HØGSKOLEN I BERGEN

TU Clausthal

Modal Logics for Games and Multi-Agent Systems

Thomas Ågotnes and Wojtek Jamroga
(Bergen University College / TU Clausthal)

Time and place:

11–15 August 2008, hours 9.15-10.45

Organization:	Parts 2, 4, 7, 9 & 10:	Thomas Ågotnes,
	Parts 1, 3, 5, 6 & 8:	Wojtek Jamroga

Lecture Overview I

1. **Introduction.** Multi-agent systems. Modal logic. Epistemic logic. Axioms and systems of modal logic. Correspondence theory.
2. **Coalition logic.** Strategic games and coalition logic (CL). Axiomatisation of CL.
3. **ATL.** Multi-step games and alternating-time temporal logic (ATL).
4. **More about ATL.** Axiomatisation; bisimulation; the role of memory; revocability of strategies.

Lecture Overview II

5. **Strategic reasoning for imperfect information (part I).** Strategic reasoning for imperfect information scenarios. Problems with ATEL. Economic solution: ATLir.
6. **Strategic reasoning for imperfect information (part II).** General solution: CSL. Properties of constructive knowledge. Semantics for constructive normal form.
7. **Characterising solution concepts.** Non-cooperative games: characterising solution concepts in modal logic.

Lecture Overview III

8. **Reasoning about rational play.** Reasoning about rational play in ATLP. Temporalized solution concepts.
9. **Axiomatisation of coalitional games (part I).** Cooperative games. Examples of games. Solution concepts. Axiomatisation in modal logic.
10. **Axiomatisation of coalitional games (part II).** Axiomatisation of coalitional games: completeness proof.

Basic Reading I

[Alur *et al.*2002] R. Alur, T. A. Henzinger, and O. Kupferman.
Alternating-time Temporal Logic.
Journal of the ACM, 49:672–713, 2002.

[van der Hoek and Pauly2006] W. van der Hoek and M. Pauly.
Modal logic for games and information.
In P. Blackburn, J. van Benthem, and F. Wolter, editors, *Handbook of Modal Logic*, pages 1077–1148. Elsevier Science Publishers B.V.: Amsterdam, The Netherlands, 2006.

Basic Reading II

[Jamroga and Ågotnes2007] W. Jamroga and T. Ågotnes.
Constructive knowledge: What agents can achieve under incomplete information.
Journal of Applied Non-Classical Logics, 17(4):423–475, 2007.

[Ågotnes *et al.*2008] Thomas Ågotnes, Wiebe van der Hoek, and Michael Wooldridge.
Reasoning about coalitional games, 2008.
Manuscript, to appear.

Agent Systems and Modal Logic

1.1 Agents

- **Multi-agent system (MAS):** a system that involves several autonomous entities that act in the same environment
- The entities are called **agents**

- **Multi-agent system (MAS):** a system that involves several autonomous entities that act in the same environment
- The entities are called **agents**
- So, what is an agent precisely?

- **Multi-agent system (MAS):** a system that involves several autonomous entities that act in the same environment
- The entities are called **agents**

- So, what is an agent precisely?
- No commonly accepted definition

For some authors, agents are:

- A new paradigm for computation

For some authors, agents are:

- A new paradigm for computation
- A new paradigm for design

For some authors, agents are:

- A new paradigm for computation
- A new paradigm for design
- A new paradigm for programming

For some authors, agents are:

- A new paradigm for computation
- A new paradigm for design
- A new paradigm for programming

Our claim:

MAS is a philosophical metaphor that induces a specific way of seeing the world.

Intuition:

We are agents!

Intuition:

We are agents!

The metaphor:

- Makes us use specific vocabulary

Intuition:

We are agents!

The metaphor:

- Makes us use specific vocabulary
- Makes us use specific conceptual structures

Intuition:

We are agents!

The metaphor:

- Makes us use specific vocabulary
- Makes us use specific conceptual structures
- So:
- A new paradigm for thinking and talking about the world

Features of agents

An agent can/should possibly be:

Features of agents

An agent can/should possibly be:

- Autonomous: operates without direct intervention of others, has some kind of control over its actions and internal state

Features of agents

An agent can/should possibly be:

- Autonomous: operates without direct intervention of others, has some kind of control over its actions and internal state
- Reactive: reacts to changes in the environment

Features of agents

An agent can/should possibly be:

- Autonomous: operates without direct intervention of others, has some kind of control over its actions and internal state
- Reactive: reacts to changes in the environment
- Pro-active: takes the initiative

Features of agents

An agent can/should possibly be:

- Autonomous: operates without direct intervention of others, has some kind of control over its actions and internal state
- Reactive: reacts to changes in the environment
- Pro-active: takes the initiative
- Goal-directed: acts to achieve a goal

Features of agents

An agent can/should possibly be:

- Autonomous: operates without direct intervention of others, has some kind of control over its actions and internal state
- Reactive: reacts to changes in the environment
- Pro-active: takes the initiative
- Goal-directed: acts to achieve a goal
- Social: interacts with others (cooperation, communication, coordination, competition)

Features of agents

- Embodied: has sensors and effectors to read from and make changes to the environment

Features of agents

- Embodied: has sensors and effectors to read from and make changes to the environment
- Intelligent:

Features of agents

- Embodied: has sensors and effectors to read from and make changes to the environment
- Intelligent: ...whatever it means

Features of agents

- Embodied: has sensors and effectors to read from and make changes to the environment
- Intelligent: ...whatever it means
- Rational: always does the right thing

Is there any essential (and commonly accepted) feature of an agent?

Is there any essential (and commonly accepted) feature of an agent?

An agent acts.

Is there any essential (and commonly accepted) feature of an agent?

An agent acts.

Agents can be described mathematically by a function

$$act : \text{set of percept sequences} \mapsto \text{set of actions}$$

Is there any essential (and commonly accepted) feature of an agent?

An agent acts.

Agents can be described mathematically by a function

$$act : \text{set of percept sequences} \mapsto \text{set of actions}$$

Note that, in game theory, such a function is called a strategy.

Is there any essential (and commonly accepted) feature of an agent?

An agent acts.

Agents can be described mathematically by a function

$$act : \text{set of percept sequences} \mapsto \text{set of actions}$$

Note that, in game theory, such a function is called a strategy.

In planning, it is called a conditional plan.

1.2 Modal Logic

Modal logic

Modal logic is an extension of classical logic by new connectives $\Box$ and $\Diamond$: necessity and possibility.

Modal logic

Modal logic is an extension of classical logic by new connectives $\Box$ and $\Diamond$: necessity and possibility.

- $\Box\varphi$ means that φ is necessarily true

Modal logic

Modal logic is an extension of classical logic by new connectives $\Box$ and $\Diamond$: necessity and possibility.

- $\Box\varphi$ means that φ is necessarily true
- $\Diamond\varphi$ means that φ is possibly true

Modal logic

Modal logic is an extension of classical logic by new connectives $\Box$ and $\Diamond$: necessity and possibility.

- $\Box\varphi$ means that φ is necessarily true
- $\Diamond\varphi$ means that φ is possibly true

Independently of the precise definition, the following holds:

$$\Diamond\varphi \leftrightarrow \neg\Box\neg\varphi$$

More precisely, necessity/possibility is interpreted as follows:

- p is necessary $\Leftrightarrow$ p is true in all possible scenarios
- p is possible $\Leftrightarrow$ p is true in at least one possible scenario

More precisely, necessity/possibility is interpreted as follows:

- p is necessary $\Leftrightarrow$ p is true in all possible scenarios
- p is possible $\Leftrightarrow$ p is true in at least one possible scenario

$\leadsto$ possible worlds semantics

Definition 1.1 (Kripke structure)

A Kripke structure is a tuple $S = \langle \mathcal{W}, \mathcal{R} \rangle$, where $\mathcal{W}$ is a set of possible worlds, and $\mathcal{R}$ is a binary relation on worlds, called accessibility relation.

For multiple modalities $\Box_1, \Diamond_1, \ldots, \Box_k, \Diamond_k$, we use a family of relations $\mathcal{R}_1, \ldots, \mathcal{R}_k$

Definition 1.1 (Kripke structure)

A Kripke structure is a tuple $S = \langle \mathcal{W}, \mathcal{R} \rangle$, where $\mathcal{W}$ is a set of possible worlds, and $\mathcal{R}$ is a binary relation on worlds, called accessibility relation.

For multiple modalities $\Box_1, \Diamond_1, \ldots, \Box_k, \Diamond_k$, we use a family of relations $\mathcal{R}_1, \ldots, \mathcal{R}_k$

Definition 1.2 (Kripke model)

Let Π be a set of atomic propositions $(p, q, r, \ldots)$. A possible worlds model $M = \langle \mathcal{S}, \pi \rangle$ consists of a Kripke structure S, and a valuation of propositions $\pi : \mathcal{W} \to 2^{\Pi}$.

Definition 1.3 (Semantics of modal logic)

The truth of formulae is relative to a Kripke model $M = \langle \mathcal{W}, \mathcal{R}, \pi \rangle$, and a world $w \in \mathcal{W}$. It can be defined through the following clauses:

Definition 1.3 (Semantics of modal logic)

The truth of formulae is relative to a Kripke model $M = \langle \mathcal{W}, \mathcal{R}, \pi \rangle$, and a world $w \in \mathcal{W}$. It can be defined through the following clauses:

- $M, w \models p$ iff $p \in \pi(w)$;

Definition 1.3 (Semantics of modal logic)

The truth of formulae is relative to a Kripke model $M = \langle \mathcal{W}, \mathcal{R}, \pi \rangle$, and a world $w \in \mathcal{W}$. It can be defined through the following clauses:

- $M, w \models p$ iff $p \in \pi(w)$;
- $M, w \models \neg\varphi$ iff not $M, w \models \varphi$;

Definition 1.3 (Semantics of modal logic)

The truth of formulae is relative to a Kripke model $M = \langle \mathcal{W}, \mathcal{R}, \pi \rangle$, and a world $w \in \mathcal{W}$. It can be defined through the following clauses:

- $M, w \models p$ iff $p \in \pi(w)$;
- $M, w \models \neg\varphi$ iff not $M, w \models \varphi$;
- $M, w \models \varphi \wedge \psi$ iff $M, w \models \varphi$ and $M, w \models \psi$;

Definition 1.3 (Semantics of modal logic)

The truth of formulae is relative to a Kripke model $M = \langle \mathcal{W}, \mathcal{R}, \pi \rangle$, and a world $w \in \mathcal{W}$. It can be defined through the following clauses:

- $M, w \models p$ iff $p \in \pi(w)$;
- $M, w \models \neg\varphi$ iff not $M, w \models \varphi$;
- $M, w \models \varphi \wedge \psi$ iff $M, w \models \varphi$ and $M, w \models \psi$;
- $M, w \models \Box\varphi$ iff, for every $w' \in \mathcal{W}$ such that $w\mathcal{R}w'$, we have $M, w' \models \varphi$.

Logic for Agents

Modal logic is a generic framework.

Logic for Agents

Modal logic is a generic framework.

Various modal logics:

- knowledge $\rightsquigarrow$ epistemic logic,
- beliefs $\rightsquigarrow$ doxastic logic,
- obligations $\rightsquigarrow$ deontic logic,
- actions $\rightsquigarrow$ dynamic logic,
- time $\rightsquigarrow$ temporal logic,
- ability $\rightsquigarrow$ strategic logic,
- and combinations of the above

Logic for Agents

Modal logic is a generic framework.

Various modal logics:

- knowledge $\rightsquigarrow$ epistemic logic,
- beliefs $\rightsquigarrow$ doxastic logic,
- obligations $\rightsquigarrow$ deontic logic,
- actions $\rightsquigarrow$ dynamic logic,
- time $\rightsquigarrow$ temporal logic,
- ability $\rightsquigarrow$ strategic logic,
- and combinations of the above

Modal logic seems very well suited for reasoning about various dimensions of multi-agent systems!

1.3 Epistemic Logic

Epistemic logic

- We interpret $\Box_i\varphi$ as "agent i knows that φ"
- $\Box_i$ is usually written as K_i in epistemic logic

Epistemic logic

- We interpret $\Box_i\varphi$ as "agent i knows that φ"
- $\Box_i$ is usually written as K_i in epistemic logic

- Possible worlds: states of the system, situations
- Modal relations $\mathcal{R}_i$: indistinguishability of states for agent i
- We assume that $\mathcal{R}_i$ are equivalence relations

Epistemic logic

- We interpret $\Box_i\varphi$ as "agent i knows that φ"
- $\Box_i$ is usually written as K_i in epistemic logic

- Possible worlds: states of the system, situations
- Modal relations $\sim_i$: indistinguishability of states for agent i
- We assume that $\sim_i$ are equivalence relations

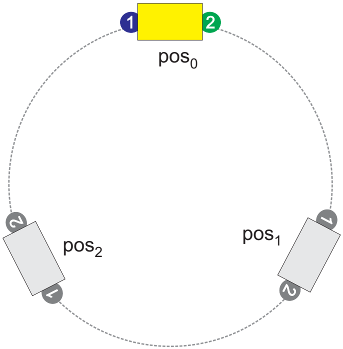
Epistemic logic

- We interpret $\Box_i \varphi$ as "agent i knows that φ"
- $\Box_i$ is usually written as K_i in epistemic logic

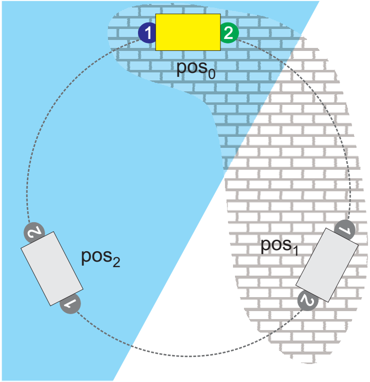
- Possible worlds: states of the system, situations
- Modal relations $\sim_i$: indistinguishability of states for agent i
- We assume that $\sim_i$ are equivalence relations

