Interactions of Knowledge and Strategies

Francesco Belardinelli

Laboratoire IBISC, Université d'Evry

LAMAS SING, Gdańsk – 24 September 2015

Motivation and Background

- **Perfect Information**: fixed-point characterisations of ATL operators

$$\langle\langle\Sigma\rangle\rangle G\phi \quad \leftrightarrow \quad \phi \wedge \langle\langle\Sigma\rangle\rangle X\langle\langle\Sigma\rangle\rangle G\phi \tag{1}$$

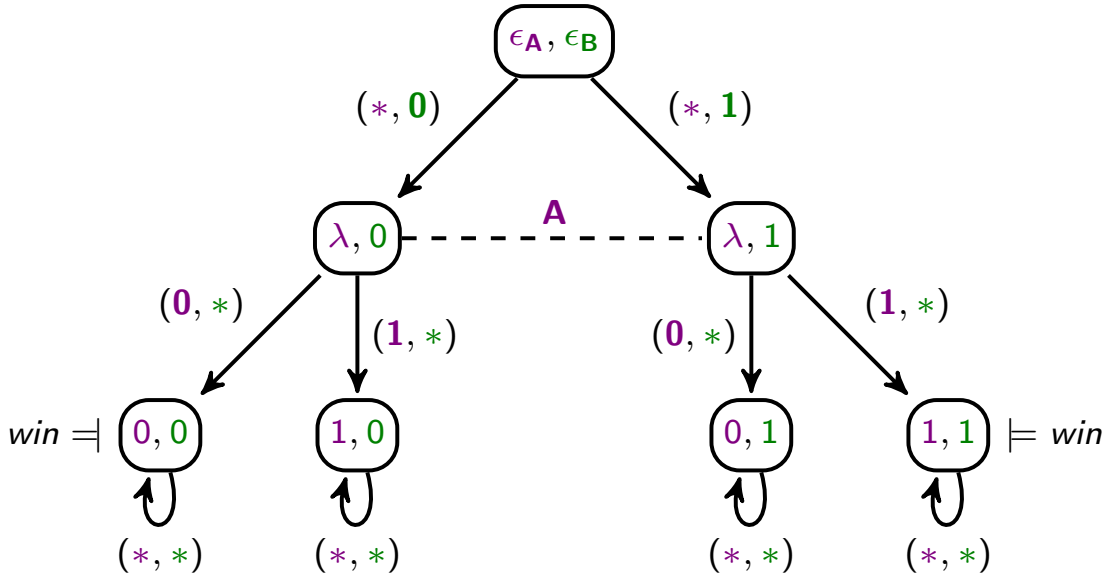
$$\langle\langle\Sigma\rangle\rangle F\phi \quad \leftrightarrow \quad \phi \vee \langle\langle\Sigma\rangle\rangle X\langle\langle\Sigma\rangle\rangle F\phi \tag{2}$$

$$\langle\langle\Sigma\rangle\rangle(\phi U\phi') \quad \leftrightarrow \quad \phi' \vee (\phi \wedge \langle\langle\Sigma\rangle\rangle X\langle\langle\Sigma\rangle\rangle(\psi U\phi')) \tag{3}$$

- **Useful Validities**: techniques for satisfiability [GS09] and model checking [AHK02, BDJ10]
- **The Problem**: (1)-(3) do not hold in the imperfect information semantics!

