# Comparing Variants of Strategic Ability

Wojtek Jamroga, University of Luxembourg
(joint work with Nils Bulling, Clausthal University of Technology)

March 1, 2012, Dynamics in Logic II

# LAMAS: Logical Aspects of Multi-Agent Systems

**LAMAS 2012:** 5th of June 2012 at AAMAS2012 (Valencia, Spain)

**http://icr.uni.lu/lamas2012/**

# Outline

# Outline

# Introduction

- Strategic logics: ATL, coalition logic, stit
- Basic issue: "Can agent $a$ (coalition $A$) bring about $\varphi$?"
- Semantic variants of ATL: encapsulate various notions of ability

# Introduction

- Strategic logics: ATL, coalition logic, stit
- Basic issue: "Can agent $a$ (coalition $A$) bring about $\varphi$?"
- Semantic variants of ATL: encapsulate various notions of ability
- **We study the relationship between standard variants of ATL on the level of valid sentences**
- Surprisingly, nobody has studied it before

# Motivation

- "Hardcore" logicians: logic = set of validities
- Thus, by comparing validity sets we compare logics in the traditional sense

# Motivation

- "Hardcore" logicians: logic = set of validities
- Thus, by comparing validity sets we compare logics in the traditional sense
- Validities capture general properties of games under consideration
- If two variants of ATL generate the same valid sentences then the underlying notions of ability induce the same kind of games

# Motivation

- "Hardcore" logicians: logic = set of validities
- Thus, by comparing validity sets we compare logics in the traditional sense
- Validities capture general properties of games under consideration
- If two variants of ATL generate the same valid sentences then the underlying notions of ability induce the same kind of games
- First step towards devising algorithms for satisfiability checking

# Outline

# ATL: What Agents Can Achieve

- ATL: Alternating-time Temporal Logic [Alur et al. 1997-2002]
- Temporal logic meets game theory
- Main idea: cooperation modalities

# ATL: What Agents Can Achieve

- ATL: Alternating-time Temporal Logic [Alur et al. 1997-2002]
- Temporal logic meets game theory
- Main idea: cooperation modalities

$\langle\langle A \rangle\rangle \Phi$: coalition $A$ has a collective strategy to enforce $\Phi$

# ATL: What Agents Can Achieve

- ATL: Alternating-time Temporal Logic [Alur et al. 1997-2002]
- Temporal logic meets game theory
- Main idea: cooperation modalities

**$\langle\langle A \rangle\rangle \Phi$: coalition $A$ has a collective strategy to enforce $\Phi$**

$\rightsquigarrow$ $\Phi$ can include temporal operators: $\bigcirc$ (next), $\Diamond$ (sometime in the future), $\Box$ (always in the future), $\mathcal{U}$ (strong until)

# ATL: What Agents Can Achieve

Example formulae:

- $\langle\langle robber \rangle\rangle \Diamond$open:
  "The robber has a strategy to eventually get the vault open no matter how the other agents act"

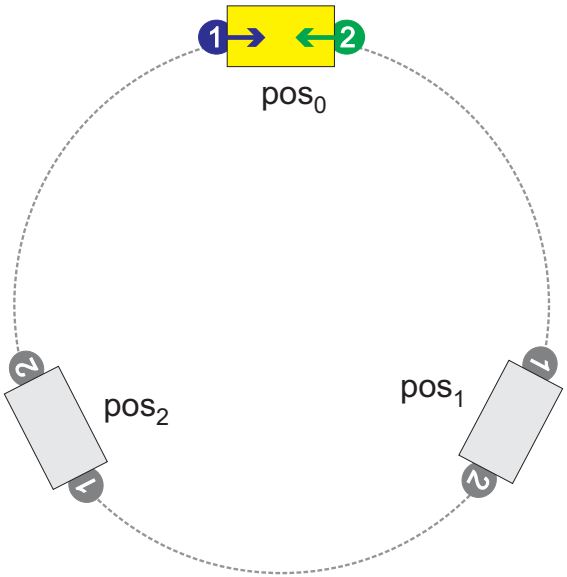
# ATL: What Agents Can Achieve

Example formulae:

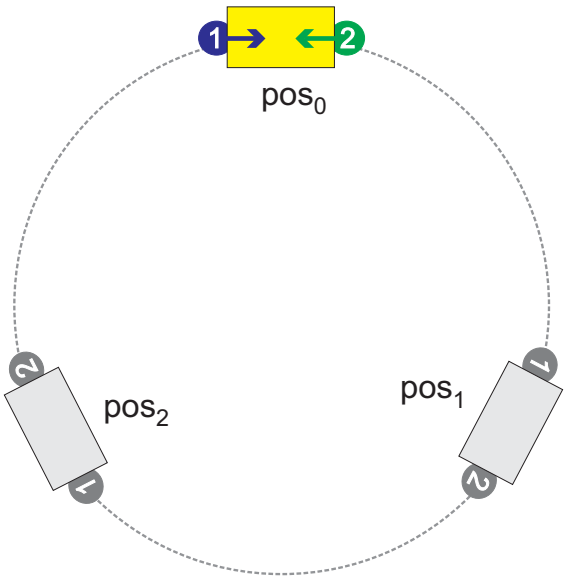
- $\langle\langle robber \rangle\rangle \Diamond$open:
  "The robber has a strategy to eventually get the vault open no matter how the other agents act"