- $M, w \models K_i \varphi$ iff φ holds in all worlds that look the same as w

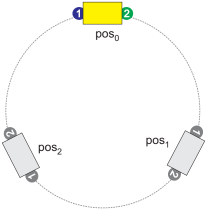
Example: Robots and Carriage

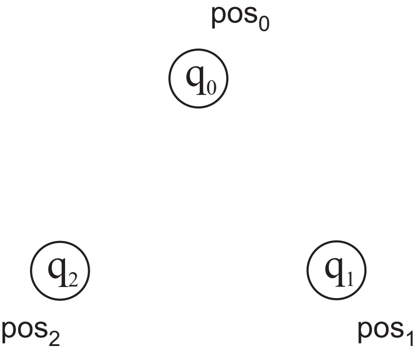


Example: Robots and Carriage

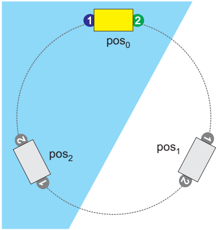


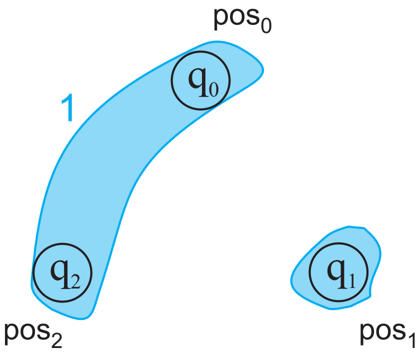
Example: Robots and Carriage



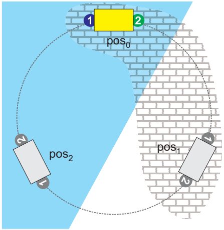


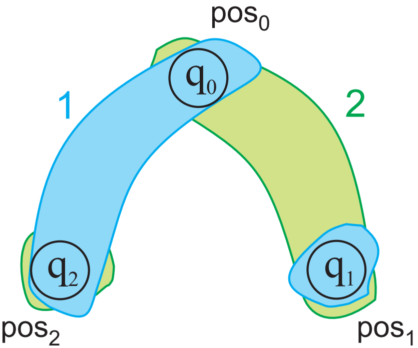
Example: Robots and Carriage



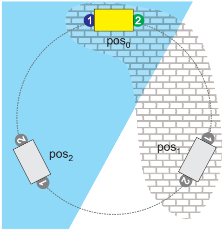


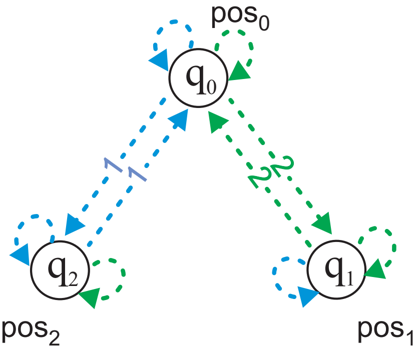
Example: Robots and Carriage



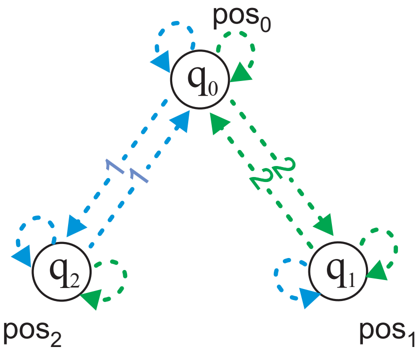


Example: Robots and Carriage

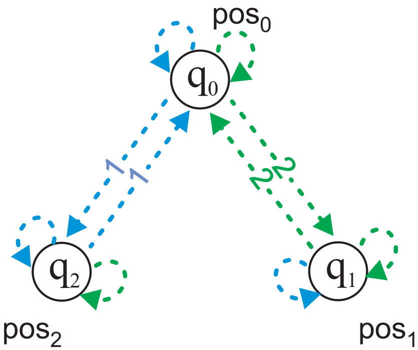




Example: Robots and Carriage

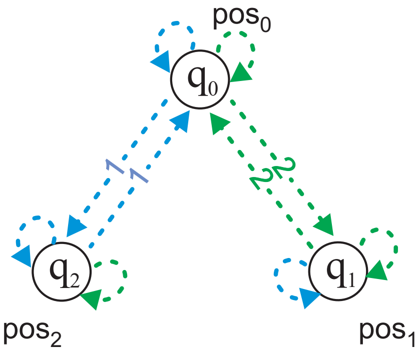


Example: Robots and Carriage



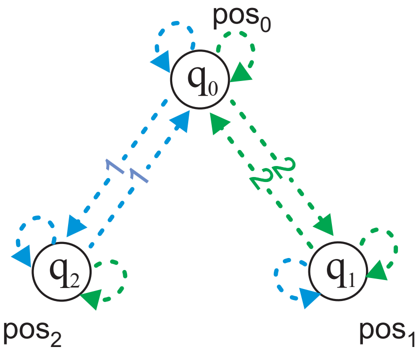
$$\mathsf{pos}_2 \rightarrow \neg K_1 \mathsf{pos}_2$$

Example: Robots and Carriage



$$\mathsf{pos}_2 \rightarrow \neg K_1 \mathsf{pos}_2$$
$$\mathsf{pos}_2 \rightarrow K_1 \neg \mathsf{pos}_1$$

Example: Robots and Carriage



$$\mathsf{pos}_2 \rightarrow \neg K_1 \mathsf{pos}_2$$
$$\mathsf{pos}_2 \rightarrow K_1 \neg \mathsf{pos}_1$$
$$\mathsf{pos}_2 \rightarrow K_2 K_1 \neg \mathsf{pos}_1$$

Logical omniscience

If φ is valid then $K_i\varphi$ also holds

Logical omniscience

If φ is valid then $K_i\varphi$ also holds
Problem!

Logical omniscience

If φ is valid then $K_i\varphi$ also holds
Problem!

Example

Do the whites have a winning strategy in chess?

Collective knowledge

A group of agents A can "know" that φ in several different epistemic modes:

Collective knowledge

A group of agents A can “know” that φ in several different epistemic modes:

- $E_A\varphi$: everybody in A knows that φ (A have mutual knowledge that φ)

Collective knowledge

A group of agents A can “know” that φ in several different epistemic modes:

- $E_A\varphi$: everybody in A knows that φ (A have mutual knowledge that φ)
- $C_A\varphi$: it is a common knowledge among A that φ

Collective knowledge

A group of agents A can “know” that φ in several different epistemic modes:

- $E_A\varphi$: everybody in A knows that φ (A have mutual knowledge that φ)
- $C_A\varphi$: it is a common knowledge among A that φ
- $D_A\varphi$: A have distributed knowledge that φ

Collective knowledge

A group of agents A can "know" that φ in several different epistemic modes:

- $E_A\varphi$: everybody in A knows that φ (A have mutual knowledge that φ)
- $C_A\varphi$: it is a common knowledge among A that φ
- $D_A\varphi$: A have distributed knowledge that φ

Multi-agent Epistemic Logic (**MAEL**$_n$): $\mathbf{K}_n$ plus modalities for mutual, common, and distributed knowledge

Collective knowledge: semantics

- $M, q \models E_A\varphi$ iff $M, q' \models \varphi$ for every q' such that $q \sim_A^E q'$, where $\sim_A^E = \bigcup_{i \in A} \sim_i$

Collective knowledge: semantics

- $M, q \models E_A\varphi$ iff $M, q' \models \varphi$ for every q' such that $q \sim_A^E q'$, where $\sim_A^E = \bigcup_{i \in A} \sim_i$
- $M, q \models C_A\varphi$ iff $M, q' \models \varphi$ for every q' such that $q \sim_A^C q'$, where $\sim_A^C$ is the transitive closure of $\sim_A^E$

Collective knowledge: semantics

- $M, q \models E_A\varphi$ iff $M, q' \models \varphi$ for every q' such that $q \sim_A^E q'$, where $\sim_A^E = \bigcup_{i \in A} \sim_i$
- $M, q \models C_A\varphi$ iff $M, q' \models \varphi$ for every q' such that $q \sim_A^C q'$, where $\sim_A^C$ is the transitive closure of $\sim_A^E$
- $M, q \models D_A\varphi$ iff $M, q' \models \varphi$ for every q' such that $q \sim_A^E q'$, where $\sim_A^E = \bigcap_{i \in A} \sim_i$

1.4 Axioms

As in classical logic, one can ask about a complete axiom system. Is there a calculus that allows to derive all sentences that are true in all Kripke models?

As in classical logic, one can ask about a complete axiom system. Is there a calculus that allows to derive all sentences that are true in all Kripke models?

Definition 1.4 (System K)

System **K** is an extension of the propositional calculus by the axiom

$$\textbf{K} \quad (\Box\varphi \wedge \Box(\varphi \rightarrow \psi)) \rightarrow \Box\psi$$

and the inference rule

$$\textbf{(Necessitation)} \quad \frac{\varphi}{\Box\varphi}.$$

Theorem 1.5 (Soundness/completeness of system K)

*System **K** is sound and complete with respect to the class of all Kripke models.*

Theorem 1.5 (Soundness/completeness of system K)

*System **K** is sound and complete with respect to the class of all Kripke models.*

Note: with n modalities, the calculus is called $\mathbf{K}_n$, and the theorem extends in a straightforward way.

Definition 1.6 (Extending K with axioms D, T, 4, 5)

System **K** is often extended by the following axioms:

K	$(\Box\varphi \wedge \Box(\varphi \rightarrow \psi)) \rightarrow \Box\psi$
D	$\neg\Box(\varphi \wedge \neg\varphi)$
T	$\Box\varphi \rightarrow \varphi$
4	$\Box\varphi \rightarrow \Box\Box\varphi$
5	$\neg\Box\varphi \rightarrow \Box\neg\Box\varphi$

Definition 1.6 (Extending K with axioms D, T, 4, 5)

System **K** is often extended by the following axioms:

K	$(\Box\varphi \wedge \Box(\varphi \rightarrow \psi)) \rightarrow \Box\psi$
D	$\neg\Box(\varphi \wedge \neg\varphi)$
T	$\Box\varphi \rightarrow \varphi$
4	$\Box\varphi \rightarrow \Box\Box\varphi$
5	$\neg\Box\varphi \rightarrow \Box\neg\Box\varphi$

Best known extensions of system **K**:

- **S5 = KDT45**: the standard logic of knowledge
- **KD45**: the standard logic of beliefs

Theorem 1.7 (Sound/complete subsystems of KDT45)

Let $\mathbf{X}$ *be any subset of* $\{\mathbf{D}, \mathbf{T}, \mathbf{4}, \mathbf{5}\}$ *and let* $\mathcal{X}$ *be any subset of* $\{serial, reflexive, transitive, euclidean\}$ *corresponding to* $\mathbf{X}$*.*
Then $\mathbf{K} \cup \mathbf{X}$ *is sound and complete with respect to Kripke models the accessibility relation of which satisfies* $\mathcal{X}$*.*

Theorem 1.7 (Sound/complete subsystems of KDT45)

Let $\mathbf{X}$ *be any subset of* $\{\mathbf{D}, \mathbf{T}, \mathbf{4}, \mathbf{5}\}$ *and let* $\mathcal{X}$ *be any subset of* $\{serial, reflexive, transitive, euclidean\}$ *corresponding to* $\mathbf{X}$.
Then $\mathbf{K} \cup \mathbf{X}$ *is sound and complete with respect to Kripke models the accessibility relation of which satisfies* $\mathcal{X}$.

Corollary 1.8

*System **S5** is sound and complete with respect to Kripke models with equivalence accessibility relations.*

Theorem 1.7 (Sound/complete subsystems of KDT45)

Let $\mathbf{X}$ *be any subset of* $\{\mathbf{D}, \mathbf{T}, \mathbf{4}, \mathbf{5}\}$ *and let* $\mathcal{X}$ *be any subset of* $\{serial, reflexive, transitive, euclidean\}$ *corresponding to* $\mathbf{X}$.
Then $\mathbf{K} \cup \mathbf{X}$ *is sound and complete with respect to Kripke models the accessibility relation of which satisfies* $\mathcal{X}$.

Corollary 1.8

*System **S5** is sound and complete with respect to Kripke models with equivalence accessibility relations.*

Theorem 1.9

Deciding if φ *is a theorem of **S5** is PSPACE-complete.*

Axioms for collective knowledge

Axioms for $\mathbf{MAEL}_n$ extend $\mathbf{S5}_n$ with schemes:

- Axioms for E_A: $E_A\varphi \leftrightarrow \bigwedge_{i\in A} E_i\varphi$

Axioms for collective knowledge

Axioms for $\mathbf{MAEL}_n$ extend $\mathbf{S5}_n$ with schemes:

- Axioms for E_A: $E_A\varphi \leftrightarrow \bigwedge_{i\in A} E_i\varphi$
- Segerberg's axioms for E_A and C_A:

$$\begin{array}{ll} \mathbf{MIX}_A: & C_A\varphi \rightarrow (\varphi \wedge E_A C_A\varphi) \\ \mathbf{IND}_A: & \varphi \wedge C_A(\varphi \rightarrow E_A\varphi) \rightarrow C_A\varphi \end{array}$$

Axioms for collective knowledge

Axioms for **MAEL**$_n$ extend **S5**$_n$ with schemes:

- Axioms for E_A: $E_A\varphi \leftrightarrow \bigwedge_{i\in A} E_i\varphi$
- Segerberg's axioms for E_A and C_A:

$$\begin{array}{ll} \mathbf{MIX}_A: & C_A\varphi \to (\varphi \wedge E_A C_A\varphi) \\ \mathbf{IND}_A: & \varphi \wedge C_A(\varphi \to E_A\varphi) \to C_A\varphi \end{array}$$

- Axioms for D_A:

$$\begin{array}{ll} \mathbf{S5(D_A)}: & \text{The } S5 \text{ axioms for } D_A, \\ \mathbf{D_i}: & D_i\varphi \leftrightarrow E_i\varphi, \\ \mathbf{INCL(D)}: & D_A\varphi \to D_{\mathrm{B}}\varphi \text{ whenever } A \subseteq \mathrm{B}. \end{array}$$

Axioms for collective knowledge

Axioms for **MAEL**$_n$ extend **S5**$_n$ with schemes:

- Axioms for E_A: $E_A\varphi \leftrightarrow \bigwedge_{i\in A} E_i\varphi$
- Segerberg's axioms for E_A and C_A:

$$\begin{array}{ll} \textbf{MIX}_A: & C_A\varphi \rightarrow (\varphi \wedge E_A C_A\varphi) \\ \textbf{IND}_A: & \varphi \wedge C_A(\varphi \rightarrow E_A\varphi) \rightarrow C_A\varphi \end{array}$$

- Axioms for D_A:

$$\begin{array}{ll} \mathbf{S5(D_A)}: & \text{The } S5 \text{ axioms for } D_A, \\ \mathbf{D_i}: & D_i\varphi \leftrightarrow E_i\varphi, \\ \mathbf{INCL(D)}: & D_A\varphi \rightarrow D_{\mathrm{B}}\varphi \text{ whenever } A \subseteq \mathrm{B}. \end{array}$$

- ...plus, for each $i \in \mathbb{A}\mathrm{gt}$, the inference rule

$$\frac{\varphi}{E_i\varphi}$$

Theorem 1.10 (Soundness/completeness of MAEL$_n$)

The axiom system for ***MAEL$_n$*** *is sound and complete.*

Theorem 1.10 (Soundness/completeness of MAEL$_n$)

*The axiom system for **MAEL**$_n$ is sound and complete.*

Theorem 1.11

*Deciding if φ is a theorem of **MAEL**$_n$ is EXPTIME-complete. It remains EXPTIME-complete even if only common or distributed knowledge operators are included.*

1.5 References

Multi-agent systems:

[Weiss1999] G. Weiss, editor.
Multiagent Systems. A Modern Approach to Distributed Artificial Intelligence.
MIT Press: Cambridge, Mass, 1999.

[Wooldridge2002] M. Wooldridge.
An Introduction to Multi Agent Systems.
John Wiley & Sons, 2002.

[fipa] Foundation for Intelligent Physical Agents.
FIPA home page.
`http://www.fipa.org/`.

Modal logic:

[Blackburn *et al.*2001] P. Blackburn, M. de Rijke, and V. Venema.
Modal Logic.
Cambridge Univ. Press, 2001.

[kripke63a] Kripke, S. (1963b).
Semantical considerations on modal logic.
Acta Philosophica Fennica 16, 83–94.

[Huth00] Huth, M. & Ryan, M.
Logic in Computer Science: Modeling and reasoning about systems.
Cambridge University Press.

Reasoning about knowledge:

[Fagin *et al.*1995] R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi.
Reasoning about Knowledge.
MIT Press: Cambridge, MA, 1995.

[Halpern1995] J. Y. Halpern.
Reasoning about knowledge: a survey.
In D. M. Gabbay, C. J. Hogger, and J. A. Robinson, editors, *The Handbook of Logic in Artificial Intelligence and Logic Programming, Volume IV*, pages 1–34. Oxford University Press, 1995.

[van der Hoek and Verbrugge2002] W. van der Hoek and R. Verbrugge.
Epistemic logic: A survey.
Game Theory and Applications, 8:53–94, 2002.