The Problem

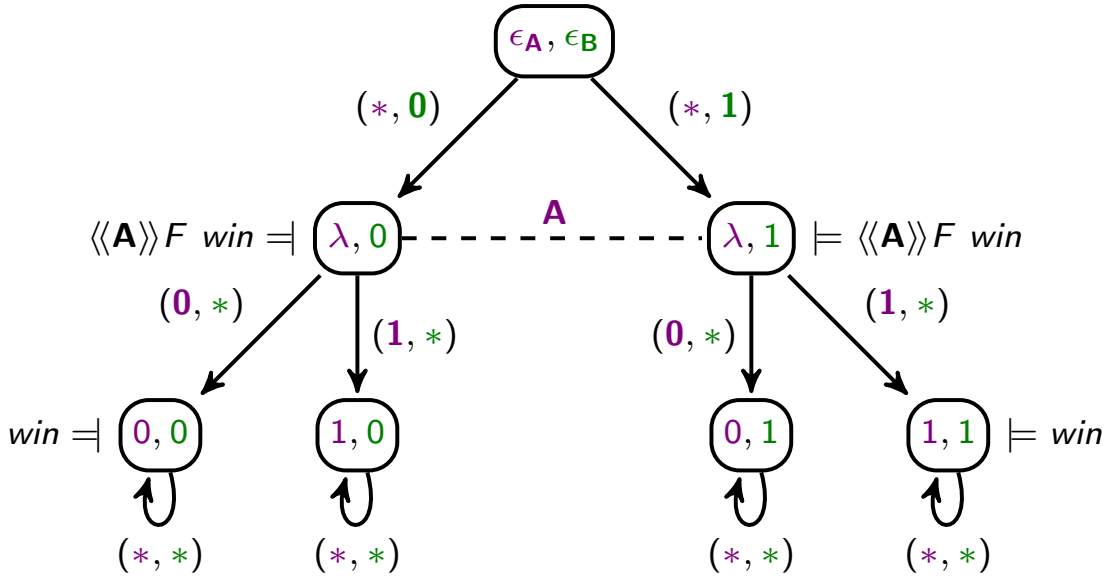
ATL with Imperfect Information



- **B**ob chooses secretly between 0 and 1
- at the next step **A**nne also chooses between 0 and 1
- **A**nne wins the game iff the values provided by the two players coincide
- the dotted line indicates epistemic indistinguishability

The Problem

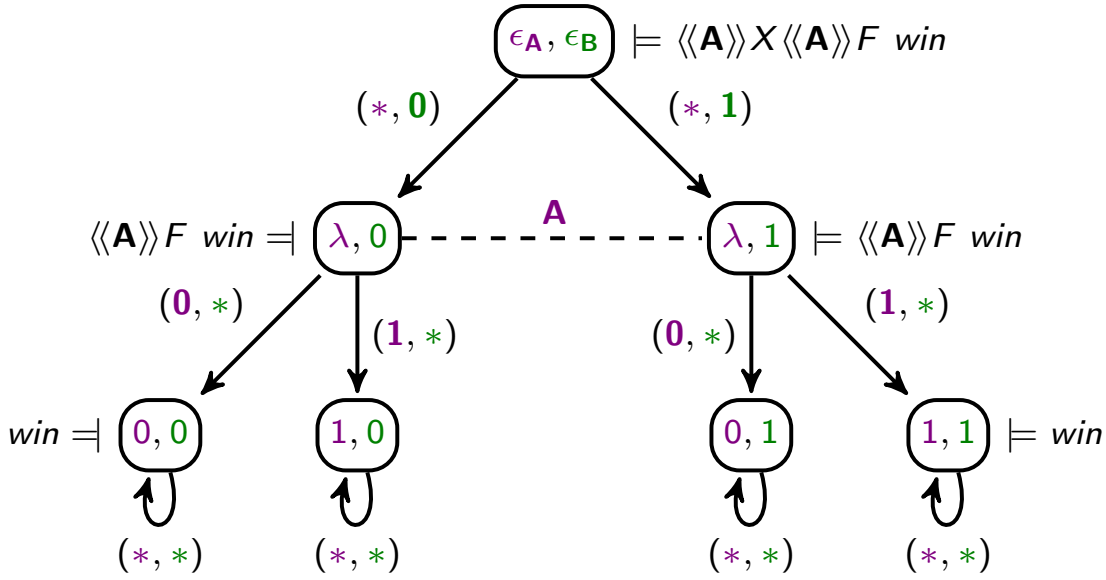
ATL with Imperfect Information



- **B**ob chooses secretly between 0 and 1
- at the next step **A**nne also chooses between 0 and 1
- **A**nne wins the game iff the values provided by the two players coincide
- the dotted line indicates epistemic indistinguishability