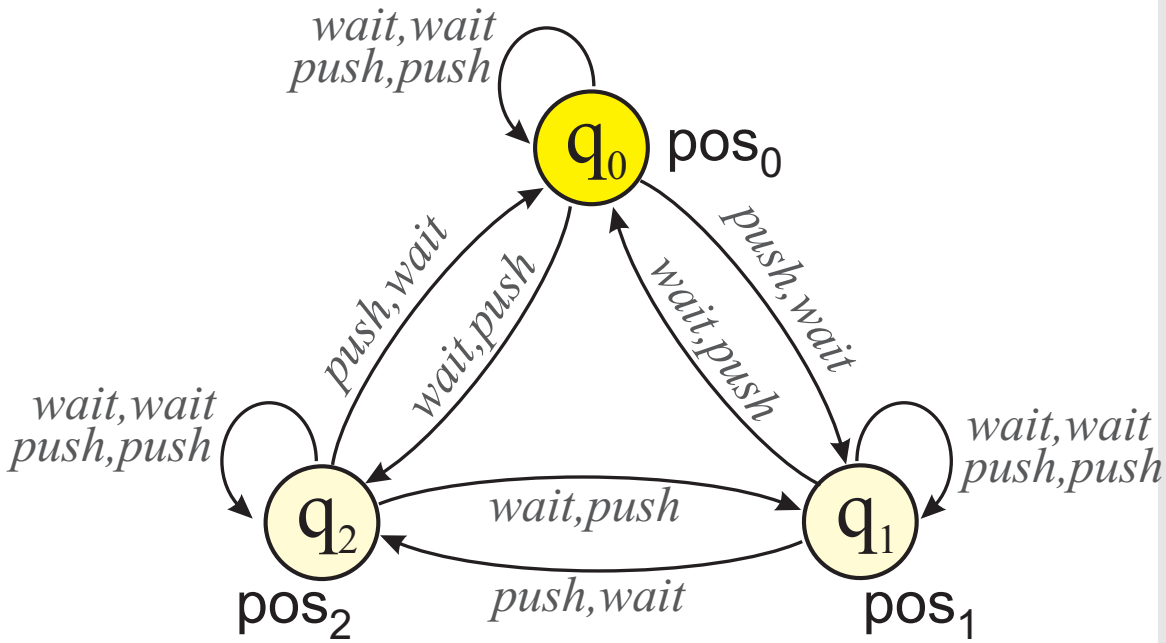
- $\langle\langle bank \rangle\rangle \Box \neg$robbery:
  "The bank can protect itself against being robbed"

# Example: Robots and Carriage



# Example: Robots and Carriage



# Strategies and Abilities

## Definition (Strategy)

A **strategy** is a conditional plan.

# Strategies and Abilities

## Definition (Strategy)

A **strategy** is a conditional plan.
We represent strategies by functions $s_a : St \to Act$.

# Strategies and Abilities

## Definition (Strategy)

A **strategy** is a conditional plan.
We represent strategies by functions $s_a : St \to Act$.

$\rightsquigarrow$ memoryless strategies

# Strategies and Abilities

## Definition (Strategy)

A **strategy** is a conditional plan.
We represent strategies by functions $s_a : St \to Act$.

$\rightsquigarrow$ memoryless strategies

Alternative: perfect recall strategies $s_a : St^+ \to Act$

# Strategies and Abilities

## Definition (Strategy)

A **strategy** is a conditional plan.
We represent strategies by functions $s_a : St \to Act$.

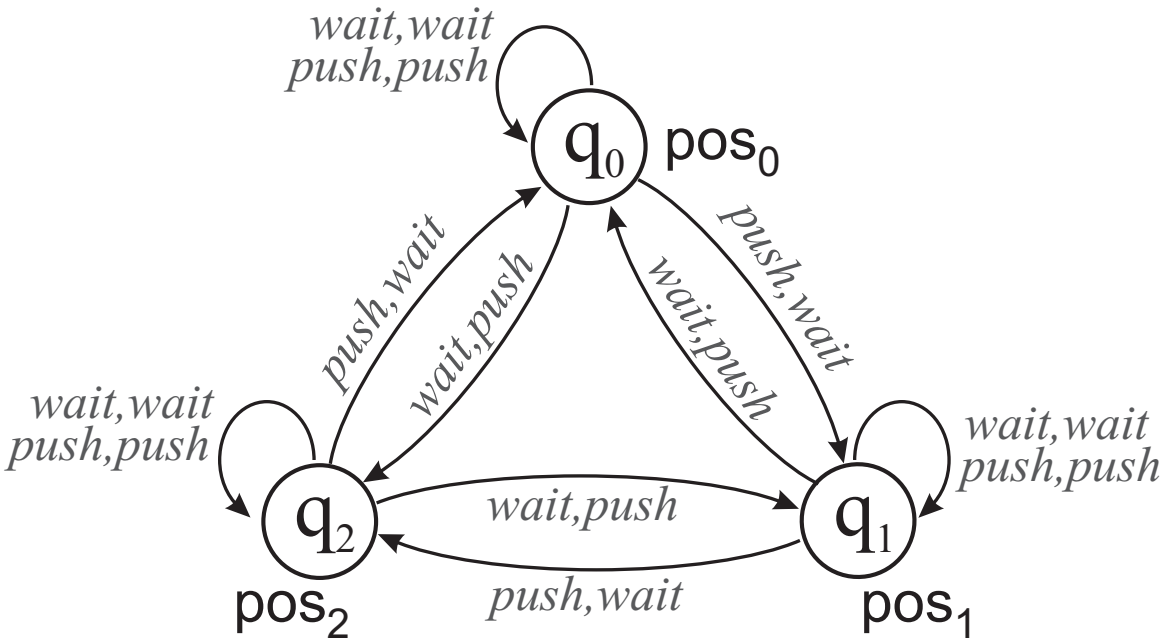
$\rightsquigarrow$ memoryless strategies

