

On Quantified Observability Analysis in Multi-Agent Systems

ECAI 2023

Chunyan Mu[1] Jun Pang[2]

[1] University of Aberdeen [2] University of Luxembourg

4 October, 2023

Outline



1. Motivations and overview
2. Modelling POMASs
3. Observability Specification: the logic oPATL
4. Verification of Observability Properties
5. Conclusions

Outline



1. Motivations and overview
2. Modelling POMASs
3. Observability Specification: the logic oPATL
4. Verification of Observability Properties
5. Conclusions

Motivations and overview



Motivations

- In MASs, agents' observability on system behaviours may improve the overall team performance, but
  - may also release sensitive information to malicious (inside/outside) agents and cause information leakage.
- Information exposure issue should also play a role in making decisions for agents.
- A quantified observability analyses can thus be useful in MASs to assist decision-making in MASs design to address such concerns.

Motivations and overview



What do we do?

- A framework for specification and verification of quantified observability properties in MASs:
  - The model: Partially Observable multi-player Markov chain game
  - Specification: the logic of oPATL, to express (probabilistic) observability properties
  - Verification: probabilistic model checking, to allow automatically verify quantified quantified observability properties in MASs

Motivations and overview



What do we do?

- A framework for specification and verification of quantified observability properties in MASs:
  - The model: Partially Observable multi-player Markov chain game
  - Specification: the logic of oPATL, to express (probabilistic) observability properties
  - Verification: probabilistic model checking, to allow automatically verify quantified quantified observability properties in MASs
- Implementation and case studies:
  - prototype: built upon the PRISM model checker
  - examples: voting, supply-chain, p2p-network

Outline



1. Motivations and overview
2. Modelling POMASs
3. Observability Specification: the logic oPATL
4. Verification of Observability Properties
5. Conclusions

Partially Observable Multiagent Systems (PGS)



Definition: probabilistic game structure

A *PGS* is a tuple $\mathcal{G} = (S, \text{Act}, \delta)$, where:

- S is a finite set of *states*;
- $\text{Act} = \prod_{j \in \text{Ag}} \text{Act}_j$ is a finite set of joint actions (decisions) of the agents in Ag, $\text{Act}_j \subseteq \Sigma$ is the set of actions that $j \in \text{Ag}$ can perform;
- $\delta : S \to 2^{\text{Dist}(\text{Act} \times S)}$ is the *probabilistic transition relation*; for state $s \in S$, $\delta(s)$ is the distribution for next state; s is a *terminal state* if $\delta(s) = \emptyset$.

Partially Observable Multiagent Systems



Definition: Observations

- Let Θ be a finite alphabet for observables, and $\Theta^{\epsilon} = obs \cup \{\epsilon\}$ where ϵ denotes the invisible/hidden action.
- An observation function on paths is a labelled-based function

$$obs : \mathrm{Paths}_{\mathcal{G}}(s) \rightarrow (\Theta_1 \times \Theta_2 \times \cdots \times \Theta_n)^{\infty}$$

where $\Theta_j \subseteq \Theta$ denotes a finite set of observables for $j \in \mathrm{Ag}$.

Partially Observable Multiagent Systems



Definition: POMAS

A POMASis a tuple $\mathcal{M} = (\mathrm{Ag}, \mathcal{G}, s_0, \mathrm{Ap}, L, \{obs_i\}_{i \in \mathrm{Ag}})$, where:

- $\mathrm{Ag} = \{1, \ldots, n\}$ is a finite set of intelligent *agents*;
- $\mathcal{G} = (S, \mathrm{Act}, \delta)$ is a *fully PGS*;
- $s_0 \in S$ is the *initial state*;
- Ap is a finite set of *atomic propositions*;
- $L : S \to 2^{\mathrm{Ap}}$ is the state labelling function mapping each state to a set of atomic state proposition taken from set Ap;
- $obs_i : \mathrm{Paths}_{\mathcal{G}}(s_0) \to (\Theta_1 \times \cdots \times \Theta_n)^{\infty}$ is an observation function for agent $i \in \mathrm{Ag}$.