Coalition Logic

2.1 Strategic Games

Game Theory

- Game theory is concerned with understanding what happens when rational decision-makers interact
- Non-cooperative game theory: actions of individual players are taken as primitives
  - Strategic games: players simultaneously choose complete strategies at the beginning of the game
  - Extensive games: players can postpone decisions about which actions to choose until they are needed
- Cooperative (or coalitional) game theory: actions of coalitions, i.e. groups of players, are taken as primitives

Game Theory

- Game theory is concerned with understanding what happens when rational decision-makers interact
- Non-cooperative game theory: actions of individual players are taken as primitives
  - Strategic games: players simultaneously choose complete strategies at the beginning of the game
  - Extensive games: players can postpone decisions about which actions to choose until they are needed
- Cooperative (or coalitional) game theory: actions of coalitions, i.e. groups of players, are taken as primitives

This lecture: (repeated) strategic game forms. (Lecture 3: extensive form games; lectures 7 and 8: solution concepts for non-cooperative games; lectures 9 and 10: solution concepts for coalitional games)

Strategic game form

A game form is a tuple

$$G = (N, \{\Sigma_i : i \in N\}, o, S)$$

where

- N is a nonempty finite set of players
- S is a nonempty set of states (or outcomes or consequences)
- Σ_i is the nonempty set of actions (strategies) for agent i
- $o : \times_{i \in N} \Sigma_i \to S$ associates a state with a strategy profile

Strategic game form

A game form is a tuple

$$G = (N, \{\Sigma_i : i \in N\}, o, S)$$

where

- N is a nonempty finite set of players
- S is a nonempty set of states (or outcomes or consequences)
- Σ_i is the nonempty set of actions (strategies) for agent i
- $o : \times_{i \in N} \Sigma_i \to S$ associates a state with a strategy profile

Notation:

- σ_G where $G \subseteq N$: a partial strategy profile $\sigma_G \subseteq \times_{i \in G} \Sigma_i$
- σ_{-i}: same as $\sigma_{N \setminus \{i\}}$
- Γ_S^N is the set of all game forms with players N and

Strategic game

When

$$G = (N, \{\Sigma_i : i \in N\}, o, S)$$

and, for each player $i \in N$,

$$\succeq_i \subseteq S \times S$$

is a preference relation (complete, reflexive, transitive), then

$$(G, \{\succeq_i : i \in N\})$$

is a strategic game

Example: Prisoners' Dilemma

		Bill	
		Cooperate	**Defect**
Ann	**Cooperate**	Ann:-1, Bill:-1	Ann:-4, Bill: 0
	Defect	Ann:0, Bill:-4	Ann:-3, Bill: -3

Example: Prisoners' Dilemma

		Bill	
		Cooperate	**Defect**
Ann	**Cooperate**	Ann:-1, Bill:-1	Ann:-4, Bill: 0
	Defect	Ann:0, Bill:-4	Ann:-3, Bill: -3

On game form:

$$PD = (N, \{\Sigma_i : i \in N\}, o, S)$$

where

- $N = \{Ann, Bill\}$
- $\Sigma_{Ann} = \Sigma_{Bill} = \{C, D\}$
- $S = \{-1/-1, -4/0, 0/-4, -3/-3\}$
- $o(C,C) = -1/-1$, $o(C,D) = -4/0$, etc.

2.2 Coalition Logic: Introduction

Modal Logic and Games

- (Extensive form) games look like Kripke structures! (van Benthem)
- Here we use strategic games and Marc Paulys Coalition Logic as a starting point

Coalition Logic (Pauly 2001)

- We can interpret modal logic in transition systems, and reason about how the system possibly or necessarily will evolve
- Game frames: transition systems where the transitions are determined by a strategic game form in each state (and where the outcomes are, again, states)
- Coalition logic: about what coalitions can do (or ensure or make come about) – coalitional power
- Main construct, $C \subseteq N$:

  $\langle\langle C\rangle\rangle\varphi$ C can make φ come about

- Alternative interpretations:
  - A logic of game frames
  - A logic of coalitional effectivity

Example 2.1

Two individuals, A and B, must choose between two outcomes, p and q. We want a mechanism that will allow them to choose which will satisfy the following requirements: We want an outcome to be possible – that is, we want the two agents to choose, collectively, either p or q. We do not want them to be able to bring about both outcomes simultaneously. Finally, we do not want either agent to be able to unilaterally dictate an outcome – we want them both to have "equal power".

Example 2.1

Two individuals, A and B, must choose between two outcomes, p and q. We want a mechanism that will allow them to choose which will satisfy the following requirements: We want an outcome to be possible – that is, we want the two agents to choose, collectively, either p or q. We do not want them to be able to bring about both outcomes simultaneously. Finally, we do not want either agent to be able to unilaterally dictate an outcome – we want them both to have "equal power".

These requirements may be formally and naturally represented using Coalition Logic, as follows:

Example 2.1

Two individuals, A and B, must choose between two outcomes, p and q. We want a mechanism that will allow them to choose which will satisfy the following requirements: We want an outcome to be possible – that is, we want the two agents to choose, collectively, either p or q. We do not want them to be able to bring about both outcomes simultaneously. Finally, we do not want either agent to be able to unilaterally dictate an outcome – we want them both to have "equal power".

These requirements may be formally and naturally represented using Coalition Logic, as follows:

$$\langle\langle A, B\rangle\rangle p \quad \langle\langle A, B\rangle\rangle q$$

Example 2.1

Two individuals, A and B, must choose between two outcomes, p and q. We want a mechanism that will allow them to choose which will satisfy the following requirements: We want an outcome to be possible – that is, we want the two agents to choose, collectively, either p or q. We do not want them to be able to bring about both outcomes simultaneously. Finally, we do not want either agent to be able to unilaterally dictate an outcome – we want them both to have “equal power”.

These requirements may be formally and naturally represented using Coalition Logic, as follows:

$$\langle\!\langle A, B\rangle\!\rangle p \quad \langle\!\langle A, B\rangle\!\rangle q \quad \neg\langle\!\langle A, B\rangle\!\rangle (p \wedge q)$$

Example 2.1

Two individuals, A and B, must choose between two outcomes, p and q. We want a mechanism that will allow them to choose which will satisfy the following requirements: We want an outcome to be possible – that is, we want the two agents to choose, collectively, either p or q. We do not want them to be able to bring about both outcomes simultaneously. Finally, we do not want either agent to be able to unilaterally dictate an outcome – we want them both to have "equal power".

These requirements may be formally and naturally represented using Coalition Logic, as follows:

$$\begin{array}{lll} \langle\!\langle A,B\rangle\!\rangle p & \langle\!\langle A,B\rangle\!\rangle q & \neg\langle\!\langle A,B\rangle\!\rangle(p\wedge q) \\ \neg\langle\!\langle A\rangle\!\rangle q & \neg\langle\!\langle A\rangle\!\rangle p & \neg\langle\!\langle B\rangle\!\rangle p \\ \neg\langle\!\langle B\rangle\!\rangle q & & \end{array}$$

2.3 From Game Forms to Effectivity Functions

- We want to reason about coalitional power: can a coalition C achieve an outcome $X \subseteq S$?
- Coalitional power can be explicitly formalised by effectivity functions

- We want to reason about coalitional power: can a coalition C achieve an outcome $X \subseteq S$?
- Coalitional power can be explicitly formalised by effectivity functions

An effectivity function E is a function:

$$E : \wp(N) \to \wp(\wp(S))$$

$X \in E(C)$ means that the coalition C is effective for the outcome X

(Note: the litterature differs in conditions imposed on eff. functions. More about this later.)

A game form G induces an effectivity function E_G:

$$X \in E_G(C) \Leftrightarrow \exists \sigma_C \forall \sigma_{N \setminus C} o(\sigma_C, \sigma_{N \setminus C}) \in X$$

(where σ_C is a tuple of strategies for C)

A game form G induces an effectivity function E_G:

$$X \in E_G(C) \Leftrightarrow \exists \sigma_C \forall \sigma_{N \setminus C} o(\sigma_C, \sigma_{N \setminus C}) \in X$$

(where σ_C is a tuple of strategies for C)

- This form of effectivity is called α-effectivity
- C is effective for $X \subseteq S$ iff C has a strategy such that the next state will be in the set X, no matter which strategies the players in $N \setminus C$ use

Example

		Bill	
		Cooperate	**Defect**
Ann	**Cooperate**	Ann:-1, Bill:-1	Ann:-4, Bill: 0
	Defect	Ann:0, Bill:-4	Ann:-3, Bill: -3

PD induces the effectivity function E_{PD}:

$$E_{PD}(\emptyset) = \{S\}$$
$$E_{PD}(\{Ann\}) = \{\ \{-1/-1, -4/0\},\ \{0/-4, -3/-3\}\ \}^+$$
$$E_{PD}(\{Bill\}) = \{\ \{-1/-1, 0/-4\},\ \{-4/0, -3/-3\}\ \}^+$$
$$E_{PD}(\{Ann, Bill\}) = \wp(S) \setminus \emptyset$$

where X^+ is X closed under outcome-monotonicity (i.e.: $X \subseteq X^+$, and if $Y \in X^+$ and $Y \subseteq Y' \subseteq S$ then $Y' \in X^+$)

2.4 Coalition Logic

Syntax

$$
\begin{array}{lll}
\phi ::= & & \\
& \bot & \text{contradiction} \\
& p & \text{atomic prop.} \\
& \neg\phi & \text{negation} \\
& \phi \vee \phi & \text{disjunction} \\
& \langle\langle C \rangle\rangle \phi & C \text{ can enforce } \phi
\end{array}
$$

where $C \subseteq N$

Game Frames

A game frame is a pair

$$(S, \gamma)$$

where S are the states and

$$\gamma : S \to \Gamma_S^N$$

associates a strategic game form to each state

Game Frame Example: Repeated PD

Recall our game form: $PD = (N, \{\Sigma_i : i \in N\}, o, S)$ where

- $N = \{Ann, Bill\}$
- $\Sigma_{Ann} = \Sigma_{Bill} = \{C, D\}$
- $S = \{-1/-1, -4/0, 0/-4, -3/-3\}$
- $o(C,C) = -1/-1$, $o(C,D) = -4/0$, etc.

Game Frame Example: Repeated PD

Recall our game form: $PD = (N, \{\Sigma_i : i \in N\}, o, S)$ where

- $N = \{Ann, Bill\}$
- $\Sigma_{Ann} = \Sigma_{Bill} = \{C, D\}$
- $S = \{-1/-1, -4/0, 0/-4, -3/-3\}$
- $o(C, C) = -1/-1$, $o(C, D) = -4/0$, etc.

The game frame

$$\mathcal{F}_{PD} = (S, \gamma)$$

where

$$S = \{-1/-1, -4/0, 0/-4, -3/-3\}$$
$$\gamma(s) = PD \text{ for all } s \in S$$

models repeated play of prisoners' dilemma

Models

- A model is a pair

$$\mathcal{M} = (\mathcal{F}, V)$$

  where
  - $\mathcal{F}$ is a game frame
  - V assigns propositional atoms to states

Interpretation

Truth of a formula in a state s of a model $\mathcal{M}$:

$$
\begin{array}{lcl}
\mathcal{M}, s \not\models \bot & & \\
\mathcal{M}, s \models p & \Leftrightarrow & s \in V(p) \text{ (} p \text{ atomic prop.)} \\
\mathcal{M}, s \models \neg\phi & \Leftrightarrow & \mathcal{M}, s \not\models \phi \\
\mathcal{M}, s \models \phi \vee \psi & \Leftrightarrow & \mathcal{M}, s \models \phi \text{ or } \mathcal{M}, s \models \psi \\
\mathcal{M}, s \models \langle\langle C \rangle\rangle \phi & \Leftrightarrow & \phi^{\mathcal{M}} \in E_{\gamma(s)}(C)
\end{array}
$$

where $\phi^{\mathcal{M}} = \{s \in S : \mathcal{M}, s \models \phi\}$

Interpretation: Example

$$\mathcal{M}_{PD} = ((S, \gamma), V)$$

where

$S = \{1/1, 4/0, 0/4, 3/3\}$ $\gamma(s) = PD$ for each $s \in S$

V assigns truth values to propositions of the type

- $A = 1$ (Ann gets one year)
- $B \geq 3$ (Bill gets at least three years)

in the natural way (e.g., $A = 1 \in V(1/1)$)

Interpretation: Example

$$\mathcal{M}_{PD} = ((S, \gamma), V)$$

where

$S = \{1/1, 4/0, 0/4, 3/3\}$ $\gamma(s) = PD$ for each $s \in S$

V assigns truth values to propositions of the type

- $A = 1$ (Ann gets one year)
- $B \geq 3$ (Bill gets at least three years)

in the natural way (e.g., $A = 1 \in V(1/1)$)

Let s be any state.

- $\mathcal{M}_{PD}, s \models \langle\langle Ann \rangle\rangle B \geq 3$

Interpretation: Example

$$\mathcal{M}_{PD} = ((S, \gamma), V)$$

where

$S = \{1/1, 4/0, 0/4, 3/3\}$ $\gamma(s) = PD$ for each $s \in S$

V assigns truth values to propositions of the type

- $A = 1$ (Ann gets one year)
- $B \geq 3$ (Bill gets at least three years)

in the natural way (e.g., $A = 1 \in V(1/1)$)

Let s be any state.

- $\mathcal{M}_{PD}, s \models \langle\!\langle Ann \rangle\!\rangle B \geq 3$
- $\mathcal{M}_{PD}, s \models \langle\!\langle \{Ann, Bill\} \rangle\!\rangle A = 1 \wedge B = 1$

Interpretation: Example

$$\mathcal{M}_{PD} = ((S, \gamma), V)$$

where

$S = \{1/1, 4/0, 0/4, 3/3\}$ $\qquad$ $\gamma(s) = PD$ for each $s \in S$

V assigns truth values to propositions of the type

- $A = 1$ (Ann gets one year)
- $B \geq 3$ (Bill gets at least three years)

in the natural way (e.g., $A = 1 \in V(1/1)$)

Let s be any state.