Alternative: perfect recall strategies $s_a : St^+ \to Act$

## Semantics of ATL

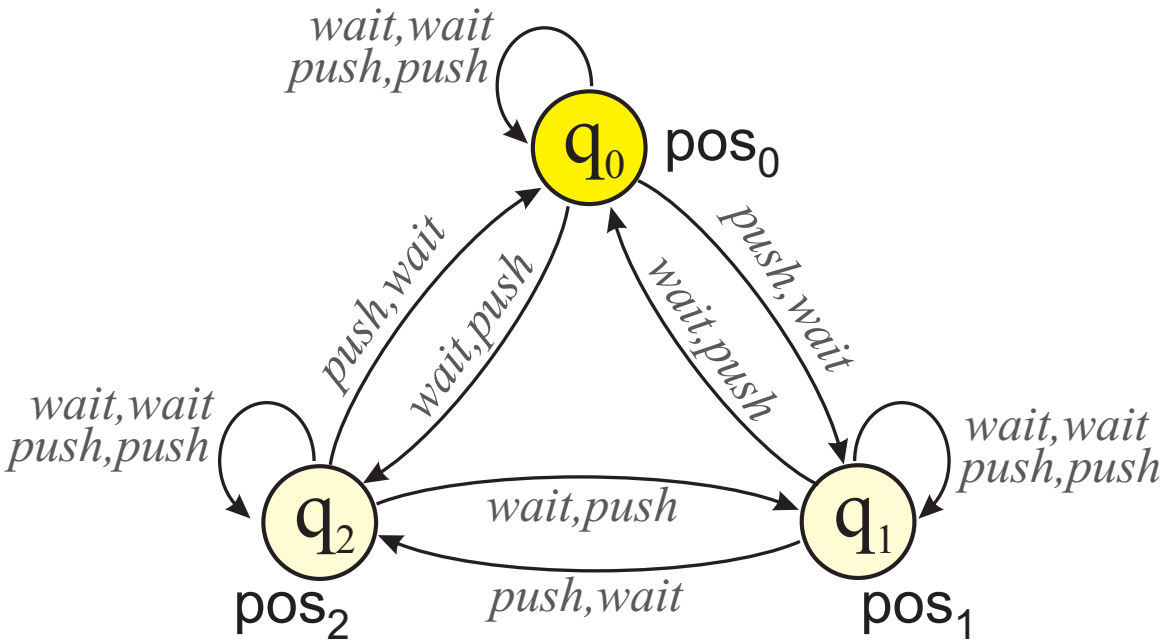
$M, q \models \langle\langle A \rangle\rangle \Phi$ iff there is a collective strategy $s_A$ such that, for every path $\lambda$ that may result from execution of $s_A$ from $q$ on, we have that $M, \lambda \models \Phi$.

# Example: Robots and Carriage



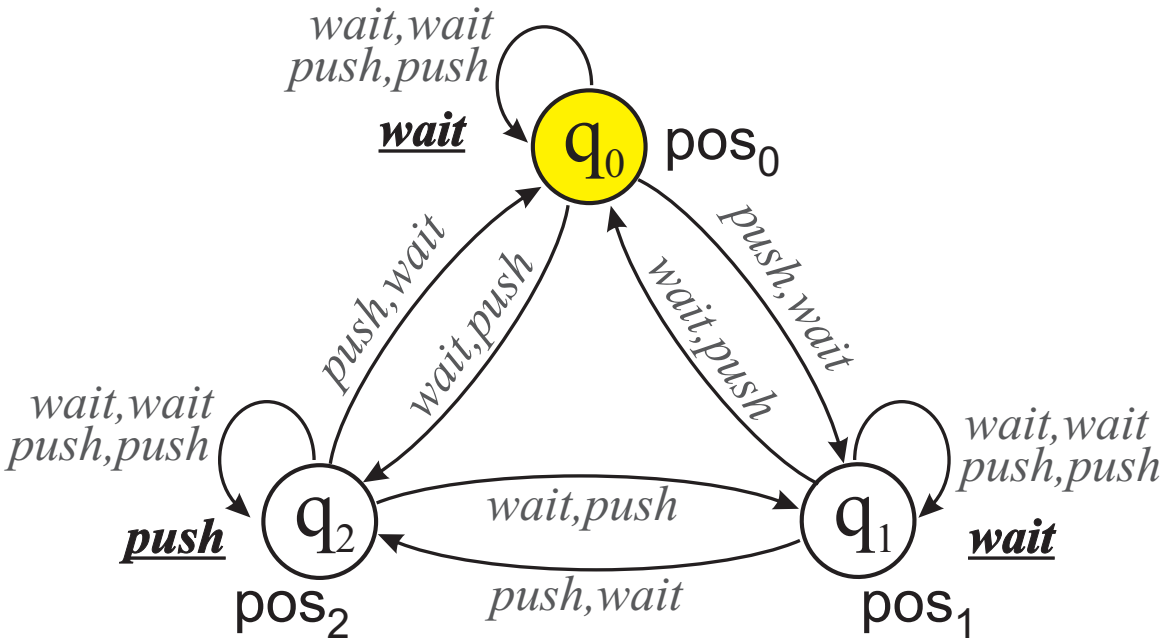

$$\mathsf{pos}_0 \rightarrow \langle\langle 1 \rangle\rangle \Box \neg \mathsf{pos}_1$$