Partially Observable Multiagent Systems



Example: Voting

1. The chair initiates the voting procedure;
2. Voters simultaneously propose their votes;
3. Each voter commits his vote once he makes his decision;
4. If a voter is waiting during this process, he can partially observe the behaviour of other voters and gather indications, such as identifying the dominant candidate;
5. Based on their observations, the voter can make their own voting decision;
6. Once all voters have committed their votes, the chair closes the voting session.

Partially Observable Multiagent Systems



Example: Voting

Actions	ζ_1(Action)	ζ_2(Action)	ζ_3(Action)	Descriptions
$\mathtt{Op}_0$	$\mathtt{Op}_0$	$\mathtt{Op}_0$	$\mathtt{Op}_0$	the chair opens a voting session
$\mathtt{Cl}_0$	$\mathtt{Cl}_0$	$\mathtt{Cl}_0$	$\mathtt{Cl}_0$	the chair closes a voting session
$\mathtt{W}_i$	$\mathtt{W}$	$\mathtt{W}$	$\mathtt{W}$	agent i is waiting, $i \in \{0,1,2,3\}$
$\mathtt{V}_1^X$	X_1	ϵ	X	voter 1 votes candidate X
$\mathtt{V}_1^Y$	Y_1	ϵ	Y	voter 1 votes candidate Y
$\mathtt{V}_2^X$	ϵ	X_2	X	voter 2 votes candidate X
$\mathtt{V}_2^Y$	ϵ	Y_2	Y	voter 2 votes candidate Y
$\mathtt{V}_3^X$	ϵ	ϵ	X_3	voter 3 votes candidate X
$\mathtt{V}_3^Y$	ϵ	ϵ	Y_3	voter 3 votes candidate Y

Partially Observable Multiagent Systems



Example: Voting

- Assume voter 2 is the observer, the actions of voting X by 1 and 3 are not visible to voter 2.
- Consider a path:

$$\rho = s_0 \xrightarrow{\mathtt{Op}_0\mathtt{W}_1\mathtt{W}_2\mathtt{W}_3} s_1 \xrightarrow{\frac{1}{2}.\mathtt{W}_0\mathtt{V}_1^X\mathtt{W}_2\mathtt{W}_3} s_2 \xrightarrow{\frac{1}{2}.\mathtt{W}_0\mathtt{W}_1\mathtt{V}_2^Y\mathtt{W}_3} s_3 \xrightarrow{\frac{2}{3}.\mathtt{W}_0\mathtt{W}_1\mathtt{W}_2\mathtt{V}_3^Y} s_4 \xrightarrow{\mathtt{Cl}_0\ \mathtt{W}_1\mathtt{W}_2\mathtt{W}_3} s_5$$

- The observation from 2's view:

$$obs_2(\rho) \quad = \quad \mathtt{Op}_0\mathtt{WWW}\ \mathtt{WWW}\ \mathtt{WW}Y_2\mathtt{W}\ \mathtt{WWW}\ \mathtt{Cl}_0\mathtt{WWW} \quad w.p. \quad \frac{1}{6}.$$

Partially Observable Multiagent Systems



Definition: Mixed Strategy

A *mixed strategy* for agent i is a function π_i:

$$\pi_i \triangleq obs_i(\text{Paths}_{\mathcal{G}}(s_0)) \to \text{Dist}(\text{Act}_i)$$

such that, if $\pi_i(\rho)(a_i) > 0$ then $a_i \in \text{Act}_i(last(\rho))$. The set of all strategies of agent i is denoted Π_i.

Definition: Strategy profile

A *strategy profile* for POMAS $\mathcal{M}$ is a tuple

$$\pi = (\pi_1, \ldots, \pi_n) \in \Pi_1 \times \cdots \times \Pi_n$$

producing a strategy for each agent of the system.