- $\mathcal{M}_{PD}, s \models \langle\!\langle Ann \rangle\!\rangle B \geq 3$
- $\mathcal{M}_{PD}, s \models \langle\!\langle \{Ann, Bill\} \rangle\!\rangle A = 1 \wedge B = 1$
- $\mathcal{M}_{PD}, s \models \neg \langle\!\langle Ann \rangle\!\rangle A = 0$

2.5 Effectivity Properties

Playable Effectivity Functions

An effectivity function

$$E : \wp(N) \rightarrow \wp(\wp(S))$$

is playable iff:

Playable Effectivity Functions

An effectivity function

$$E : \wp(N) \rightarrow \wp(\wp(S))$$

is playable iff:

1. $\forall C \subseteq N$: $\emptyset \notin E(C)$

Playable Effectivity Functions

An effectivity function

$$E : \wp(N) \to \wp(\wp(S))$$

is playable iff:

1. $\forall C \subseteq N$: $\emptyset \notin E(C)$
2. $\forall C \subseteq N$: $S \in E(C)$

Playable Effectivity Functions

An effectivity function

$$E : \wp(N) \to \wp(\wp(S))$$

is playable iff:

1. $\forall C \subseteq N$: $\emptyset \notin E(C)$
2. $\forall C \subseteq N$: $S \in E(C)$
3. $\forall X \subseteq S$: $S \setminus X \notin E(\emptyset) \Rightarrow X \in E(N)$ (N-maximality)

Playable Effectivity Functions

An effectivity function

$$E : \wp(N) \to \wp(\wp(S))$$

is playable iff:

1. $\forall C \subseteq N$: $\emptyset \notin E(C)$
2. $\forall C \subseteq N$: $S \in E(C)$
3. $\forall X \subseteq S$: $S \setminus X \notin E(\emptyset) \Rightarrow X \in E(N)$ (N-maximality)
4. $\forall C$: $\forall X \subseteq X' \subseteq S$: $X \in E(C) \Rightarrow X' \in E(C)$ (outcome-monotonicity)

Playable Effectivity Functions

An effectivity function

$$E : \wp(N) \to \wp(\wp(S))$$

is playable iff:

1. $\forall C \subseteq N$: $\emptyset \notin E(C)$
2. $\forall C \subseteq N$: $S \in E(C)$
3. $\forall X \subseteq S$: $S \setminus X \notin E(\emptyset) \Rightarrow X \in E(N)$ (N-maximality)
4. $\forall C$: $\forall X \subseteq X' \subseteq S$: $X \in E(C) \Rightarrow X' \in E(C)$ (outcome-monotonicity)
5. $\forall C_1 \subseteq N$: $\forall C_2 \subseteq N$: $\forall X_1 \subseteq S$: $\forall X_2 \subseteq S$: ($C_1 \cap C_2 = \emptyset$ and $X_1 \in E(C_1)$ and $X_2 \in E(C_2)$) $\Rightarrow X_1 \cap X_2 \in E(C_1 \cup C_2)$ (superadditivity)

Strategic Game Forms vs. Effectivity Functions

Q: which effectivity functions are induced by strategic game forms?
A: exactly the playable effectivity functions

Theorem 2.2 (Pauly)

An effectivity function is playable iff it is induced by some strategic game form

- Since the induced effectivity function is the only information about the fame frame used in the interpretation, this means that coalition logic can be seen as a logic of playable effectivity functions
- Alternative – and equivalent – models:

$$\mathcal{M} = ((S, E), V)$$

  where $E(s)$ associates a playable effectivity function to each state
- This is a neighbourhood semantics: we get a neigbourhood relation for each coalition
- Technically easier to work with

Coalition Logic: Axiomatisation

A sound and complete axiomatisation of all models (ref. playability properties):

$\neg\langle\langle C\rangle\rangle\bot$	$(\bot)$
$\langle\langle C\rangle\rangle\neg\bot$	$(\top)$
$(\neg\langle\langle\emptyset\rangle\rangle\neg\phi) \rightarrow \langle\langle N\rangle\rangle\phi$	(N)
$\langle\langle C\rangle\rangle(\phi \wedge \psi) \rightarrow \langle\langle C\rangle\rangle\psi$	(M)
$(\langle\langle C_1\rangle\rangle\phi_1 \wedge \langle\langle C_2\rangle\rangle\phi_2) \rightarrow \langle\langle C_1 \cup C_2\rangle\rangle(\phi_1 \wedge \phi_2)$ when $C_1 \cap C_2 = \emptyset$	(S)
$\frac{\phi,\phi\rightarrow\psi}{\psi}$	(MP)
$\frac{\phi\leftrightarrow\psi}{\langle\langle C\rangle\rangle\phi\leftrightarrow\langle\langle C\rangle\rangle\psi}$	(EQ)

Playability properties again

1. $\forall C \subseteq N$: $\emptyset \notin E(C)$
2. $\forall C \subseteq N$: $S \in E(C)$
3. $\forall X \subseteq S$: $S \setminus X \notin E(\emptyset) \Rightarrow X \in E(N)$ (N-maximality)
4. $\forall C$: $\forall X \subseteq X' \subseteq S$: $X \in E(C) \Rightarrow X' \in E(C)$ (outcome-monotonicity)
5. $\forall C_1 \subseteq N$: $\forall C_2 \subseteq N$: $\forall X_1 \subseteq S$: $\forall X_2 \subseteq S$: ($C_1 \cap C_2 = \emptyset$ and $X_1 \in E(C_1)$ and $X_2 \in E(C_2)$) $\Rightarrow X_1 \cap X_2 \in E(C_1 \cup C_2)$ (superadditivity)

Completeness

- Can be shown by an canonical model construction
- Standard Lindenbaum argument: every consistent set of formulae can be extended to a max. cons. set

Canonical Model

$$\mathcal{M}^c = ((S^c, E^c), V^c)$$

S^c: all MCSs

Canonical Model

$$\mathcal{M}^c = ((S^c, E^c), V^c)$$

S^c: all MCSs

V^c: $s \in V^c(p) \Leftrightarrow p \in s$

Canonical Model

$$\mathcal{M}^c = ((S^c, E^c), V^c)$$

S^c: all MCSs

V^c: $s \in V^c(p) \Leftrightarrow p \in s$

E^c: associates with each s the canonical effectivity function $E^c(s)$:

$$X \in E^c(s)(G) \Leftrightarrow$$
$$\begin{cases} \exists \phi : \tilde{\phi} \subseteq X \text{ and } \langle\langle G \rangle\rangle \phi \in s & G \neq N \\ \forall \phi : \text{ if } \tilde{\phi} \subseteq S^c \setminus X \text{ then } \langle\langle \emptyset \rangle\rangle \phi \notin s & G = N \end{cases}$$

where $\tilde{\phi} = \{s \in S^c : \phi \in s\}$

Completeness

Truth Lemma

$$\mathcal{M}^c, s \models \phi \Leftrightarrow \phi \in s$$

for any MCS s and any ϕ

Completeness

Truth Lemma

$$\mathcal{M}^c, s \models \phi \Leftrightarrow \phi \in s$$

for any MCS s and any ϕ

- Easy to show by induction over ϕ, using the fact that we can derive the rule

$$\frac{\phi \to \psi}{\langle\langle C \rangle\rangle \phi \to \langle\langle C \rangle\rangle \psi}$$

It remains to be shown that $E^c(s)$ is playable.

It remains to be shown that $E^c(s)$ is playable.

Proposition: Alternative playability properties

An effectivity function E is playable iff:

1. $\forall C \neq N$: $\emptyset \notin E(C)$
2. $\forall C \neq N$: $S \in E(C)$
3. $\forall X \subseteq S$: $S \setminus X \notin E(\emptyset) \Rightarrow X \in E(N)$ (N-maximality)
4. $\forall C \neq N$: $\forall X \subseteq X' \subseteq S$: $X \in E(C) \Rightarrow X' \in E(C)$ (outcome-monotonicity)
5. $\forall C_1 \subset N$: $\forall C_2 \subset N$ s.t. $C_1 \cup C_2 \neq N$: $\forall X_1 \subseteq S$: $\forall X_2 \subseteq S$: $(C_1 \cap C_2 = \emptyset$ and $X_1 \in E(C_1)$ and $X_2 \in E(C_2)) \Rightarrow X_1 \cap X_2 \in E(C_1 \cup C_2)$ (superadditivity)
6. $\forall C \subseteq N$: $\forall X \subseteq S$: if $X \in E(C)$ then $S \setminus X \notin E(N \setminus C)$ (regularity)

It remains to be shown that $E^c(s)$ is playable.

Proposition: Alternative playability properties

An effectivity function E is playable iff:

1. $\forall C \neq N$: $\emptyset \notin E(C)$
2. $\forall C \neq N$: $S \in E(C)$
3. $\forall X \subseteq S$: $S \setminus X \notin E(\emptyset) \Rightarrow X \in E(N)$ (N-maximality)
4. $\forall C \neq N$: $\forall X \subseteq X' \subseteq S$: $X \in E(C) \Rightarrow X' \in E(C)$ (outcome-monotonicity)
5. $\forall C_1 \subset N$: $\forall C_2 \subset N$ s.t. $C_1 \cup C_2 \neq N$: $\forall X_1 \subseteq S$: $\forall X_2 \subseteq S$: $(C_1 \cap C_2 = \emptyset$ and $X_1 \in E(C_1)$ and $X_2 \in E(C_2)) \Rightarrow X_1 \cap X_2 \in E(C_1 \cup C_2)$ (superadditivity)
6. $\forall C \subseteq N$: $\forall X \subseteq S$: if $X \in E(C)$ then $S \setminus X \notin E(N \setminus C)$ (regularity)

Relatively easy to show that these must hold in $\mathcal{M}^c$

Superadditivity

Let $C_1 \cap C_2 = \emptyset$, $C_1 \cup C_2 \neq N$, $X_1 \in E^c(s)(C_1)$, $X_2 \in E^c(s)(C_2)$:

Superadditivity

Let $C_1 \cap C_2 = \emptyset$, $C_1 \cup C_2 \neq N$, $X_1 \in E^c(s)(C_1)$, $X_2 \in E^c(s)(C_2)$:

- $\langle\langle C_1 \rangle\rangle \phi_1, \langle\langle C_2 \rangle\rangle \phi_2 \in s$ for some ϕ_1, ϕ_2 s.t. $\tilde{\phi}_1 \subseteq X_1$ and $\tilde{\phi}_2 \subseteq X_2$

Superadditivity

Let $C_1 \cap C_2 = \emptyset$, $C_1 \cup C_2 \neq N$, $X_1 \in E^c(s)(C_1)$, $X_2 \in E^c(s)(C_2)$:

- $\langle\langle C_1 \rangle\rangle \phi_1, \langle\langle C_2 \rangle\rangle \phi_2 \in s$ for some ϕ_1, ϕ_2 s.t. $\tilde{\phi_1} \subseteq X_1$ and $\tilde{\phi_2} \subseteq X_2$
- By the superadditivity axiom: $\langle\langle C_1 \cup C_2 \rangle\rangle (\phi_1 \wedge \phi_2) \in s$

Superadditivity

Let $C_1 \cap C_2 = \emptyset$, $C_1 \cup C_2 \neq N$, $X_1 \in E^c(s)(C_1)$, $X_2 \in E^c(s)(C_2)$:

- $\langle\langle C_1 \rangle\rangle \phi_1, \langle\langle C_2 \rangle\rangle \phi_2 \in s$ for some ϕ_1, ϕ_2 s.t. $\tilde{\phi_1} \subseteq X_1$ and $\tilde{\phi_2} \subseteq X_2$
- By the superadditivity axiom: $\langle\langle C_1 \cup C_2 \rangle\rangle (\phi_1 \wedge \phi_2) \in s$
- $\widetilde{\phi_1 \wedge \phi_2} \subseteq X_1 \cap X_2$

Superadditivity

Let $C_1 \cap C_2 = \emptyset$, $C_1 \cup C_2 \neq N$, $X_1 \in E^c(s)(C_1)$, $X_2 \in E^c(s)(C_2)$:

- $\langle\langle C_1 \rangle\rangle \phi_1, \langle\langle C_2 \rangle\rangle \phi_2 \in s$ for some ϕ_1, ϕ_2 s.t. $\tilde{\phi}_1 \subseteq X_1$ and $\tilde{\phi}_2 \subseteq X_2$
- By the superadditivity axiom: $\langle\langle C_1 \cup C_2 \rangle\rangle (\phi_1 \wedge \phi_2) \in s$
- $\widetilde{\phi_1 \wedge \phi_2} \subseteq X_1 \cap X_2$
- Thus, $X_1 \cap X_2 \in E^c(s)(C_1 \cup C_2)$

Computational Complexity

- The problem:

  Given coalition logic formula ϕ is there some model that satisfies ϕ?

- Complexity: PSPACE-complete

2.6 Quantified Coalition Logic

Lack of Succinctness in CL

Take the property:

agent 1 is necessary to achieve φ

Lack of Succinctness in CL

Take the property:

agent 1 is necessary to achieve φ

Its expression in CL is exponentially long in the number of agents in the system. If $Ag = \{1, 2, 3, 4\}$:

$$\neg\langle\langle\{\}\rangle\rangle\varphi \wedge \neg\langle\langle\{2\}\rangle\rangle\varphi \wedge \neg\langle\langle\{3\}\rangle\rangle\varphi \wedge \neg\langle\langle\{4\}\rangle\rangle\varphi \wedge \neg\langle\langle\{2,3\}\rangle\rangle\varphi \wedge \neg\langle\langle\{3,4\}\rangle\rangle\varphi \wedge \neg\langle\langle\{2,4\}\rangle\rangle\varphi \wedge \neg\langle\langle\{2,3,4\}\rangle\rangle\varphi$$

Quantified Coalition Logic (QCL)

In QCL $\langle\langle\cdot\rangle\rangle$ is replaced by a collection of unary modal operators indexed by a coalition predicate P, in order to make the logic more succinct:

- $\langle P\rangle\varphi$: there exists some coalition satisfying P which can achieve φ
- $[P]\varphi$: every coalition satisfying P can achieve φ

Quantified Coalition Logic (QCL)

In QCL $\langle\!\langle \cdot \rangle\!\rangle$ is replaced by a collection of unary modal operators indexed by a coalition predicate P, in order to make the logic more succinct:

$\langle P \rangle \varphi$: there exists some coalition satisfying P which can achieve φ

$[P]\varphi$: every coalition satisfying P can achieve φ

Examples of predicates (C' a coalition, n a number):

- $supseteq(C')$: satisfied by C iff $C \supseteq C'$
- $geq(n)$: satisfied by C iff $|C| \geq n$
- $gt(n)$: satisfied by C iff $|C| > n$
- $maj(n) \equiv geq(\lceil (n+1)/2 \rceil)$
- Boolean combinations

QCL: Example

agent 1 is necessary to achieve φ

QCL: Example

agent 1 is necessary to achieve φ

$$\neg\langle\neg supseteq\{1\}\rangle\varphi$$

QCL Example: voting

An electorate of n voters wishes to select one of two outcomes ω_1 and ω_2. They want to use a simple majority voting protocol, so that outcome ω_i will be selected iff a majority of the n voters state a preference for it. No coalition of less than majority size should be able to select an outcome, and any majority should be able to choose the outcome (i.e., the selection procedure is not influenced by the "names" of the agents in a coalition).

QCL Example: voting

An electorate of n voters wishes to select one of two outcomes ω_1 and ω_2. They want to use a simple majority voting protocol, so that outcome ω_i will be selected iff a majority of the n voters state a preference for it. No coalition of less than majority size should be able to select an outcome, and any majority should be able to choose the outcome (i.e., the selection procedure is not influenced by the "names" of the agents in a coalition).

$$([maj(n)]\omega_1) \wedge ([maj(n)]\omega_2)$$

QCL Example: voting

An electorate of n voters wishes to select one of two outcomes ω_1 and ω_2. They want to use a simple majority voting protocol, so that outcome ω_i will be selected iff a majority of the n voters state a preference for it. No coalition of less than majority size should be able to select an outcome, and any majority should be able to choose the outcome (i.e., the selection procedure is not influenced by the "names" of the agents in a coalition).

$$([maj(n)]\omega_1) \wedge ([maj(n)]\omega_2)$$

$$(\neg\langle\neg maj(n)\rangle\omega_1) \wedge (\neg\langle\neg maj(n)\rangle\omega_2)$$

QCL

- QCL is no more expressive than Coalition Logic
- But is exponentially more succinct
- The model checking problem can be solved in polynomial time – assuming an explicit representation of models
- The model checking problem assuming an RML representation of models is PSPACE-complete.
- The satisfiability problem is PSCPACE-complete.

2.7 References

[1] M. J. Osborne and A. Rubinstein.
A Course in Game Theory.
The MIT Press: Cambridge, MA, 1994.

[2] M. Pauly.
Logic for Social Software.
PhD thesis, University of Amsterdam, 2001.
ILLC Dissertation Series 2001-10.

[3] M. Pauly.
A modal logic for coalitional power in games.
Journal of Logic and Computation, 12(1):149–166, 2002.

[4] W. van der Hoek and M. Pauly.
Modal logic for games and information.
In P. Blackburn, J. van Benthem, and F. Wolter, editors, *Handbook of Modal Logic*, pages 1077–1148. Elsevier Science Publishers B.V.: Amsterdam, The Netherlands, 2006.