# Example: Robots and Carriage



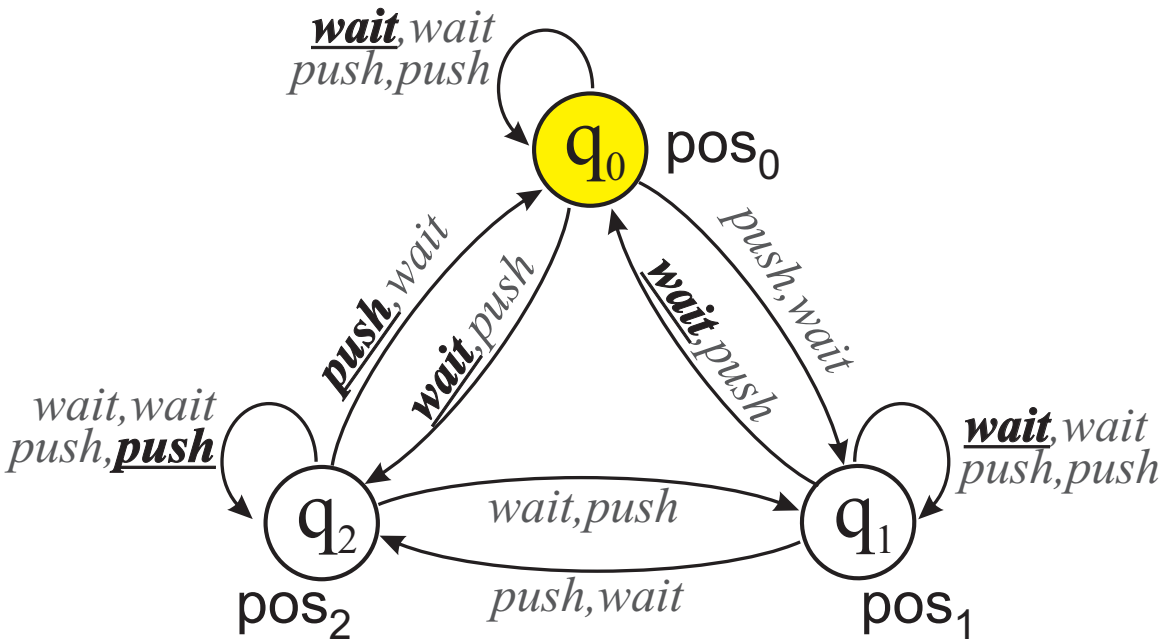

$$\mathsf{pos}_0 \rightarrow \langle\langle 1 \rangle\rangle \Box \neg \mathsf{pos}_1$$

# Example: Robots and Carriage



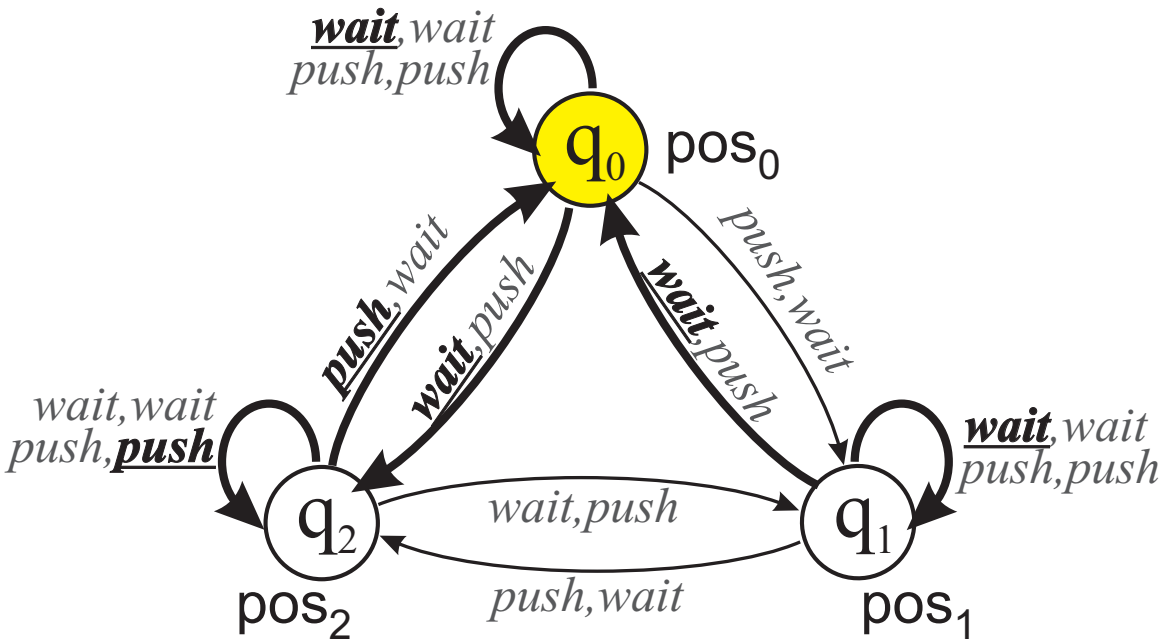

$$pos_0 \rightarrow \langle\langle 1 \rangle\rangle \Box \neg pos_1$$

# Example: Robots and Carriage



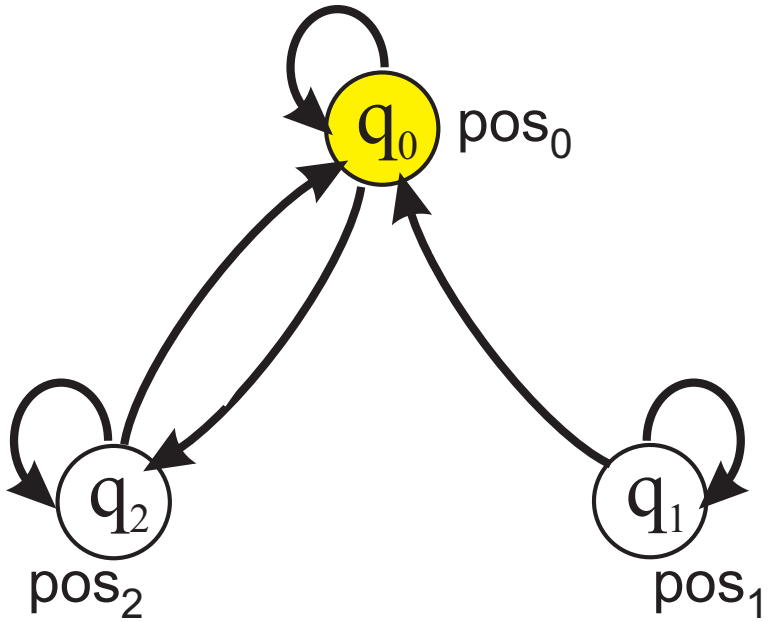

$\mathsf{pos}_0 \rightarrow \langle\langle 1 \rangle\rangle \Box \neg \mathsf{pos}_1$