Partially Observable Multiagent Systems



Example: strategy in voting

- Assume voter 3 is the observer, and consider the path:

$$\rho = s_0 \xrightarrow{\mathtt{Op}_0\mathtt{W}_1\mathtt{W}_2\mathtt{W}_3} s_1 \xrightarrow{\frac{1}{4}.\mathtt{W}_0 V_1^X V_2^Y \mathtt{W}_3} s_2 \xrightarrow{\frac{1}{3}.\mathtt{W}_0\mathtt{W}_1\mathtt{W}_2 V_3^X} s_3 \xrightarrow{\mathtt{Cl}_0\mathtt{W}_1\mathtt{W}_2\mathtt{W}_3} s_4$$

Partially Observable Multiagent Systems



Example: strategy in voting

- Assume voter 3 is the observer, and consider the path:
  $\rho = s_0 \xrightarrow{\mathtt{Op}_0\mathtt{W}_1\mathtt{W}_2\mathtt{W}_3} s_1 \xrightarrow{\frac{1}{4}.\mathtt{W}_0 V_1^X V_2^Y \mathtt{W}_3} s_2 \xrightarrow{\frac{1}{3}.\mathtt{W}_0\mathtt{W}_1\mathtt{W}_2 V_3^X} s_3 \xrightarrow{\mathtt{Cl}_0\mathtt{W}_1\mathtt{W}_2\mathtt{W}_3} s_4$
- The observation on the above path from voter 3's view would then be:
  $obs_3(\rho) = \mathtt{Op}_0\mathtt{WWW}\ \mathtt{W}X\mathtt{Y}\mathtt{W}\ \mathtt{WWW}X_3\ \mathtt{Cl}_0\mathtt{WWW}$ *w.p.* $\frac{1}{12}$.

Partially Observable Multiagent Systems



Example: strategy in voting

- Assume voter 3 is the observer, and consider the path:
  $\rho = s_0 \xrightarrow{\mathtt{Op}_0\mathtt{W}_1\mathtt{W}_2\mathtt{W}_3} s_1 \xrightarrow{\frac{1}{4}.\mathtt{W}_0 V_1^X V_2^Y \mathtt{W}_3} s_2 \xrightarrow{\frac{1}{3}.\mathtt{W}_0\mathtt{W}_1\mathtt{W}_2 V_3^X} s_3 \xrightarrow{\mathtt{Cl}_0\mathtt{W}_1\mathtt{W}_2\mathtt{W}_3} s_4$
- The observation on the above path from voter 3's view would then be:
  $obs_3(\rho) = \mathtt{Op}_0\mathtt{WWW}\ \mathtt{W}X\mathtt{Y}\mathtt{W}\ \mathtt{WWW}X_3\ \mathtt{Cl}_0\mathtt{WWW}$ *w.p.* $\frac{1}{12}$.
- The observer's information of knowledge obtained from her observation might influence her decision on voting.

Partially Observable Multiagent Systems



Example: strategy in voting

- Assume voter 3 is the observer, and consider the path:
  $\rho = s_0 \xrightarrow{\mathtt{Op}_0\mathtt{W}_1\mathtt{W}_2\mathtt{W}_3} s_1 \xrightarrow{\frac{1}{4}.\mathtt{W}_0 V_1^X V_2^Y \mathtt{W}_3} s_2 \xrightarrow{\frac{1}{3}.\mathtt{W}_0\mathtt{W}_1\mathtt{W}_2 V_3^X} s_3 \xrightarrow{\mathtt{Cl}_0\mathtt{W}_1\mathtt{W}_2\mathtt{W}_3} s_4$
- The observation on the above path from voter 3's view would then be:
  $obs_3(\rho) = \mathtt{Op}_0\mathtt{WWW}\ \mathtt{W}X\mathtt{Y}\mathtt{W}\ \mathtt{WWW}X_3\ \mathtt{Cl}_0\mathtt{WWW}$ *w.p.* $\frac{1}{12}$.
- The observer's information of knowledge obtained from her observation might influence her decision on voting.
- Assume the observer (voter 3) is not able to see whom other voters have voted, but she can see how many ballots each candidate has received
- A basic strategy to reflect such an influence is that:
  - she will vote the dominant candidate if there is one, otherwise
  - she will vote the candidates under her preferred distribution.