The Problem

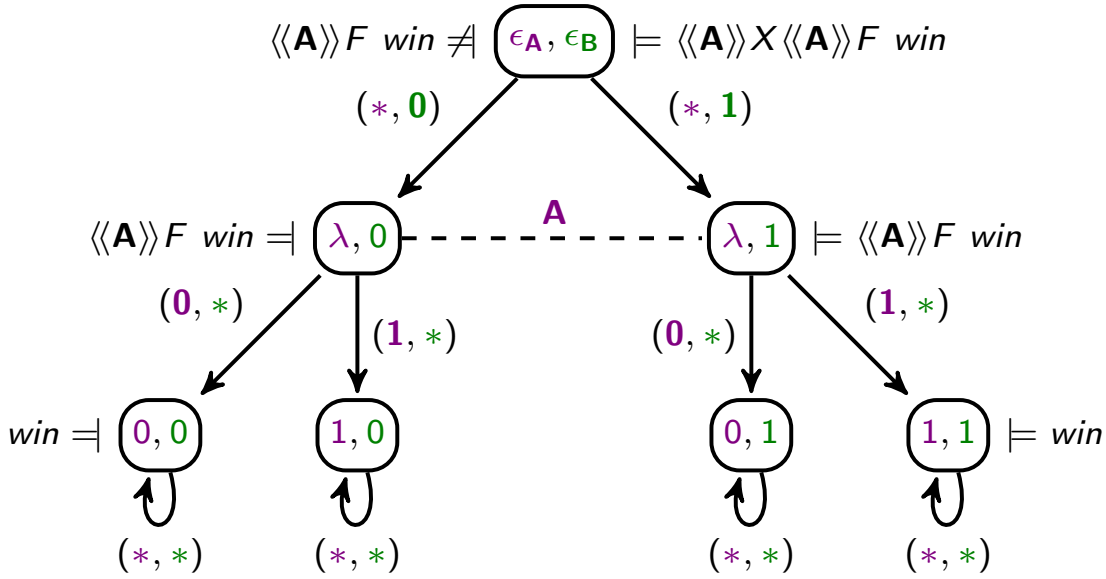
ATL with Imperfect Information



- **B**ob chooses secretly between 0 and 1
- at the next step **A**nne also chooses between 0 and 1
- **A**nne wins the game iff the values provided by the two players coincide
- the dotted line indicates epistemic indistinguishability