# Example: Robots and Carriage



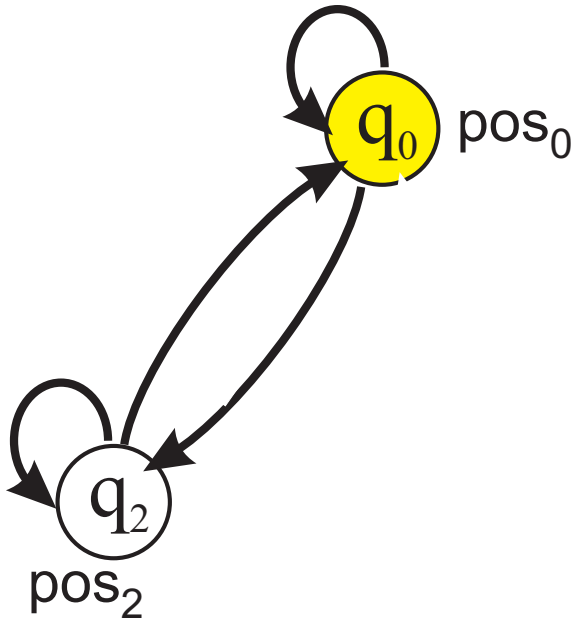

$$\mathrm{pos}_0 \rightarrow \langle\langle 1 \rangle\rangle \Box \neg \mathrm{pos}_1$$

# Example: Robots and Carriage




$$pos_0 \rightarrow \langle\langle 1 \rangle\rangle \Box \neg pos_1$$

# Example: Robots and Carriage




$$pos_0 \rightarrow \langle\langle 1 \rangle\rangle \Box \neg pos_1$$

# Semantic Variants of ATL

- Basic semantics of ATL assumes perfect information - not very realistic
- Semantic variants for more realistic cases defined in (Jamroga 2003), (Jonker 2003), (Schobbens 2004), (Jamroga & van der Hoek 2004), (Agotnes 2004), ...
- Encapsulate different assumptions about agents and abilities

# Semantic Variants of ATL

Memory of agents:

- Perfect recall **(R)** vs. imperfect recall strategies **(r)**

# Semantic Variants of ATL

Memory of agents:

- Perfect recall **(R)** vs. imperfect recall strategies **(r)**

Available information:

- Perfect information **(I)** vs. imperfect information strategies **(i)**

# Semantic Variants of ATL

Memory of agents:

- Perfect recall **(R)** vs. imperfect recall strategies **(r)**

Available information:

- Perfect information **(I)** vs. imperfect information strategies **(i)**

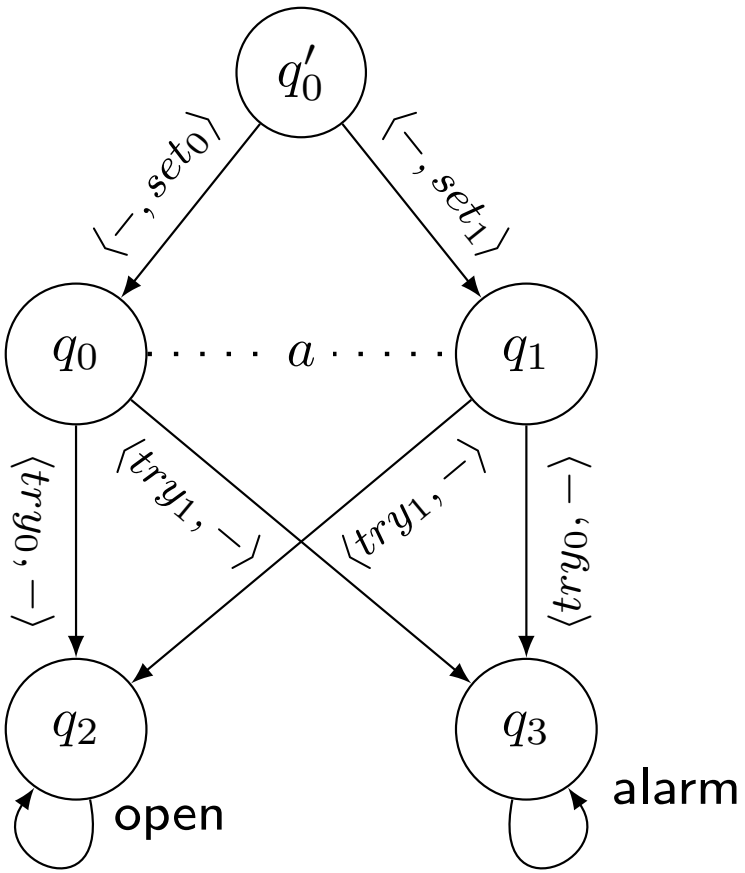
Success of strategies:

- Objectively **($i_o$)** vs. subjectively successful strategies **($i_s$)**

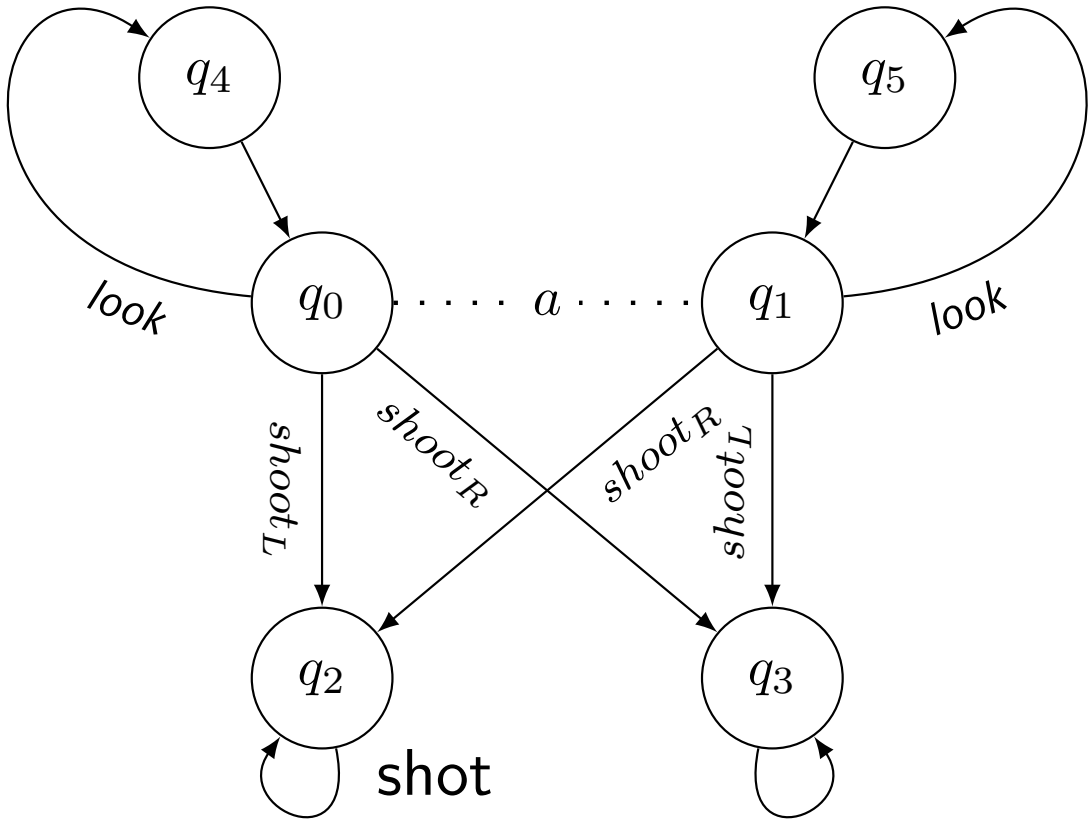
# Example: Robbing a Bank

# Example: Robbing a Bank

# Example: Robbing a Bank



# Example: Poor Duck Problem



# Connection to Game Theory

- $M, q \models \langle\langle A \rangle\rangle \gamma \approx$ extensive game

# Connection to Game Theory

- $M, q \models \langle\langle A \rangle\rangle \gamma \approx$ extensive game

- $A$ splits agents into proponents and opponents
- $\models$ and $\gamma$ define the winning condition
  $\rightsquigarrow$ infinite 2-player zero-sum game with binary payoffs

# Connection to Game Theory

- $M, q \models \langle\langle A \rangle\rangle \gamma \approx$ extensive game
- $A$ splits agents into proponents and opponents
- $\models$ and $\gamma$ define the winning condition
  $\rightsquigarrow$ infinite 2-player zero-sum game with binary payoffs
- Model checking $\rightleftharpoons$ game solving

# Connection to Game Theory

- $M, q \models \langle\langle A \rangle\rangle \gamma \approx$ extensive game

- $A$ splits agents into proponents and opponents
- $\models$ and $\gamma$ define the winning condition
  $\rightsquigarrow$ infinite 2-player zero-sum game with binary payoffs

- Model checking $\rightleftharpoons$ game solving
- Satisfiability $\rightleftharpoons$ mechanism design

# Connection to Game Theory

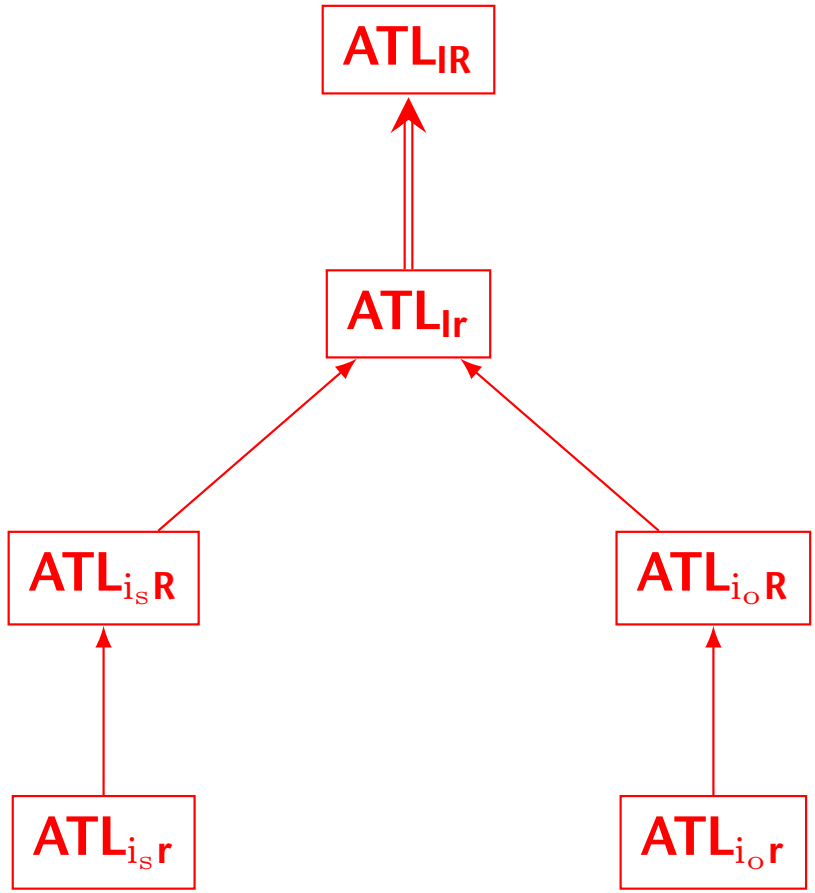
- $M, q \models \langle\langle A \rangle\rangle \gamma \approx$ extensive game

- $A$ splits agents into proponents and opponents
- $\models$ and $\gamma$ define the winning condition
  $\rightsquigarrow$ infinite 2-player zero-sum game with binary payoffs