Outline



1. Motivations and overview
2. Modelling POMASs
3. Observability Specification: the logic oPATL
4. Verification of Observability Properties
5. Conclusions

Observability Specification: the logic oPATL



Observability and Opacity

Let $\mathcal{M} = (\mathrm{Ag}, \mathcal{G}, s_0, \mathrm{Ap}, L, \{obs_i\}_{i \in \mathrm{Ag}})$. Given a predicate φ over $\mathrm{Paths}_{\mathcal{G}}(s_0)$, we say φ is opaque w.r.t. obs_i if:

- for every path $\rho \in [\![\varphi]\!]$, there is a path $\rho' \in [\![\bar{\varphi}]\!]$ s.t.

$$obs(\rho) = obs(\rho')$$

- i.e., under obs_i, all paths satisfying φ are covered by paths in $[\![\bar{\varphi}]\!]$:

$$obs_i([\![\varphi]\!]) \subseteq obs_i([\![\bar{\varphi}]\!])$$

where $[\![\bar{\varphi}]\!] \triangleq \mathrm{Paths}_{\mathcal{G}}(s_0) \setminus [\![\varphi]\!]$.

Observability Specification: the logic oPATL



The logic oPATL: syntax

Three classes of formulae: **state formulae** ϕ, **path formulae** ψ, and **observability formulae** Φ:

$$
\begin{array}{lcl}
\phi & ::= & a \mid \neg\phi \mid \phi \wedge \phi \mid \mathbf{P}_{\bowtie p}\langle \mathrm{A} \rangle [\psi] \mid \mathcal{D}_{\bowtie p}\langle \mathrm{A} \rangle [\Phi] \\
\psi & ::= & \mathbf{X}\phi \mid \phi \mathbf{U} \phi \mid \phi \mathbf{R} \phi \mid \neg\psi \mid \psi \wedge \psi \\
\Phi & ::= & \mathbf{O}_i [\psi] \mid \neg\Phi \mid \Phi \wedge \Phi
\end{array}
$$

Observability Specification: the logic oPATL



Voting Example: observability

- Consider the property ψ: *"eventually candidate X wins"*
- The probabilistic observability property of voter 2:

$$\mathcal{D}_{\leq \frac{1}{3}}\langle 1,2,3\rangle[\mathbf{O}_2[\psi]] = \texttt{true}$$

since:

$$\begin{aligned}&\mathsf{Prob}([\![\langle 1,2,3\rangle \mathbf{O}_2[\psi]]\!])\\ =&\mathsf{Prob}(\{\frac{1}{12}.\texttt{Op}_0 V_1^X V_2^X V_3^X \texttt{Cl}_0, \frac{1}{6}.\texttt{Op}_0 V_1^X V_2^X V_3^Y \texttt{Cl}_0\}) = \frac{1}{4} < \frac{1}{3}\end{aligned}$$

Outline



1. Motivations and overview
2. Modelling POMASs
3. Observability Specification: the logic oPATL
4. Verification of Observability Properties
5. Conclusions

Verification of Observability Properties



Verification of $\mathcal{D}_{\bowtie p}[\mathbf{O}_i\langle\mathrm{A}\rangle\psi]$