[5] Thomas Ågotnes, Wiebe van der Hoek, and Michael Wooldridge.
Quantified coalition logic.
Synthese, 2008.

ATL

ATL: What Agents Can Achieve

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

ATL: What Agents Can Achieve

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

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

3.1 The Logic

Syntax

$$\varphi ::= \mathsf{p} \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\!\langle A \rangle\!\rangle \gamma,$$
$$\gamma ::= \varphi \mid \neg\gamma \mid \gamma \wedge \gamma \mid \bigcirc\gamma \mid \Diamond\gamma \mid \Box\gamma \mid \gamma\,\mathcal{U}\,\gamma.$$

Syntax

$$\varphi ::= \mathsf{p} \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\langle A\rangle\rangle\gamma,$$
$$\gamma ::= \varphi \mid \neg\gamma \mid \gamma \wedge \gamma \mid \bigcirc\gamma \mid \Diamond\gamma \mid \Box\gamma \mid \gamma\,\mathcal{U}\gamma.$$

In fact, “eventually” and “always” can be derived from “until”:

- $\Diamond\gamma \equiv \top\,\mathcal{U}\gamma$
- $\Box\gamma \equiv \neg\Diamond\neg\gamma$

Syntax

$$\varphi ::= \mathsf{p} \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\langle A \rangle\rangle \gamma,$$
$$\gamma ::= \varphi \mid \neg\gamma \mid \gamma \wedge \gamma \mid \bigcirc\gamma \mid \Diamond\gamma \mid \Box\gamma \mid \gamma\,\mathcal{U}\,\gamma.$$

In fact, "eventually" and "always" can be derived from "until":

- $\Diamond\gamma \equiv \top\,\mathcal{U}\,\gamma$
- $\Box\gamma \equiv \neg\Diamond\neg\gamma$

- "Vanilla" ATL: every temporal operator preceded by exactly one cooperation modality
- ATL*: no syntactic restrictions

- $\langle\!\langle jamesbond \rangle\!\rangle \Diamond (\mathsf{ski} \wedge \neg \mathsf{getBurned})$:
  "James Bond can go skiing without getting burned"

- $\langle\langle jamesbond \rangle\rangle \Diamond(\mathsf{ski} \wedge \neg \mathsf{getBurned})$:
  "James Bond can go skiing without getting burned"

- $\langle\langle jamesbond\rangle\rangle\Diamond(\mathsf{ski} \land \neg\mathsf{getBurned})$:
  "James Bond can go skiing without getting burned"

- $\langle\langle jamesbond, bondsgirl\rangle\rangle\mathsf{fun}\,\mathcal{U}\,\mathsf{shot}$:
  "James Bond and his girlfriend are able to have fun until someone shoots at them"

- ATL extends the branching-time logic CTL:

  $\mathsf{A} \equiv \langle\langle\emptyset\rangle\rangle$ ("for all paths")
  $\mathsf{E} \equiv \langle\langle\mathbb{A}\mathrm{gt}\rangle\rangle$ ("there is a path")

ATL Models: Concurrent Game Structures

- Agents, actions, transitions, atomic propositions
- Atomic propositions + interpretation
- Actions are abstract

Definition 3.1 (Concurrent Game Structure)

A **concurrent game structure** is a tuple
$M = \langle \mathbb{A}\mathrm{gt}, St, \pi, Act, d, o \rangle$, where:

Definition 3.1 (Concurrent Game Structure)

A **concurrent game structure** is a tuple $M = \langle \mathbb{A}\mathrm{gt}, St, \pi, Act, d, o \rangle$, where:

- $\mathbb{A}\mathrm{gt}$: a finite set of all agents

Definition 3.1 (Concurrent Game Structure)

A **concurrent game structure** is a tuple
$M = \langle \mathbb{A}\mathrm{gt}, St, \pi, Act, d, o \rangle$, where:

- $\mathbb{A}\mathrm{gt}$: a finite set of all agents
- St: a set of states

Definition 3.1 (Concurrent Game Structure)

A **concurrent game structure** is a tuple $M = \langle \mathbb{A}\mathrm{gt}, St, \pi, Act, d, o \rangle$, where:

- $\mathbb{A}\mathrm{gt}$: a finite set of all agents
- St: a set of states
- π: a valuation of propositions

Definition 3.1 (Concurrent Game Structure)

A **concurrent game structure** is a tuple $M = \langle \mathbb{A}\mathrm{gt}, St, \pi, Act, d, o \rangle$, where:

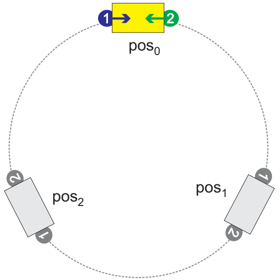
- $\mathbb{A}\mathrm{gt}$: a finite set of all agents
- St: a set of states
- π: a valuation of propositions
- Act: a finite set of (atomic) actions

Definition 3.1 (Concurrent Game Structure)

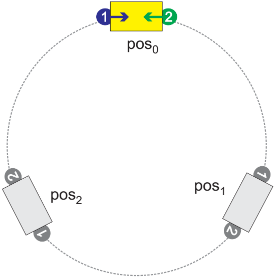
A **concurrent game structure** is a tuple
$M = \langle \mathbb{A}\mathrm{gt}, St, \pi, Act, d, o \rangle$, where:

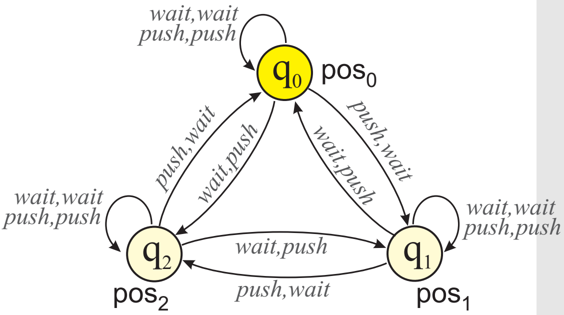
- $\mathbb{A}\mathrm{gt}$: a finite set of all agents
- St: a set of states
- π: a valuation of propositions
- Act: a finite set of (atomic) actions
- $d : \mathbb{A}\mathrm{gt} \times St \rightarrow 2^{Act}$ defines actions available to an agent in a state
- o: a deterministic transition function that assigns outcome states $q' = o(q, \alpha_1, \ldots, \alpha_k)$ to states and tuples of actions

Example: Robots and Carriage



Example: Robots and Carriage





Definition 3.2 (Strategy)

A **strategy** is a conditional plan.

Definition 3.2 (Strategy)

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

Definition 3.2 (Strategy)

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

$\leadsto$ memoryless agents

Definition 3.2 (Strategy)

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

$\rightsquigarrow$ memoryless agents

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

Definition 3.2 (Strategy)

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

$\rightsquigarrow$ memoryless agents

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

Function $out(q, s_A)$ returns the set of all paths that may result from agents A executing strategy s_A from state q onward.

Definition 3.3 (Semantics of ATL*)

$M, q \models p$ iff p is in $\pi(q)$;
$M, q \models \neg\varphi$ iff $M, q \not\models \varphi$;
$M, q \models \varphi_1 \wedge \varphi_2$ iff $M, q \models \varphi_1$ and $M, q \models \varphi_2$;

Definition 3.3 (Semantics of ATL*)

$M, q \models p$ iff p is in $\pi(q)$;
$M, q \models \neg\varphi$ iff $M, q \not\models \varphi$;
$M, q \models \varphi_1 \wedge \varphi_2$ iff $M, q \models \varphi_1$ and $M, q \models \varphi_2$;
$M, \lambda \models \neg\gamma$ iff $M, q \not\models \gamma$ etc.;

Definition 3.3 (Semantics of ATL*)

$M, q \models p$	iff p is in $\pi(q)$;
$M, q \models \neg\varphi$	iff $M, q \not\models \varphi$;
$M, q \models \varphi_1 \land \varphi_2$	iff $M, q \models \varphi_1$ and $M, q \models \varphi_2$;
$M, \lambda \models \neg\gamma$	iff $M, q \not\models \gamma$ etc.;
$M, q \models \langle\langle A \rangle\rangle \Phi$	iff there is a collective strategy s_A such that, for every path $\lambda \in out(q, s_A)$, we have $M, \lambda \models \Phi$.

Definition 3.3 (Semantics of ATL*)

$M, q \models p$	iff p is in $\pi(q)$;
$M, q \models \neg\varphi$	iff $M, q \not\models \varphi$;
$M, q \models \varphi_1 \wedge \varphi_2$	iff $M, q \models \varphi_1$ and $M, q \models \varphi_2$;
$M, \lambda \models \neg\gamma$	iff $M, q \not\models \gamma$ etc.;
$M, q \models \langle\langle A \rangle\rangle\Phi$	iff there is a collective strategy s_A such that, for every path $\lambda \in out(q, s_A)$, we have $M, \lambda \models \Phi$.
$M, \lambda \models \bigcirc\gamma$	iff $M, \lambda[1..\infty] \models \gamma$;

Definition 3.3 (Semantics of ATL*)

$M, q \models p$	iff p is in $\pi(q)$;
$M, q \models \neg\varphi$	iff $M, q \not\models \varphi$;
$M, q \models \varphi_1 \wedge \varphi_2$	iff $M, q \models \varphi_1$ and $M, q \models \varphi_2$;
$M, \lambda \models \neg\gamma$	iff $M, q \not\models \gamma$ etc.;
$M, q \models \langle\langle A \rangle\rangle \Phi$	iff there is a collective strategy s_A such that, for every path $\lambda \in out(q, s_A)$, we have $M, \lambda \models \Phi$.
$M, \lambda \models \bigcirc\gamma$	iff $M, \lambda[1..\infty] \models \gamma$;
$M, \lambda \models \Box\gamma$	iff $M, \lambda[i..\infty] \models \gamma$ for all $i \geq 0$;

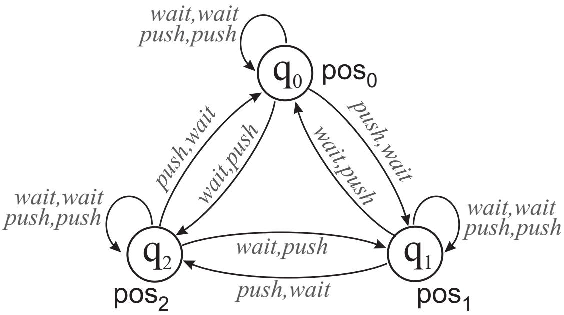
Definition 3.3 (Semantics of ATL*)

$M, q \models p$	iff p is in $\pi(q)$;
$M, q \models \neg\varphi$	iff $M, q \not\models \varphi$;
$M, q \models \varphi_1 \wedge \varphi_2$	iff $M, q \models \varphi_1$ and $M, q \models \varphi_2$;
$M, \lambda \models \neg\gamma$	iff $M, q \not\models \gamma$ etc.;
$M, q \models \langle\langle A \rangle\rangle \Phi$	iff there is a collective strategy s_A such that, for every path $\lambda \in out(q, s_A)$, we have $M, \lambda \models \Phi$.
$M, \lambda \models \bigcirc\gamma$	iff $M, \lambda[1..\infty] \models \gamma$;
$M, \lambda \models \Box\gamma$	iff $M, \lambda[i..\infty] \models \gamma$ for all $i \geq 0$;
$M, \lambda \models \gamma_1 \mathcal{U} \gamma_2$	iff $M, \lambda[i..\infty] \models \gamma_2$ for some $i \geq 0$, and $M, \lambda[j..\infty] \models \gamma_1$ forall $0 \leq j \leq i$.

The semantics of "vanilla" ATL can be given entirely in terms of models and states:

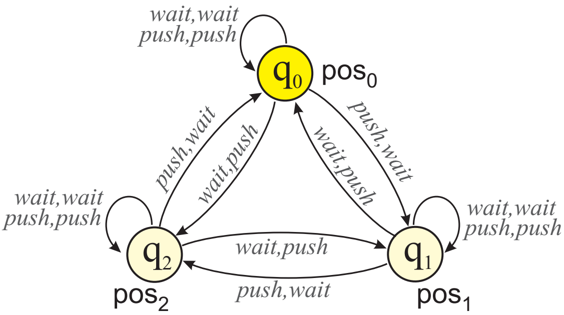
$M, q \models p$	iff p is in $\pi(q)$;
$M, q \models \neg\varphi$	iff $M, q \not\models \varphi$;
$M, q \models \varphi_1 \wedge \varphi_2$	iff $M, q \models \varphi_1$ and $M, q \models \varphi_2$;
$M, q \models \langle\langle A \rangle\rangle \bigcirc \varphi$	iff there is s_A such that, for every $\lambda \in out(q, s_A)$, we have $M, \lambda[1] \models \varphi$;
$M, q \models \langle\langle A \rangle\rangle \Box \varphi$	iff there is s_A such that, for every $\lambda \in out(q, s_A)$, we have $M, \lambda[i] \models \varphi$ for all $i \geq 0$;
$M, q \models \langle\langle A \rangle\rangle \varphi_1 \mathcal{U} \varphi_2$	iff there is s_A such that, for every $\lambda \in out(q, s_A)$, we have $M, \lambda[i] \models \varphi_2$ for some $i \geq 0$ and $M, \lambda[j] \models \varphi_1$ for all $0 \leq j \leq i$.

Example: Robots and Carriage



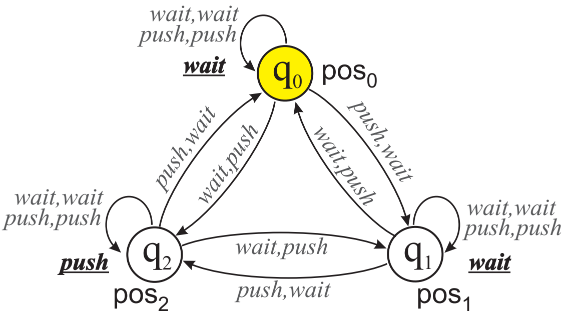
$$\mathsf{pos}_0 \rightarrow \langle\langle 1 \rangle\rangle \Box \neg \mathsf{pos}_1$$

Example: Robots and Carriage



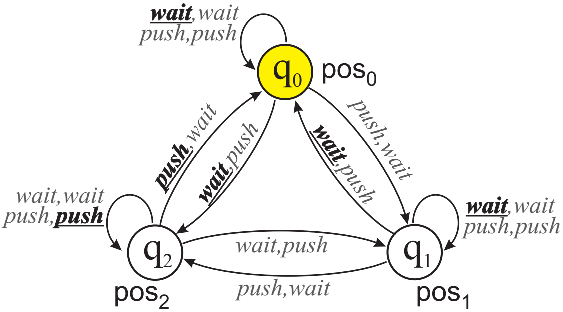
$$\mathsf{pos}_0 \longrightarrow \langle\!\langle 1 \rangle\!\rangle \Box \neg \mathsf{pos}_1$$

Example: Robots and Carriage



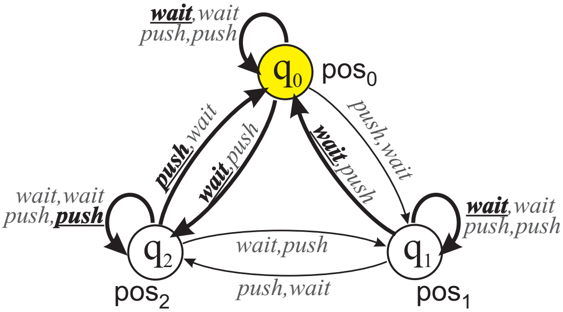
$$\mathsf{pos}_0 \rightarrow \langle\!\langle 1 \rangle\!\rangle \Box \neg \mathsf{pos}_1$$