The Problem

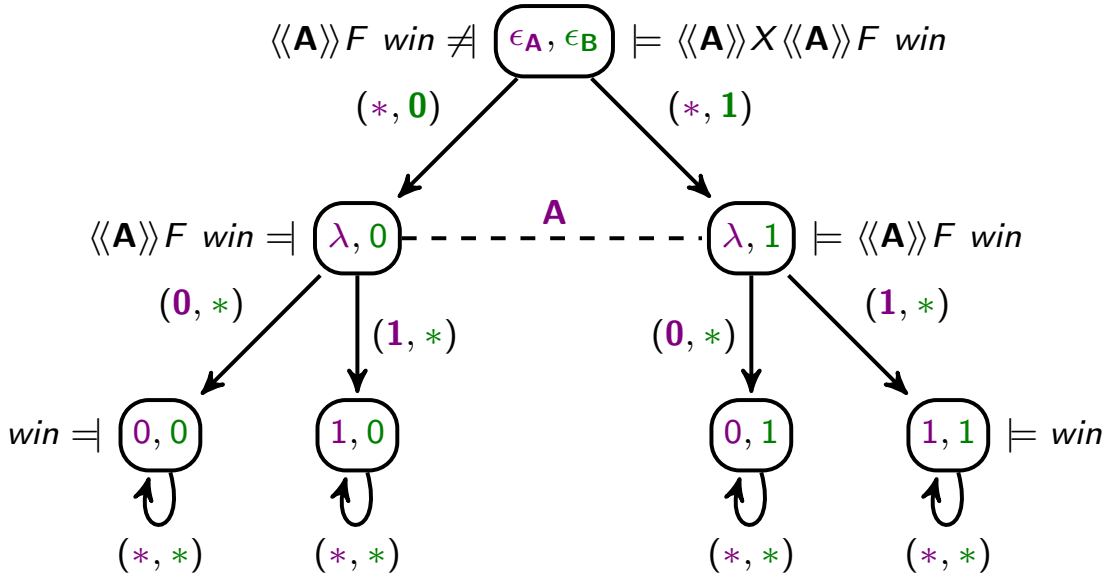
ATL with Imperfect Information



- **B**ob chooses secretly between 0 and 1
- at the next step **A**nne also chooses between 0 and 1
- **A**nne wins the game iff the values provided by the two players coincide
- the dotted line indicates epistemic indistinguishability

The Problem

ATL with Imperfect Information



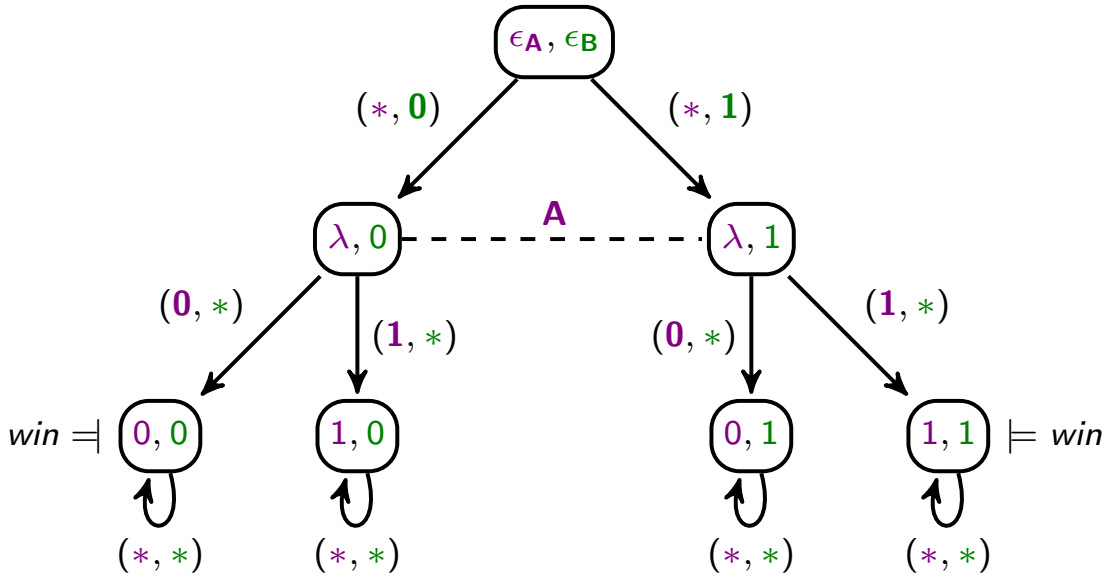
- **B**ob chooses secretly between 0 and 1
- at the next step **A**nne also chooses between 0 and 1
- **A**nne wins the game iff the values provided by the two players coincide
- the dotted line indicates epistemic indistinguishability
- **A**nne knows that there exists a strategy to win the game . . .

