap

- Environment: one follower with several objectives
- He responds to the announced strategy $\sigma_0$ by following a play with Pareto-optimal payoff

Another approach $[\text{KPV16, GMP}^+ 17]$

- Environment: several followers, each with one objective
- Stackelberg game $\mathcal{G} = (G, \Omega_0, \Omega_1, \ldots, \Omega_t)$ with an arena $G = (V, (V_i)_{i=0}^t, E, v_0)$, a set $\Pi = \{0, 1, \ldots, t\}$ of players, and an objective $\Omega_i$ for Player $i$, $i \in \Pi$
- These players respond to $\sigma_0$ with a strategy profile that is an equilibrium with respect to their own objectives
- Equilibrium: Nash equilibrium, subgame-perfect equilibrium, ...
Nash equilibrium

Let $\sigma_0$ be a strategy for Player 0.

- A $\sigma_0$-Stackelberg profile is a strategy profile $\sigma = (\sigma_0, (\sigma_i)_{i \in \Pi \setminus \{0\}})$ such that $pay_i(\langle \sigma \rangle) \geq pay_i(\langle \sigma'_i, \sigma_{-i} \rangle)$ for all players $i \in \Pi \setminus \{0\}$ and all strategies $\sigma'_i$ for Player $i$ where
  - $\langle \sigma \rangle$ is the play consistent with all strategies of $\sigma$.
  - $\langle \sigma'_i, \sigma_{-i} \rangle$ is the play consistent with all strategies of $\sigma$, except that $\sigma'_i$ replaces $\sigma_i$.

- No player $i \neq 0$ has an incentive to deviate from $\sigma_i$ in a way to increase his payoff.

Example

- Player 0: circle vertices
- Player 1: square vertices
- Player 2: diamond vertices
- Strategy $\sigma_0$: choice of $v_2 \rightarrow v_3$
Nash equilibrium

Rational synthesis problem (RS problem)
Given a Stackelberg game $G = (G, \Omega_0, \Omega_1, \ldots, \Omega_t)$, decide whether there exists a strategy $\sigma_0$ for Player 0 such that for every $\sigma_0$-Stackelberg profile $\sigma$, it holds that $\langle \sigma \rangle \in \Omega_0$

Rational verification problem (RV problem)
Given a Stackelberg game $G = (G, \Omega_0, \Omega_1, \ldots, \Omega_t)$ where the strategy $\sigma_0$ of Player 0 is fixed, decide whether for every $\sigma_0$-Stackelberg profile $\sigma$, it holds that $\langle \sigma \rangle \in \Omega_0$
Results

Theorem

For Stackelberg games

- with LTL objectives, the RS problem is 2EXPTIME-complete [KS22] as well as the RV problem [GNPW20]
- with parity objectives, the RS problem is in EXPTIME and PSPACE-hard [CFGR16] and the RV problem is co-NP-complete [Umm08]

Additional results for subgame perfect equilibria (instead of NEs) in [KPV16, BRvdB22]
Conclusion

- **Classical reactive synthesis**
  - Model of two-player zero-sum games
  - System and environment have **opposed** objectives

- Model of Stackelberg **non** zero-sum games with **one** follower

- Verification and synthesis
  - Complexity class and fixed-parameter complexity for $\omega$-regular objectives

- Model of Stackelberg non zero-sum games with **several** followers

Thanks for your attention!


Léonard Brice, Jean-François Raskin, and Marie van den Bogaard, *On the complexity of spes in parity games*, 30th EACSL Annual Conference on Computer Science Logic, CSL 2022, February 14-19, 2022, Göttingen, Germany (Virtual Conference) (Florin Manea and
Rational synthesis/verification


Example

Autonomous robotized lawnmower [Ran12]

- System: lawnmower with solar panels and fuel tank
- Environment: weather and cat
Example

Objectives

- **Büchi** objective: grass must be cut infinitely often
- **Energy** objective: battery and fuel must never drop below 0
- **Mean-payoff** objective: average time per action must be less than 10 in the long run

Controller as the following strategy

- If sunny, mow slowly
- If cloudy
  - If solar battery $\geq 1$, mow on battery
  - otherwise, if fuel level $\geq 2$, mow on fuel
  - otherwise, rest at the base