Example: Robots and Carriage



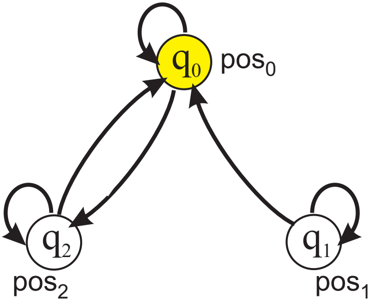
$$\mathsf{pos}_0 \rightarrow \langle\langle 1 \rangle\rangle \Box \neg \mathsf{pos}_1$$

Example: Robots and Carriage



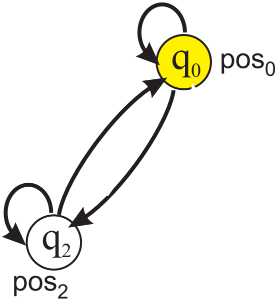
$$\mathsf{pos}_0 \rightarrow \langle\langle 1 \rangle\rangle \Box \neg \mathsf{pos}_1$$

Example: Robots and Carriage



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

Example: Robots and Carriage



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

Fixpoint Properties

Theorem 3.4

The following formulae are valid for ATL (but not for ATL!):*

- $\langle\!\langle A\rangle\!\rangle\Box\varphi \quad\leftrightarrow\quad \varphi \land \langle\!\langle A\rangle\!\rangle\bigcirc\langle\!\langle A\rangle\!\rangle\Box\varphi$
- $\langle\!\langle A\rangle\!\rangle\varphi_1\,\mathcal{U}\,\varphi_2 \quad\leftrightarrow\quad \varphi_2 \lor \varphi_1 \land \langle\!\langle A\rangle\!\rangle\bigcirc\langle\!\langle A\rangle\!\rangle\varphi_1\,\mathcal{U}\,\varphi_2.$

Fixpoint Properties

Theorem 3.4

The following formulae are valid for ATL (but not for ATL!):*

- $\langle\langle A\rangle\rangle\Box\varphi \quad\leftrightarrow\quad \varphi \wedge \langle\langle A\rangle\rangle\bigcirc\langle\langle A\rangle\rangle\Box\varphi$
- $\langle\langle A\rangle\rangle\varphi_1\mathcal{U}\varphi_2 \quad\leftrightarrow\quad \varphi_2 \vee \varphi_1 \wedge \langle\langle A\rangle\rangle\bigcirc\langle\langle A\rangle\rangle\varphi_1\mathcal{U}\varphi_2.$

Corollary

Strategy for A can be synthesized incrementally (no backtracking is necessary).

3.2 Agents, Systems, Games

Connection to Temporal Analysis of Systems

Temporal operators allow a number of useful concepts to be formally specified:

Connection to Temporal Analysis of Systems

Temporal operators allow a number of useful concepts to be formally specified:

- safety properties
- liveness properties
- fairness properties

Safety (maintenance goals):

“something bad will not happen”
“something good will always hold”

Safety (maintenance goals):

"something bad will not happen"
"something good will always hold"

Typical example:

$\Box\neg\mathsf{bankrupt}$

Safety (maintenance goals):

"something bad will not happen"
"something good will always hold"

Typical example:

$\Box\neg$bankrupt

Usually: $\Box\neg\ldots$

Safety (maintenance goals):

"something bad will not happen"
"something good will always hold"

Typical example:

$\Box\neg$bankrupt

Usually: $\Box\neg\ldots$

In ATL:

$\langle\langle os \rangle\rangle\Box\neg$crash

Liveness (achievement goals):

“something good will happen”

Liveness (achievement goals):

"something good will happen"

Typical example:

$\Diamond$rich

Usually: $\Diamond \ldots$

Liveness (achievement goals):

“something good will happen”

Typical example:

$\Diamond$rich

Usually: $\Diamond \dots$

In ATL:

$\langle\langle alice, bob\rangle\rangle\Diamond$paperAccepted

Fairness (service goals):

> *"if something is attempted/requested, then it will be successful/allocated"*

Fairness (service goals):

> *"if something is attempted/requested, then it will be successful/allocated"*

Typical examples:

$$\Box(\mathsf{attempt} \;\rightarrow\; \Diamond\mathsf{success})$$

$$\Box\Diamond\mathsf{attempt} \;\rightarrow\; \Box\Diamond\mathsf{success}$$

Fairness (service goals):

"if something is attempted/requested, then it will be successful/allocated"

Typical examples:

$\Box(\mathsf{attempt} \rightarrow \Diamond\mathsf{success})$

$\Box\Diamond\mathsf{attempt} \rightarrow \Box\Diamond\mathsf{success}$

In ATL* (!):

$$\langle\langle prod, dlr\rangle\rangle\Box(\mathsf{carRequested} \rightarrow \Diamond\mathsf{carDelivered})$$

Connection to Multi-Agent/Multi-Process Systems

- Validity $\rightleftharpoons$ General properties of systems

Connection to Multi-Agent/Multi-Process Systems

- Validity $\rightleftharpoons$ General properties of systems
- Satisfiability $\rightleftharpoons$ System synthesis

Connection to Multi-Agent/Multi-Process Systems

- Validity $\rightleftharpoons$ General properties of systems
- Satisfiability $\rightleftharpoons$ System synthesis
- Model checking $\rightleftharpoons$ Verification

Connection to Multi-Agent/Multi-Process Systems

- Validity $\rightleftharpoons$ General properties of systems
- Satisfiability $\rightleftharpoons$ System synthesis
- Model checking $\rightleftharpoons$ Verification

ATL is just another specification language in this context...

Connection to Games

- Concurrent game structure = generalized extensive game

Connection to Games

- Concurrent game structure = generalized extensive game

- $\langle\langle A\rangle\rangle\gamma$: $\langle\langle A\rangle\rangle$ splits the agents into proponents and opponents
- γ defines the winning condition

Connection to Games

- Concurrent game structure = generalized extensive game
- $\langle\langle A\rangle\rangle\gamma$: $\langle\langle A\rangle\rangle$ splits the agents into proponents and opponents
- γ defines the winning condition
  $\rightsquigarrow$ infinite 2-player, binary, zero-sum game

Connection to Games

- Concurrent game structure = generalized extensive game
- $\langle\langle A \rangle\rangle \gamma$: $\langle\langle A \rangle\rangle$ splits the agents into proponents and opponents
- γ defines the winning condition
  $\rightsquigarrow$ infinite 2-player, binary, zero-sum game
- Flexible and compact specification of winning conditions

Connection to Games

- Concurrent game structure = generalized extensive game
- $\langle\langle A \rangle\rangle\gamma$: $\langle\langle A \rangle\rangle$ splits the agents into proponents and opponents
- γ defines the winning condition
  $\rightsquigarrow$ infinite 2-player, binary, zero-sum game
- Flexible and compact specification of winning conditions
- Solving a game $\approx$ checking if $M, q \models \langle\langle A \rangle\rangle\gamma$
- Model checking ATL corresponds to game solving in game theory!

Connection to Games

What about other problems?

Connection to Games

What about other problems?

- Validity $\rightleftharpoons$ General properties of games

Connection to Games

What about other problems?

- Validity $\rightleftharpoons$ General properties of games
- Satisfiability $\rightleftharpoons$ Mechanism design

Connection to Games

What about other problems?

- Validity $\rightleftharpoons$ General properties of games
- Satisfiability $\rightleftharpoons$ Mechanism design
- E.g., building a model for $\langle\langle\emptyset\rangle\rangle\gamma_1 \wedge \langle\langle A\rangle\rangle\gamma_1 \rightleftharpoons$ Designing a game in which γ_1 is guaranteed and A can achieve γ_2

3.3 A Short Look at Satisfiability

Satisfiability of Temporal and Strategic Logics: Complexity Results

	l
CTL	EXPTIME-complete
LTL	PSPACE-complete
CTL*	2EXPTIME-complete
ATL	
ATL*	

Satisfiability of Temporal and Strategic Logics: Complexity Results

	l
CTL	EXPTIME-complete
LTL	PSPACE-complete
CTL*	2EXPTIME-complete
ATL	EXPTIME-complete
ATL*	?

Satisfiability of Temporal and Strategic Logics: Complexity Results

	l
CTL	EXPTIME-complete
LTL	PSPACE-complete
CTL*	2EXPTIME-complete
ATL	EXPTIME-complete
ATL*	?

For strategies with perfect recall:

	m, l
ATL	EXPTIME-complete
ATL*	2EXPTIME-complete

For the Interested Ones...

- Valentin Goranko and Govert van Drimmelen: Decidability and Complete Axiomatization of the Alternating-time Temporal Logic, Theoretical Computer Science, Vol. 353, 1-3, (2006), pp. 93-117.
- D. Walther, C. Lutz, F. Wolter, and M. Wooldridge. ATL Satisfiability is Indeed ExpTime-Complete. In Journal of Logic and Computation, 16:765-787, 2006.
- S. Schewe: ATL* Satisfiability is 2EXPTIME-Complete. Proceedings of ICALP, 2008.

3.4 References

[Alur et al. 2002] R. Alur, T. A. Henzinger, and O. Kupferman.
Alternating-time Temporal Logic.
Journal of the ACM, 49:672–713, 2002.

[Kupferman et al. 2000] O. Kupferman, M.Y. Vardi, and P. Wolper.
An automata-theoretic approach to branching-time model checking.
Journal of the ACM, 47(2):312–360, 2000.

More about ATL

4.1 Axiomatisation

Sound and Compl. Ax. (Goranko, van Drimmelen)

($\bot$) $\neg\langle\langle C\rangle\rangle\bigcirc\bot$

($\top$) $\langle\langle C\rangle\rangle\bigcirc\top$

(Σ) $\neg\langle\langle\emptyset\rangle\rangle\bigcirc\neg\varphi \rightarrow \langle\langle\Sigma\rangle\rangle\bigcirc\varphi$

(S) $\langle\langle C_1\rangle\rangle\bigcirc\varphi_1 \wedge \langle\langle C_2\rangle\rangle\bigcirc\varphi_2 \longrightarrow \langle\langle C_1 \cup C_2\rangle\rangle\bigcirc(\varphi_1 \wedge \varphi_2)$, where C_1 and C_2 are disjoint

$$\frac{\varphi_1, \varphi_1 \rightarrow \varphi_2}{\varphi_2}(MP) \qquad \frac{\varphi_1 \rightarrow \varphi_2}{\langle\langle C\rangle\rangle\bigcirc\varphi_1 \rightarrow \langle\langle C\rangle\rangle\bigcirc\varphi_2}(Mon)$$

Sound and Compl. Ax. (Goranko, van Drimmelen)

$(\bot)$ $\neg\langle\langle C\rangle\rangle\bigcirc\bot$

$(\top)$ $\langle\langle C\rangle\rangle\bigcirc\top$

(Σ) $\neg\langle\langle\emptyset\rangle\rangle\bigcirc\neg\varphi \rightarrow \langle\langle\Sigma\rangle\rangle\bigcirc\varphi$

(S) $\langle\langle C_1\rangle\rangle\bigcirc\varphi_1 \wedge \langle\langle C_2\rangle\rangle\bigcirc\varphi_2 \rightarrow \langle\langle C_1 \cup C_2\rangle\rangle\bigcirc(\varphi_1 \wedge \varphi_2)$, where C_1 and C_2 are disjoint

$(FP_\Box)$ $\langle\langle C\rangle\rangle\Box\varphi \leftrightarrow \varphi \wedge \langle\langle C\rangle\rangle\bigcirc\langle\langle C\rangle\rangle\Box\varphi$

$$\frac{\varphi_1, \varphi_1 \rightarrow \varphi_2}{\varphi_2}(MP) \quad \frac{\varphi_1 \rightarrow \varphi_2}{\langle\langle C\rangle\rangle\bigcirc\varphi_1 \rightarrow \langle\langle C\rangle\rangle\bigcirc\varphi_2}(Mon) \quad \frac{\varphi}{\langle\langle\emptyset\rangle\rangle\Box\varphi}(Nec)$$

Sound and Compl. Ax. (Goranko, van Drimmelen)

($\bot$) $\neg\langle\langle C\rangle\rangle\bigcirc\bot$

($\top$) $\langle\langle C\rangle\rangle\bigcirc\top$

(Σ) $\neg\langle\langle\emptyset\rangle\rangle\bigcirc\neg\varphi \rightarrow \langle\langle\Sigma\rangle\rangle\bigcirc\varphi$

(S) $\langle\langle C_1\rangle\rangle\bigcirc\varphi_1 \wedge \langle\langle C_2\rangle\rangle\bigcirc\varphi_2 \rightarrow \langle\langle C_1 \cup C_2\rangle\rangle\bigcirc(\varphi_1 \wedge \varphi_2)$, where C_1 and C_2 are disjoint

($FP_\Box$) $\langle\langle C\rangle\rangle\Box\varphi \leftrightarrow \varphi \wedge \langle\langle C\rangle\rangle\bigcirc\langle\langle C\rangle\rangle\Box\varphi$

($GFP_\Box$) $\langle\langle\emptyset\rangle\rangle\Box(\theta \rightarrow (\varphi \wedge \langle\langle C\rangle\rangle\bigcirc\theta)) \rightarrow \langle\langle\emptyset\rangle\rangle\Box(\theta \rightarrow \langle\langle C\rangle\rangle\Box\varphi)$

$$\frac{\varphi_1, \varphi_1 \rightarrow \varphi_2}{\varphi_2}(MP) \qquad \frac{\varphi_1 \rightarrow \varphi_2}{\langle\langle C\rangle\rangle\bigcirc\varphi_1 \rightarrow \langle\langle C\rangle\rangle\bigcirc\varphi_2}(Mon) \qquad \frac{\varphi}{\langle\langle\emptyset\rangle\rangle\Box\varphi}(Nec)$$

Sound and Compl. Ax. (Goranko, van Drimmelen)

($\bot$) $\neg\langle\langle C\rangle\rangle\bigcirc\bot$

($\top$) $\langle\langle C\rangle\rangle\bigcirc\top$

(Σ) $\neg\langle\langle\emptyset\rangle\rangle\bigcirc\neg\varphi \rightarrow \langle\langle\Sigma\rangle\rangle\bigcirc\varphi$

(S) $\langle\langle C_1\rangle\rangle\bigcirc\varphi_1 \wedge \langle\langle C_2\rangle\rangle\bigcirc\varphi_2 \rightarrow \langle\langle C_1 \cup C_2\rangle\rangle\bigcirc(\varphi_1 \wedge \varphi_2)$, where C_1 and C_2 are disjoint

($FP_\Box$) $\langle\langle C\rangle\rangle\Box\varphi \leftrightarrow \varphi \wedge \langle\langle C\rangle\rangle\bigcirc\langle\langle C\rangle\rangle\Box\varphi$

($GFP_\Box$) $\langle\langle\emptyset\rangle\rangle\Box(\theta \rightarrow (\varphi \wedge \langle\langle C\rangle\rangle\bigcirc\theta)) \rightarrow \langle\langle\emptyset\rangle\rangle\Box(\theta \rightarrow \langle\langle C\rangle\rangle\Box\varphi)$

($FP_\mathcal{U}$) $\langle\langle C\rangle\rangle(\varphi_1\,\mathcal{U}\,\varphi_2) \leftrightarrow \varphi_2 \vee (\varphi_1 \wedge \langle\langle C\rangle\rangle\bigcirc\langle\langle C\rangle\rangle(\varphi_1\,\mathcal{U}\,\varphi_2))$

$$\frac{\varphi_1, \varphi_1 \rightarrow \varphi_2}{\varphi_2}(MP) \qquad \frac{\varphi_1 \rightarrow \varphi_2}{\langle\langle C\rangle\rangle\bigcirc\varphi_1 \rightarrow \langle\langle C\rangle\rangle\bigcirc\varphi_2}(Mon) \qquad \frac{\varphi}{\langle\langle\emptyset\rangle\rangle\Box\varphi}(Nec)$$