Given a POMAS $\mathcal{M} = (\mathcal{G}, s_0, \mathrm{Ag}, \mathrm{Ap}, \{obs_i\}_{i\in\mathrm{Ag}})$, a task in property ψ required to be completed under a strategy π_{A} of a coalition $\mathrm{A} \in \mathrm{Ag}$, the probabilistic verification problem of observability property is to decide whether $s_0 \models_{\mathcal{M}} \mathcal{D}_{\bowtie p}(\mathbf{O}_i\ \langle\mathrm{A}\rangle[\psi])$, i.e.,

$$\mathbb{P}_{s_0}([\![\langle\mathrm{A}\rangle[\psi]]\!] \setminus obs_i^{-1}(obs_i([\![\langle\mathrm{A}\rangle[\neg\psi]]\!]))) \ \bowtie\ p.$$

Verification of Observability Properties



Checking $s \models_{\mathcal{M}} \mathcal{D}_{\bowtie p}(\mathbf{O}_i \langle\langle A \rangle\rangle[\psi])$

- Find all consistent paths Π and the corresponding traces Λ, satisfying ψ under mixed strategy π_A of coalition A:
  $\Pi = \{\mathrm{Reg}(\rho_{\pi_A}) \mid \rho_{\pi_A} \models_{\mathcal{M}} \psi\}$ $\quad$ $\Lambda = \{\mathrm{erase}(\rho) \mid \rho \in \Pi\}$.

Verification of Observability Properties



Checking $s \models_{\mathcal{M}} \mathcal{D}_{\bowtie p}(\mathbf{O}_i \langle\!\langle A \rangle\!\rangle [\psi])$

- Find all consistent paths Π and the corresponding traces Λ, satisfying ψ under mixed strategy π_A of coalition A:
  $\Pi = \{\mathrm{Reg}(\rho_{\pi_A}) \mid \rho_{\pi_A} \models_{\mathcal{M}} \psi\} \qquad \Lambda = \{\mathrm{erase}(\rho) \mid \rho \in \Pi\}.$
- Find all consistent paths Π' and the corresponding traces Λ, violating ψ under mixed strategy π_A of coalition A:
  $\Pi' = \{\mathrm{Reg}(\rho'_{\pi_A}) \mid \rho_{\pi_A} \not\models_{\mathcal{M}} \psi\} \qquad \Lambda' = \{\mathrm{erase}(\rho') \mid \rho' \in \Pi'\}.$

Verification of Observability Properties



Checking $s \models_{\mathcal{M}} \mathcal{D}_{\bowtie p}(\mathbf{O}_i \langle\!\langle A \rangle\!\rangle [\psi])$

- Find all consistent paths Π and the corresponding traces Λ, satisfying ψ under mixed strategy π_A of coalition A:
  $\Pi = \{\mathrm{Reg}(\rho_{\pi_A}) \mid \rho_{\pi_A} \models_{\mathcal{M}} \psi\} \qquad \Lambda = \{\mathrm{erase}(\rho) \mid \rho \in \Pi\}.$
- Find all consistent paths Π' and the corresponding traces Λ, violating ψ under mixed strategy π_A of coalition A:
  $\Pi' = \{\mathrm{Reg}(\rho'_{\pi_A}) \mid \rho_{\pi_A} \not\models_{\mathcal{M}} \psi\} \qquad \Lambda' = \{\mathrm{erase}(\rho') \mid \rho' \in \Pi'\}.$
- Find all ψ-opaque traces:
  $\Lambda'' = \{\lambda'' \mid \lambda'' \in \Lambda \wedge \exists \lambda' \in \Lambda'.(obs_i(\lambda') = obs_i(\lambda''))\}.$

Verification of Observability Properties



Checking $s \models_{\mathcal{M}} \mathcal{D}_{\bowtie p}(\mathbf{O}_i \langle\!\langle A \rangle\!\rangle [\psi])$