- Model checking $\rightleftharpoons$ game solving
- Satisfiability $\rightleftharpoons$ mechanism design
- Validity $\rightleftharpoons$ general properties of games

# Outline

# Validities in Variants of ATL: Subsumption Graph



# Summary in Plain Words

In terms of general properties of games we get the following:

- Perfect information is a (strict) special case of imperfect information
- Perfect recall games are (strict) special case of memoryless games
- Information type has more impact than type of recall
- Properties of objective and subjective abilities of agents are incomparable

# Outline

# Some (In)Validities

- $\langle\langle a \rangle\rangle \Diamond p \leftrightarrow p \vee \langle\langle a \rangle\rangle \bigcirc \langle\langle a \rangle\rangle \Diamond p$
  Invalid in all variants with imperfect information

# Some (In)Validities

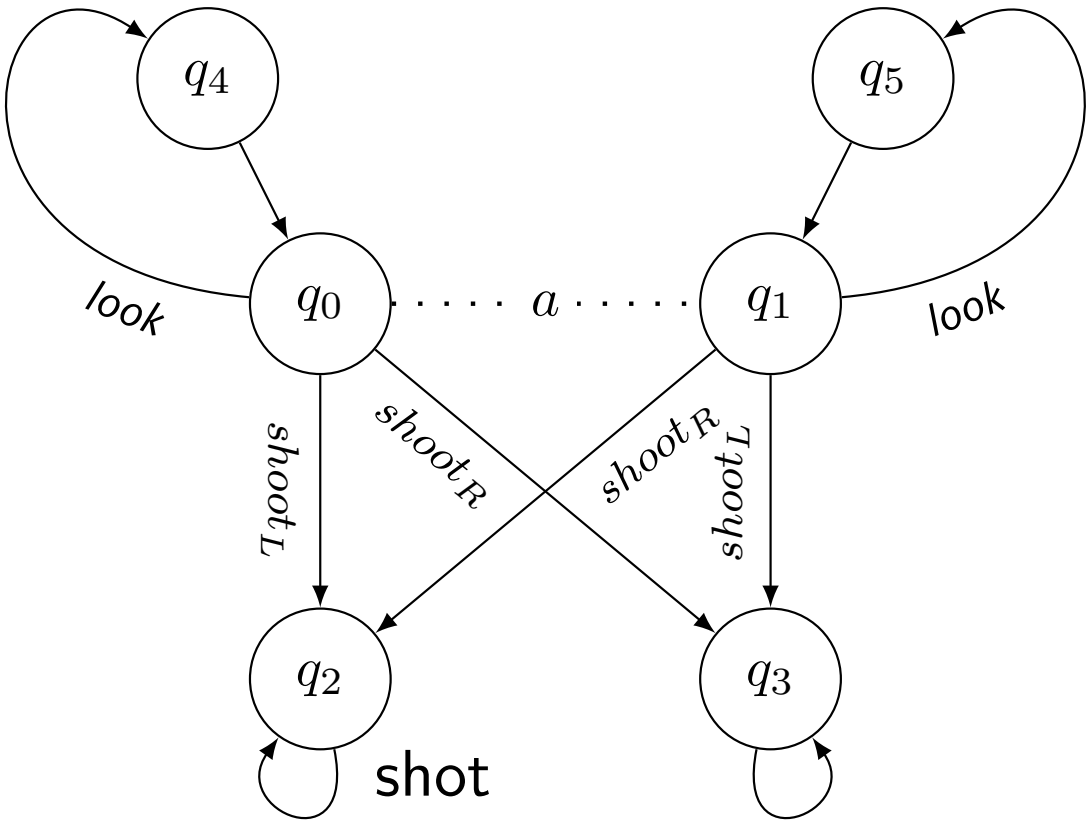
- $\langle\langle a \rangle\rangle \Diamond p \leftrightarrow p \vee \langle\langle a \rangle\rangle \bigcirc \langle\langle a \rangle\rangle \Diamond p$
  Invalid in all variants with imperfect information
- $\langle\langle a \rangle\rangle (\Diamond p_1 \wedge \Diamond p_2) \leftrightarrow \langle\langle a \rangle\rangle \Diamond (p_1 \wedge \langle\langle a \rangle\rangle \Diamond p_2 \vee p_2 \wedge \langle\langle a \rangle\rangle \Diamond p_1)$
  Invalid for imperfect recall

# Some (In)Validities

- $\langle\langle a \rangle\rangle \Diamond p \leftrightarrow p \vee \langle\langle a \rangle\rangle \bigcirc \langle\langle a \rangle\rangle \Diamond p$
  Invalid in all variants with imperfect information
- $\langle\langle a \rangle\rangle (\Diamond p_1 \wedge \Diamond p_2) \leftrightarrow \langle\langle a \rangle\rangle \Diamond (p_1 \wedge \langle\langle a \rangle\rangle \Diamond p_2 \vee p_2 \wedge \langle\langle a \rangle\rangle \Diamond p_1)$
  Invalid for imperfect recall
- $\neg\langle\langle \emptyset \rangle\rangle \Diamond \neg p \leftrightarrow \langle\langle \mathbb{A}\mathrm{gt} \rangle\rangle \Box p$
  Invalid for subjective ability