Sound and Compl. Ax. (Goranko, van Drimmelen)

$(\bot)$ $\neg\langle\langle C\rangle\rangle\bigcirc\bot$

$(\top)$ $\langle\langle C\rangle\rangle\bigcirc\top$

(Σ) $\neg\langle\langle\emptyset\rangle\rangle\bigcirc\neg\varphi\rightarrow\langle\langle\Sigma\rangle\rangle\bigcirc\varphi$

(S) $\langle\langle C_1\rangle\rangle\bigcirc\varphi_1\wedge\langle\langle C_2\rangle\rangle\bigcirc\varphi_2\rightarrow\langle\langle C_1\cup C_2\rangle\rangle\bigcirc(\varphi_1\wedge\varphi_2)$, where C_1 and C_2 are disjoint

$(FP_\Box)$ $\langle\langle C\rangle\rangle\Box\varphi\leftrightarrow\varphi\wedge\langle\langle C\rangle\rangle\bigcirc\langle\langle C\rangle\rangle\Box\varphi$

$(GFP_\Box)$ $\langle\langle\emptyset\rangle\rangle\Box(\theta\rightarrow(\varphi\wedge\langle\langle C\rangle\rangle\bigcirc\theta))\rightarrow\langle\langle\emptyset\rangle\rangle\Box(\theta\rightarrow\langle\langle C\rangle\rangle\Box\varphi)$

$(FP_\mathcal{U})$ $\langle\langle C\rangle\rangle(\varphi_1\,\mathcal{U}\,\varphi_2)\leftrightarrow\varphi_2\vee(\varphi_1\wedge\langle\langle C\rangle\rangle\bigcirc\langle\langle C\rangle\rangle(\varphi_1\,\mathcal{U}\,\varphi_2))$

$(LFP_\mathcal{U})$ $\langle\langle\emptyset\rangle\rangle\Box((\varphi_2\vee(\varphi_1\wedge\langle\langle C\rangle\rangle\bigcirc\theta))\rightarrow\theta)\rightarrow\langle\langle\emptyset\rangle\rangle\Box(\langle\langle C\rangle\rangle(\varphi_1\,\mathcal{U}\,\varphi_2)\rightarrow\theta)$

$$\frac{\varphi_1,\ \varphi_1\rightarrow\varphi_2}{\varphi_2}(MP)\qquad\frac{\varphi_1\rightarrow\varphi_2}{\langle\langle C\rangle\rangle\bigcirc\varphi_1\rightarrow\langle\langle C\rangle\rangle\bigcirc\varphi_2}(Mon)\qquad\frac{\varphi}{\langle\langle\emptyset\rangle\rangle\Box\varphi}(Nec)$$

4.2 Bisimulation and The Role of Memory

Definitions

When $\vec{a}_C \in D(q, C)$ let

$$next_{\mathcal{M}}(q, \vec{a}_C) = \{\delta(q, \vec{b}) : \vec{b} \in D(q), a_i = b_i \text{ for all } i \in C\}$$

denote the set of possible next states in CGS $\mathcal{M}$ when coalition C choose actions $\vec{a}_C$.

Definition 4.1 (Bisimulation)

Given CGS $\mathcal{M}_1 = (Q_1, \pi_1, \mathsf{act}_1, d_1, \delta_1)$; CGS $\mathcal{M}_2 = (Q_2, \pi_2, \mathsf{act}_2, d_2, \delta_2)$; $\beta \subseteq Q_1 \times Q_2$.

$\mathcal{M}_1 \rightleftarrows_\beta^C \mathcal{M}_2$ (for some $C \subseteq \Sigma$): for all q_1, q_2, $q_1 \beta q_2$ implies that

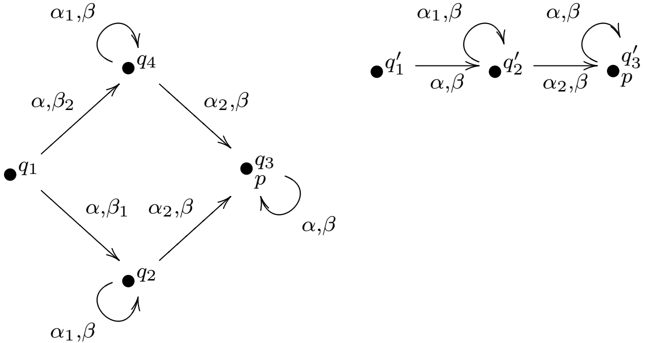
Local harmony $\pi_1(q_1) = \pi_2(q_2)$;

Forth For all joint actions $\vec{a}_C^1 \in D_1(q_1, C)$ for C, there exists a joint action $\vec{a}_C^2 \in D_2(q_2, C)$ for C such that for all states $s_2 \in next_{\mathcal{M}_2}(q_2, \vec{a}_C^2)$, there exists a state $s_1 \in next_{\mathcal{M}_1}(q_1, \vec{a}_C^1)$ such that $s_1 \beta s_2$;

Back Likewise, for 1 and 2 swapped.

$\mathcal{M}_1 \rightleftarrows_\beta \mathcal{M}_2$: $\mathcal{M}_1 \rightleftarrows_\beta^C \mathcal{M}_2$ for every $C \subseteq \Sigma$

Bisimulation: Example



$$\beta = \{(q_1, q_1'), (q_2, q_2'), (q_4, q_2'), (q_3, q_3')\}$$

Strategies and Memory

Let us discern between two definitions of the satisfaction relation:

$\models_F$: perfect recall is assumed, all strategies

$$f : Q^+ \rightarrow \mathsf{act}$$

are allowed

$\models_L$: only memoryless strategies are allowed, i.e., strategies

$$f : Q \rightarrow \mathsf{act}$$

Invariance under Bisimulation: the Memoryless Case

Theorem 4.2 (Bisimulation Characterisation)

If $\mathcal{M}_1 \rightleftarrows_\beta \mathcal{M}_2$ and $s_1 \beta s_2$, then for every ATL formula φ:

$$\mathcal{M}_1, s_1 \models_L \varphi \quad \textit{iff} \quad \mathcal{M}_2, s_2 \models_L \varphi$$

Tree-unfolding

Let $fincomp_M(q)$ denote the set of finite prefixes of computations starting in q. Let $\ell(q_0 \cdots q_k) = q_k$.

Definition 4.3 (Tree-unfolding of CGS)

Given a CGS

$$M = (Q, \pi, \mathsf{act}, d, \delta)$$

and $q \in Q$, the tree-unfolding $T(M,q)$ of M from q is defined as follows:

$$T(M,q) = (Q^*, \pi^*, \mathsf{act}, d^*, \delta^*),$$

where $Q^* = fincomp_M(q)$; $\pi^*(\sigma) = \pi(\ell(\sigma))$; $d_i^*(\sigma) = d_i(\ell(\sigma))$; and $\delta^*(\sigma, \mathbf{a}) = \sigma\delta(\ell(\sigma), \mathbf{a})$.

Lemma 4.4

For any $\mathcal{M}$, q,

$$T(\mathcal{M}, q) \rightleftarrows_{\beta} \mathcal{M}$$

where $\beta = \{(\sigma, \ell(\sigma)) \mid \sigma \in fincomp_{\mathcal{M}}(q)\}$

Lemma 4.5

For any $\mathcal{M}$, q and φ,

$$T(M,q), q \models_L \varphi \Leftrightarrow M, q \models_F \varphi$$

Memory Does not Influence Ability

Corollary 4.6

For any $\mathcal{M}$, q and φ,

$$\mathcal{M}, q \models_L \varphi \Leftrightarrow \mathcal{M}, q \models_F \varphi$$

Memory Does not Influence Ability

Corollary 4.6

For any $\mathcal{M}$, q and φ,

$$\mathcal{M}, q \models_L \varphi \Leftrightarrow \mathcal{M}, q \models_F \varphi$$

Also: the axiomatisation is sound and complete wrt. both semantics.

Invariance under Bisimulation: the Perfect Recall Case

Corollary 4.7

If $\mathcal{M}_1 \rightleftarrows_\beta \mathcal{M}_2$ *and* $s_1 \beta s_2$*, then*

$$\mathcal{M}_1, s_1 \models_F \varphi \textit{ iff } \mathcal{M}_2, s_2 \models_F \varphi$$

for every ATL formula φ*.*

ATL* and memory

For ATL* – contrary to ATL — memory matters:

ATL* and memory

For ATL* – contrary to ATL — memory matters:

Proposition

There is a model M with a state q, and a formula φ, such that

$$M, q \models_L \varphi \not\Leftrightarrow M, q \models_F \varphi$$

ATL* and memory

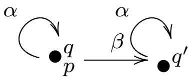
For ATL* – contrary to ATL – memory matters:

Proposition

There is a model M with a state q, and a formula φ, such that

$$M, q \models_L \varphi \not\Leftrightarrow M, q \models_F \varphi$$

M:



$$\varphi = \langle\langle a \rangle\rangle(\bigcirc p \wedge \bigcirc\bigcirc \neg p)$$

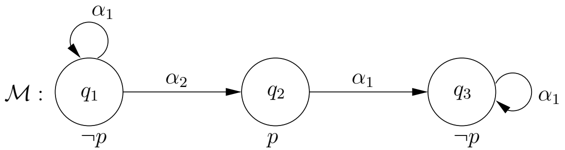
4.3 Revocability of Strategies

Example

- p: agent a controls the resource
- $\langle\langle a\rangle\rangle\bigcirc p$: a has the ability to control the resource next
- $\langle\langle a\rangle\rangle\Box\langle\langle a\rangle\rangle\bigcirc p$: a has the ability to ensure that $\langle\langle a\rangle\rangle\bigcirc p$ is always true

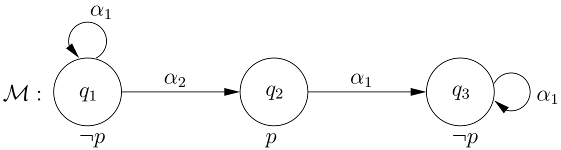
Example

- p: agent a controls the resource
- $\langle\langle a\rangle\rangle\bigcirc p$: a has the ability to control the resource next
- $\langle\langle a\rangle\rangle\Box\langle\langle a\rangle\rangle\bigcirc p$: a has the ability to ensure that $\langle\langle a\rangle\rangle\bigcirc p$ is always true



Example

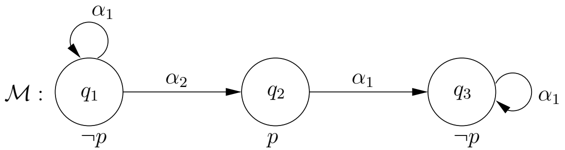
- p: agent a controls the resource
- $\langle\langle a\rangle\rangle\bigcirc p$: a has the ability to control the resource next
- $\langle\langle a\rangle\rangle\Box\langle\langle a\rangle\rangle\bigcirc p$: a has the ability to ensure that $\langle\langle a\rangle\rangle\bigcirc p$ is always true



$$\mathcal{M}, q_1 \models \langle\langle a\rangle\rangle\bigcirc p$$

Example

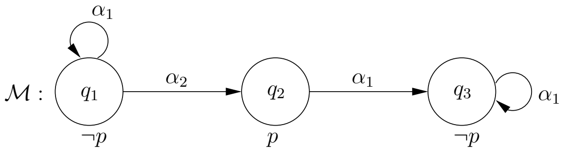
- p: agent a controls the resource
- $\langle\langle a\rangle\rangle\bigcirc p$: a has the ability to control the resource next
- $\langle\langle a\rangle\rangle\Box\langle\langle a\rangle\rangle\bigcirc p$: a has the ability to ensure that $\langle\langle a\rangle\rangle\bigcirc p$ is always true



$$\mathcal{M}, q_1 \models \langle\langle a\rangle\rangle\bigcirc p$$
$$\mathcal{M}, q_1 \models \langle\langle a\rangle\rangle\Box\langle\langle a\rangle\rangle\bigcirc p$$

Example

- p: agent a controls the resource
- $\langle\langle a\rangle\rangle\bigcirc p$: a has the ability to control the resource next
- $\langle\langle a\rangle\rangle\Box\langle\langle a\rangle\rangle\bigcirc p$: a has the ability to ensure that $\langle\langle a\rangle\rangle\bigcirc p$ is always true



$$\mathcal{M}, q_1 \models \langle\langle a\rangle\rangle\bigcirc p$$
$$\mathcal{M}, q_1 \models \langle\langle a\rangle\rangle\Box\langle\langle a\rangle\rangle\bigcirc p$$

Counterintuitive? a can ensure that she is forever able to access the resource – but only without never actually accessing it.

- In the evaluation of a formula such as $\langle\langle a\rangle\rangle\Box\varphi$, when the goal φ is evaluated the agent (a) is no longer restricted by the strategy she chose in order to get to the state where the goal is evaluated (as the example illustrates)
- In this sense, strategies in ATL are revocable
- In some contexts, it would be more natural to reason about strategies which are not revocable and completely specify the future behaviour of the agent

Alternative: Irrevocable Strategies

Irrevocable strategies can be modelled by using model updates in the semantics.

Alternative: Irrevocable Strategies

Irrevocable strategies can be modelled by using model updates in the semantics.
Assume memoryless strategies (for now).

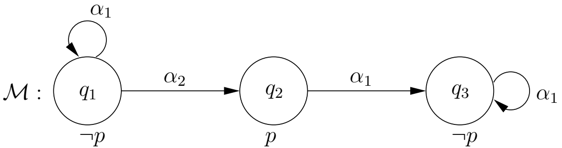
Definition 4.8 (Model Update)

Let $\mathcal{M}$ be a CGS, C a coalition, and f_C a memoryless strategy for C. The update of $\mathcal{M}$ by f_C, denoted $\mathcal{M} \dagger f_C$, is the same as $\mathcal{M}$, except that the choices of each agent $i \in C$ are fixed by the strategy f_i:

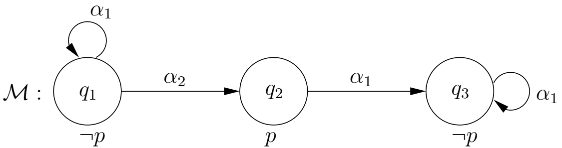
$$d_i(q) = \{f_i(q)\}$$

for each state q.