- Find all consistent paths Π and the corresponding traces Λ, satisfying ψ under mixed strategy π_A of coalition A:
  $\Pi = \{\mathrm{Reg}(\rho_{\pi_A}) \mid \rho_{\pi_A} \models_{\mathcal{M}} \psi\} \qquad \Lambda = \{\mathrm{erase}(\rho) \mid \rho \in \Pi\}.$
- Find all consistent paths Π' and the corresponding traces Λ, violating ψ under mixed strategy π_A of coalition A:
  $\Pi' = \{\mathrm{Reg}(\rho'_{\pi_A}) \mid \rho_{\pi_A} \not\models_{\mathcal{M}} \psi\} \qquad \Lambda' = \{\mathrm{erase}(\rho') \mid \rho' \in \Pi'\}.$
- Find all ψ-opaque traces:
  $\Lambda'' = \{\lambda'' \mid \lambda'' \in \Lambda \wedge \exists \lambda' \in \Lambda'.(obs_i(\lambda') = obs_i(\lambda''))\}.$
- Compute the probability of ψ-observable traces:
  $d = \mathbb{P}_{s_0}([\![\mathbf{O}_i \langle\!\langle A \rangle\!\rangle [\psi]]\!]) = \sum_{\xi \in (\Lambda \setminus \Lambda'')} \mathrm{Prob}(\xi).$

Verification of Observability Properties



Checking $s \models_{\mathcal{M}} \mathcal{D}_{\bowtie p}(\mathbf{O}_i \langle\!\langle A \rangle\!\rangle [\psi])$

- Find all consistent paths Π and the corresponding traces Λ, satisfying ψ under mixed strategy π_A of coalition A:
  $\Pi = \{\mathrm{Reg}(\rho_{\pi_A}) \mid \rho_{\pi_A} \models_{\mathcal{M}} \psi\} \qquad \Lambda = \{\mathrm{erase}(\rho) \mid \rho \in \Pi\}.$
- Find all consistent paths Π' and the corresponding traces Λ, violating ψ under mixed strategy π_A of coalition A:
  $\Pi' = \{\mathrm{Reg}(\rho'_{\pi_A}) \mid \rho_{\pi_A} \not\models_{\mathcal{M}} \psi\} \qquad \Lambda' = \{\mathrm{erase}(\rho') \mid \rho' \in \Pi'\}.$
- Find all ψ-opaque traces:
  $\Lambda'' = \{\lambda'' \mid \lambda'' \in \Lambda \wedge \exists \lambda' \in \Lambda'.(obs_i(\lambda') = obs_i(\lambda''))\}.$
- Compute the probability of ψ-observable traces:
  $d = \mathbb{P}_{s_0}([\![\mathbf{O}_i \langle\!\langle A \rangle\!\rangle [\psi]]\!]) = \sum_{\xi \in (\Lambda \setminus \Lambda'')} \mathrm{Prob}(\xi).$
- Return true if $d \bowtie p$, return false otherwise.

Verification of Observability Properties



Voting Example: observability

Probabilistic verification of $\mathcal{D}_{=?}\langle A\rangle[\mathbf{O}_i[\psi]]$

```
Result: 0.25.{
    0.08333333333333333:vX1vX2vX3cl0->X2cl0,
    0.16666666666666666:vX1vX2vY3cl0->X2cl0
} (value in the initial state)
```

Outline



1. Motivations and overview
2. Modelling POMASs
3. Observability Specification: the logic oPATL
4. Verification of Observability Properties
5. Conclusions

Conclusions and Future work



Conclusions

- A formal framework for quantified specification and verification of observability properties in MASs.
- Applications: information security analysis in MASs

Conclusions and Future work



Conclusions

- A formal framework for quantified specification and verification of observability properties in MASs.
- Applications: information security analysis in MASs

Future work

- Integration of the observability operator into **Strategic Logic** and evaluation its suitability for various scenarios
- Trade-off: performance vs. information leakage/observability
  - e.g., game-theoretic methods

Thank You! 