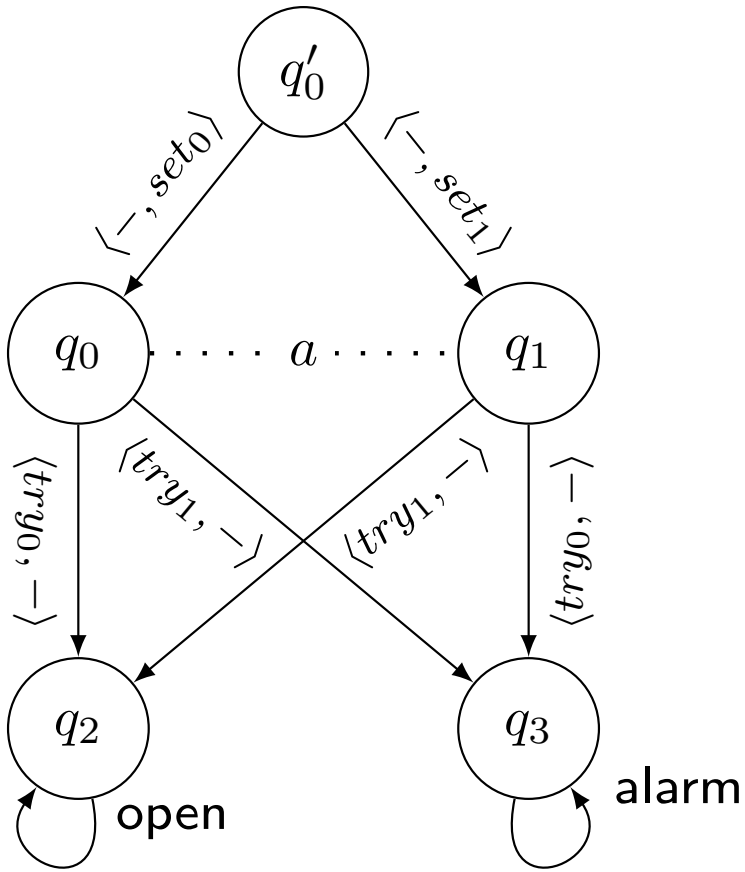
# Some (In)Validities

- $\langle\langle a \rangle\rangle \Diamond \mathsf{p} \leftrightarrow \mathsf{p} \vee \langle\langle a \rangle\rangle \bigcirc \langle\langle a \rangle\rangle \Diamond \mathsf{p}$
  Invalid in all variants with imperfect information



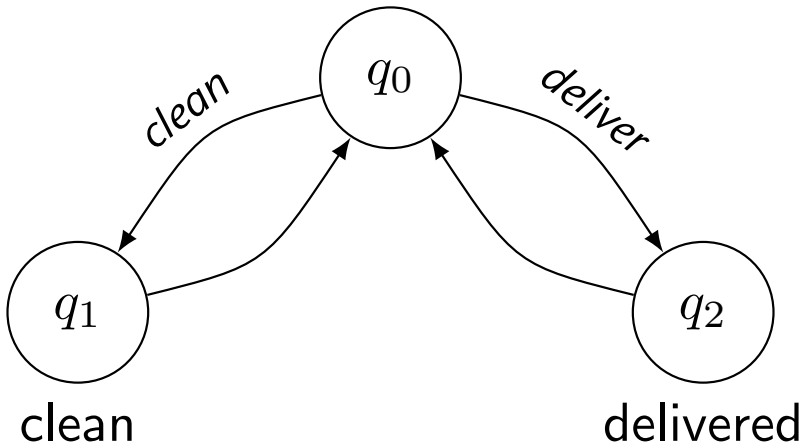
# Some (In)Validities

- $\langle\langle a \rangle\rangle \Diamond \mathsf{p} \leftrightarrow \mathsf{p} \vee \langle\langle a \rangle\rangle \bigcirc \langle\langle a \rangle\rangle \Diamond \mathsf{p}$
  Invalid in all variants with imperfect information



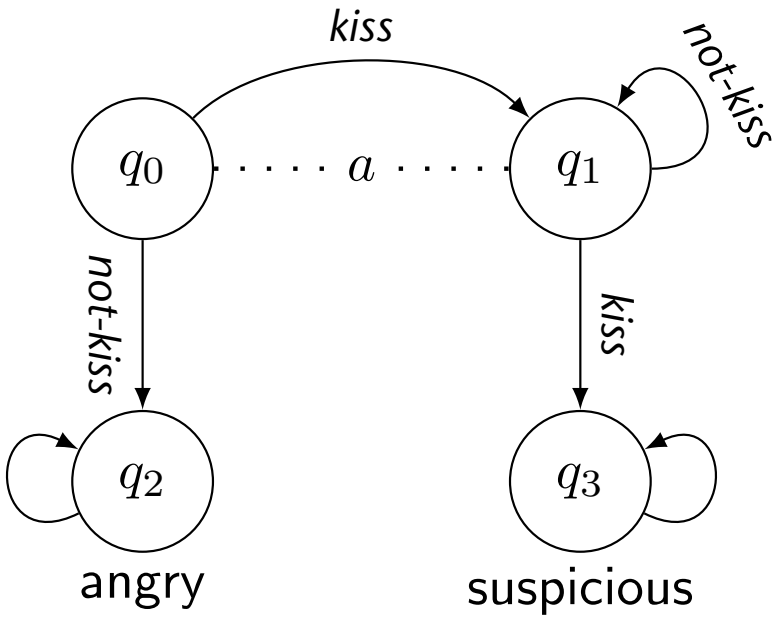
# Some (In)Validities

- $\langle\langle a \rangle\rangle(\Diamond p_1 \wedge \Diamond p_2) \leftrightarrow \langle\langle a \rangle\rangle\Diamond(p_1 \wedge \langle\langle a \rangle\rangle\Diamond p_2 \vee p_2 \wedge \langle\langle a \rangle\rangle\Diamond p_1)$
  Invalid for imperfect recall



# Some (In)Validities

- $\neg\langle\langle\emptyset\rangle\rangle\Diamond\neg p \leftrightarrow \langle\langle \mathbb{A}gt\rangle\rangle\Box p$
  Invalid for subjective ability



# Outline

# Conclusions

- All the basic semantic variants of ATL **are different** on the level of general properties they induce
- Strong pattern of subsumption
- Very natural when you see it, but by no means obvious before
- Some proofs nontrivial

# Conclusions

- All the basic semantic variants of ATL **are different** on the level of general properties they induce
- Strong pattern of subsumption
- Very natural when you see it, but by no means obvious before
- Some proofs nontrivial

- **Non-validities more important than the inclusion results**