  . . . however, she is not able to point this strategy out

  $\Leftarrow$ **A**nne has *imperfect information* of the game

The Problem

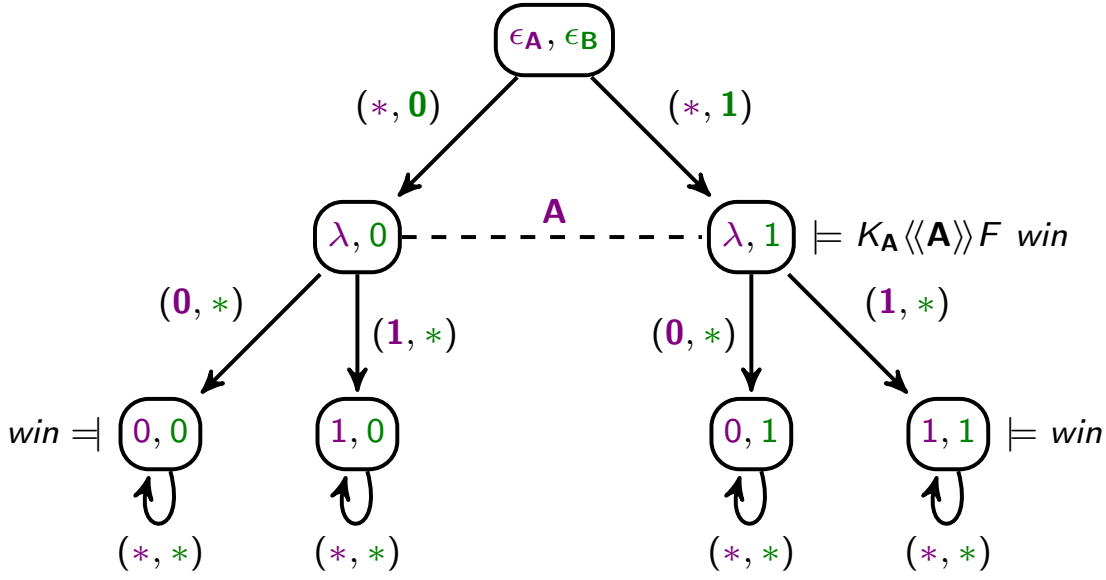
ATL with Imperfect Information



It looks like it's a question of knowledge

The Problem

ATL with Imperfect Information

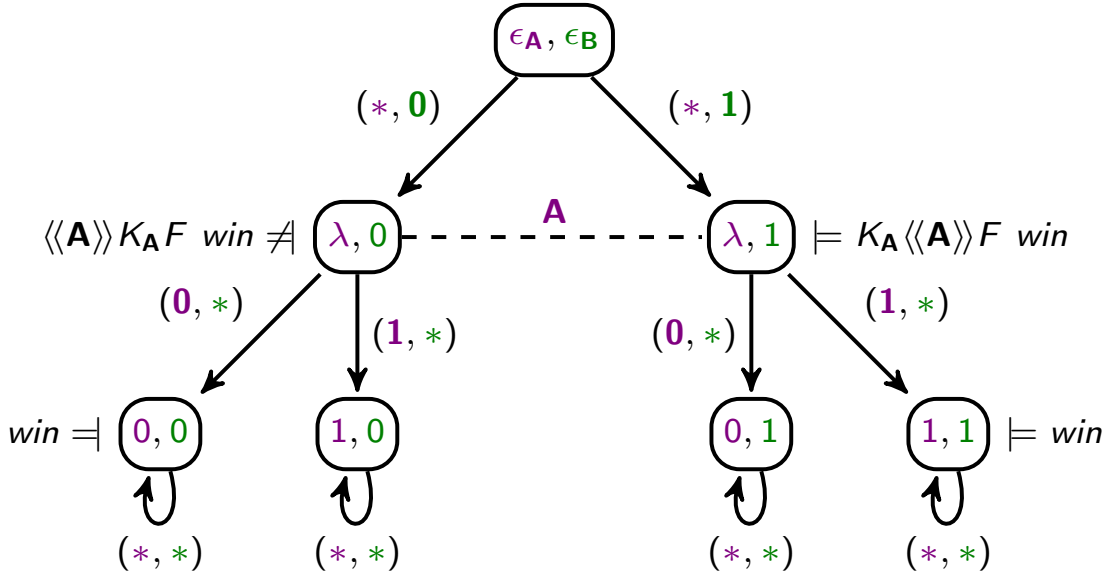


It looks like it's a question of knowledge

- **A**nne knows that there is some strategy to win (knowledge *de dicto*)

The Problem

ATL with Imperfect Information

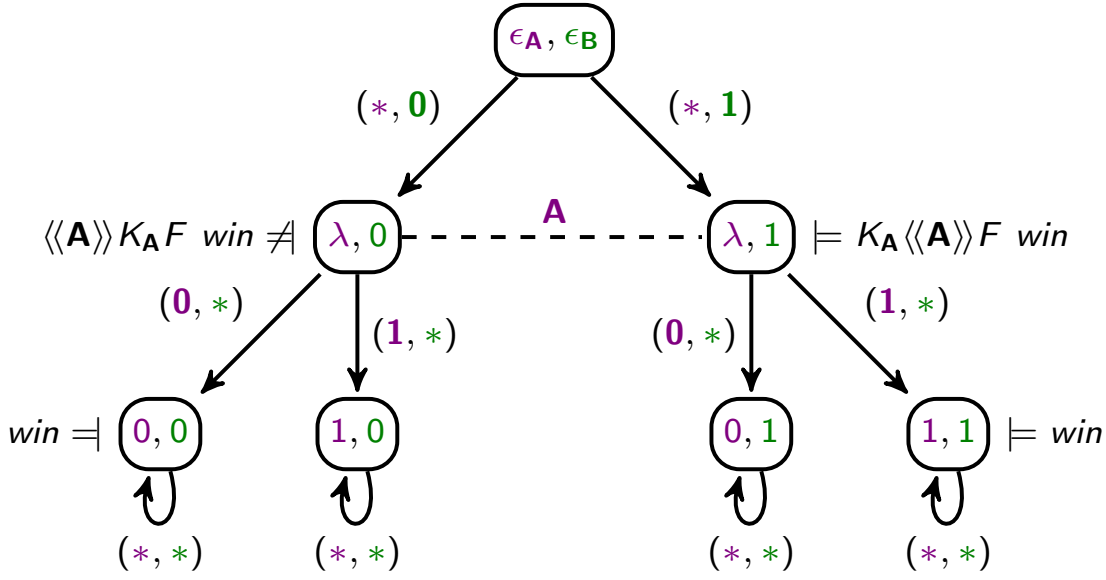


It looks like it's a question of knowledge

- **A**nne knows that there is some strategy to win (knowledge *de dicto*)
- but there is no strategy known to her to guarantee a win (knowledge *de re*)

The Problem

ATL with Imperfect Information



It looks like it's a question of knowledge

- **A**nne knows that there is some strategy to win (knowledge *de dicto*)
- but there is no strategy known to her to guarantee a win (knowledge *de re*)

Is there any way of combining ATL and epistemic operators so as to obtain something similar to (1)-(3)?

References

Rajeev Alur, Thomas A. Henzinger, and Orna Kupferman.

Alternating-time temporal logic.
J. ACM, 49(5):672–713, 2002.

Nils Bulling, Jurgen Dix, and Wojciech Jamroga.

Model checking logics of strategic ability: Complexity*.
In Mehdi Dastani, Koen V. Hindriks, and John-Jules Charles Meyer, editors, *Specification and Verification of Multi-agent Systems*, pages 125–159. Springer US, 2010.

V. Goranko and D. Shkatov.

Tableau-based decision procedures for logics of strategic ability in multiagent systems.
ACM Trans. Comput. Log., 11(1), 2009.