Model Update: Example

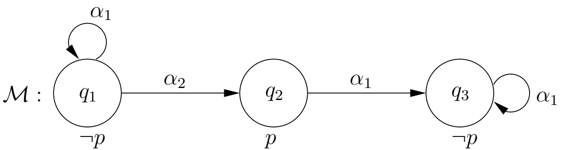


Model Update: Example

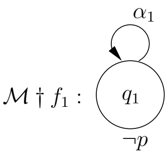


$$f_1 = \{q_1 \mapsto \alpha_1, q_2 \mapsto \alpha_1, q_3 \mapsto \alpha_1\}$$

Model Update: Example



$$f_1 = \{q_1 \mapsto \alpha_1, q_2 \mapsto \alpha_1, q_3 \mapsto \alpha_1\}$$

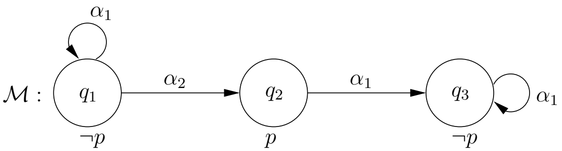


Satisfiability under Irrevocable Strategies

We can now define a new variant of the satisfiability relation:

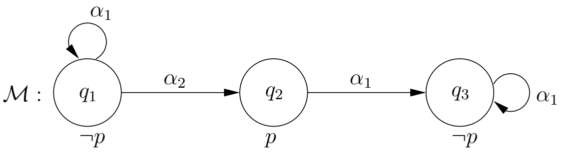
$$
\begin{array}{lcl}
\mathcal{M}, q \models_i \langle\langle C\rangle\rangle \bigcirc \phi & \Leftrightarrow & \exists f_C \forall \lambda \in comp(\mathcal{M} \dagger f_C, q, f_C) \\
 & & (\mathcal{M} \dagger f_C, \lambda[1] \models_i \phi) \\
\mathcal{M}, q \models_i \langle\langle C\rangle\rangle \Box \phi & \Leftrightarrow & \exists f_C \forall \lambda \in comp(\mathcal{M} \dagger f_C, q, f_C) \\
 & & \forall j \geq 0 (\mathcal{M} \dagger f_C, \lambda[j] \models_i \phi) \\
\mathcal{M}, q \models_i \langle\langle C\rangle\rangle (\phi_1 \mathcal{U} \phi_2) & \Leftrightarrow & \exists f_C \forall \lambda \in comp(\mathcal{M} \dagger f_C, q, f_C) \\
 & & \exists j \geq 0 (\mathcal{M} \dagger f_C, \lambda[j] \models_i \phi_2 \text{ and} \\
 & & \forall 0 \leq k < j (\mathcal{M} \dagger f_C, \lambda[k] \models_i \phi_1))
\end{array}
$$

Example (contd.)

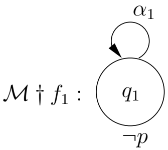


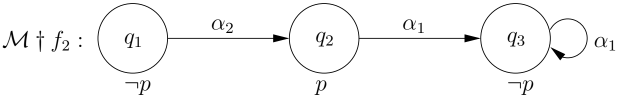
$$f_1 = \{q_1 \mapsto \alpha_1, q_2 \mapsto \alpha_1, q_3 \mapsto \alpha_1\}$$
$$f_2 = \{q_1 \mapsto \alpha_2, q_2 \mapsto \alpha_1, q_3 \mapsto \alpha_1\}$$

Example (contd.)

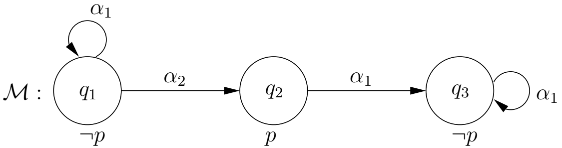


$$f_1 = \{q_1 \mapsto \alpha_1, q_2 \mapsto \alpha_1, q_3 \mapsto \alpha_1\}$$
$$f_2 = \{q_1 \mapsto \alpha_2, q_2 \mapsto \alpha_1, q_3 \mapsto \alpha_1\}$$

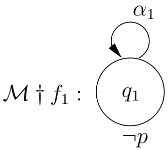


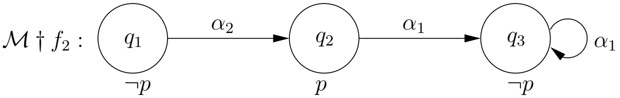


Example (contd.)



$$f_1 = \{q_1 \mapsto \alpha_1, q_2 \mapsto \alpha_1, q_3 \mapsto \alpha_1\}$$
$$f_2 = \{q_1 \mapsto \alpha_2, q_2 \mapsto \alpha_1, q_3 \mapsto \alpha_1\}$$

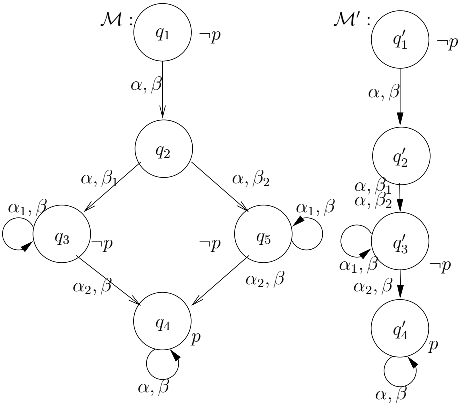




$\mathcal{M}, q_1 \models \langle\langle a \rangle\rangle \Box \langle\langle a \rangle\rangle \bigcirc p$ (standard definition)

$\mathcal{M}, q_1 \not\models_i \langle\langle a \rangle\rangle \Box \langle\langle a \rangle\rangle \bigcirc p$ (with irrevocable strategies)

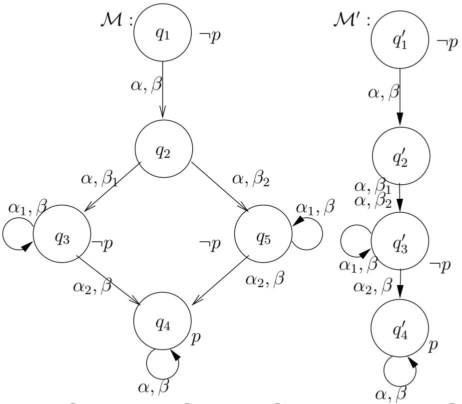
With irrevocable strategies, truth of formulae is not invariant under bisimulations:



$$\mathcal{M}, q_1 \models_i \langle\langle 1 \rangle\rangle \bigcirc ((\langle\langle 2 \rangle\rangle \bigcirc \langle\langle \emptyset \rangle\rangle \bigcirc \neg p) \wedge \langle\langle 2 \rangle\rangle \bigcirc \langle\langle \emptyset \rangle\rangle \bigcirc p)$$

(strategies: $\{q_3 \mapsto \alpha_1, q_5 \mapsto \alpha_2\}$; $\{q_2 \mapsto \beta_1\}$; $\{q_2 \mapsto \beta_2\}$)

With irrevocable strategies, truth of formulae is not invariant under bisimulations:



$$\mathcal{M}, q_1 \models_i \langle\langle 1\rangle\rangle \bigcirc ((\langle\langle 2\rangle\rangle \bigcirc \langle\langle \emptyset\rangle\rangle \bigcirc \neg p) \wedge \langle\langle 2\rangle\rangle \bigcirc \langle\langle \emptyset\rangle\rangle \bigcirc p)$$

(strategies: $\{q_3 \mapsto \alpha_1, q_5 \mapsto \alpha_2\}$; $\{q_2 \mapsto \beta_1\}$; $\{q_2 \mapsto \beta_2\}$)

$$\mathcal{M}', q_1 \not\models_i \langle\langle 1\rangle\rangle \bigcirc ((\langle\langle 2\rangle\rangle \bigcirc \langle\langle \emptyset\rangle\rangle \bigcirc \neg p) \wedge \langle\langle 2\rangle\rangle \bigcirc \langle\langle \emptyset\rangle\rangle \bigcirc p)$$

On Valid Reasoning about Irrevocable Strategies

- Formulae valid under the standard definition is not necessarily valid under irrevocable strategies. For example, the principle of uniform substitution does not hold. The ATL axiom

$$\neg\langle\langle\emptyset\rangle\rangle\bigcirc\neg p \rightarrow \langle\langle\Sigma\rangle\rangle\bigcirc p$$

is still valid with irrevocable strategies, but the result of substituting

$$\langle\langle\Sigma\rangle\rangle\bigcirc p \wedge \langle\langle\Sigma\rangle\rangle\bigcirc\neg p$$

for p in it is not valid.

On Valid Reasoning about Irrevocable Strategies

- Formulae valid under the standard definition is not necessarily valid under irrevocable strategies. For example, the principle of uniform substitution does not hold. The ATL axiom

$$\neg\langle\langle\emptyset\rangle\rangle\bigcirc\neg p \rightarrow \langle\langle\Sigma\rangle\rangle\bigcirc p$$

  is still valid with irrevocable strategies, but the result of substituting

$$\langle\langle\Sigma\rangle\rangle\bigcirc p \wedge \langle\langle\Sigma\rangle\rangle\bigcirc\neg p$$

  for p in it is not valid.
- Formulae valid under irrevocable strategies are not necessarily valid under the standard definition. Example:

$$\langle\langle C\rangle\rangle\bigcirc\langle\langle C\rangle\rangle\bigcirc\phi \leftrightarrow \langle\langle C\rangle\rangle\bigcirc\langle\langle\emptyset\rangle\rangle\bigcirc\phi$$

  for $C \neq \emptyset$

Perfect Recall

With perfect recall strategies, we cannot update the model directly. Instead, unwind it first, and recall that a perfect recall strategy in $\mathcal{M}$ is equivalent to a memoryless strategy in $T(\mathcal{M}, q)$:

$$\mathcal{M}, q \models_{mi} \varphi \Leftrightarrow^{def} T(M, q), q \models_i \varphi$$

Perfect Recall

$$\mathcal{M}, q \models_{mi} \varphi \Leftrightarrow^{def} T(M, q), q \models_i \varphi$$

Perfect Recall

$$\mathcal{M}, q \models_{mi} \varphi \Leftrightarrow^{def} T(M, q), q \models_i \varphi$$

We get that:

- Still non-invariant under bisimulation

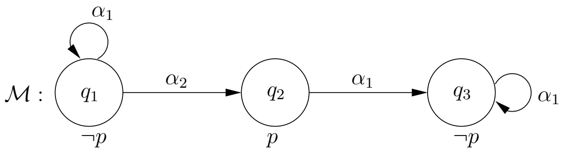
Perfect Recall

$$\mathcal{M}, q \models_{mi} \varphi \Leftrightarrow^{def} T(M, q), q \models_i \varphi$$

We get that:

- Still non-invariant under bisimulation
- With irrevocable strategies (unlike under the standard definition), memory matters:

$$\mathcal{M}, q_1 \models_{mi} \langle\langle a \rangle\rangle \bigcirc \langle\langle a \rangle\rangle \bigcirc p$$
$$\mathcal{M}, q_1 \not\models_i \langle\langle a \rangle\rangle \bigcirc \langle\langle a \rangle\rangle \bigcirc p$$



4.4 References

[1] Goranko, V. and G. van Drimmelen: 2006, 'Complete axiomatization and decidability of Alternating-time Temporal Logic'.
Theoretical Computer Science **353**(1-3), 93–117.

[2] Thomas Ågotnes, Valentin Goranko, and Wojciech Jamroga.
Alternating-time temporal logics with irrevocable strategies.
In Dov Samet, editor, *Proceedings of the 11th Conference on Theoretical Aspects of Rationality and Knowledge (TARK XI)*, pages 15–24, Brussels, Belgium, June 2007. Presses Universitaires de Louvain/ACM DL.

[3] Thomas Brihaye, Arnaud Da Costa, François Laroussinie, and Nicolas Markey.
ATL with strategy contexts.
In preparation, 2008.

Imperfect Information

How can we reason about extensive games with imperfect information?

How can we reason about extensive games with imperfect information?

Let's put ATL and epistemic logic in one box.

- We extend CGS with indistinguishability relations $\sim_a$, one per agent
- We add epistemic operators to ATL

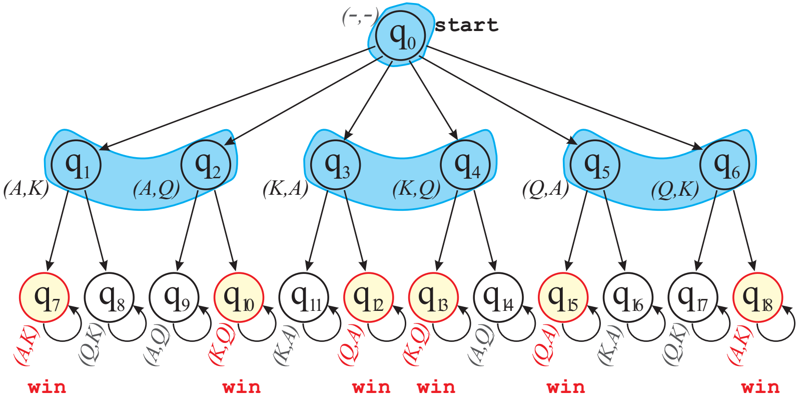
How can we reason about extensive games with imperfect information?

Let's put ATL and epistemic logic in one box.

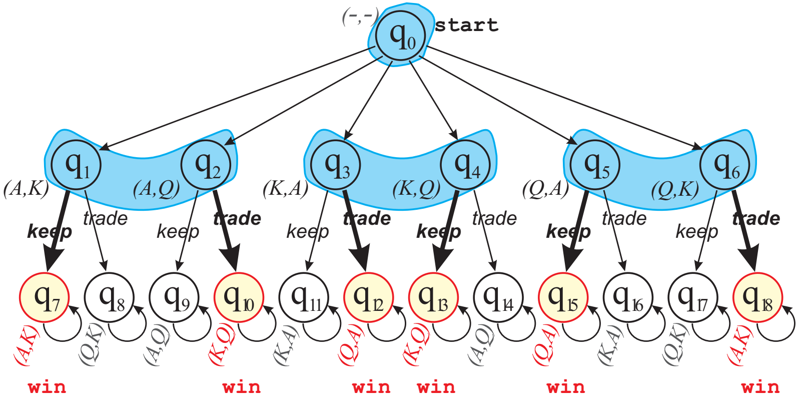
- We extend CGS with indistinguishability relations $\sim_a$, one per agent
- We add epistemic operators to ATL

$\rightsquigarrow$ Problems!

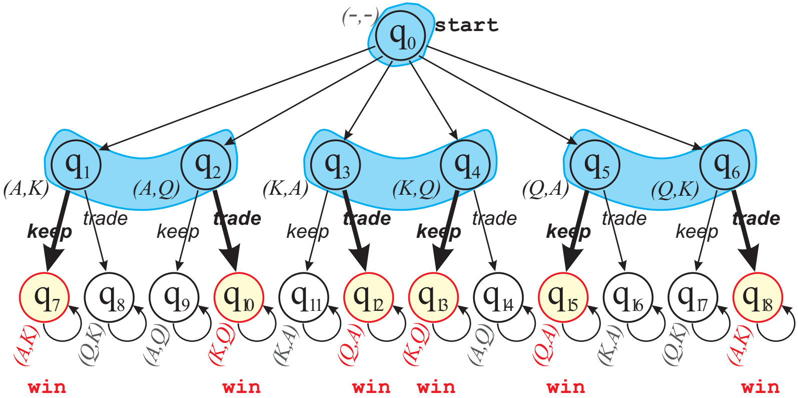
5.1 Combining Dimensions



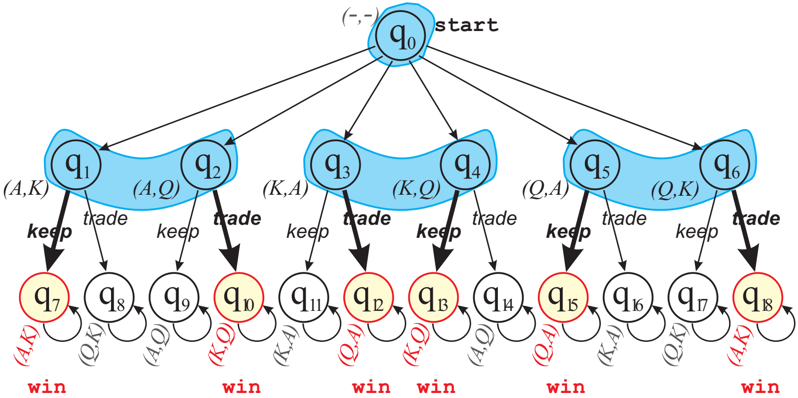
(-,-)
q0
start
q1
q2
q3
q4
q5
q6
(A,K)
(A,Q)
(K,A)
(K,Q)
(Q,A)
(Q,K)
q7
q8
q9
q10
q11
q12
q13
q14
q15
q16
q17
q18
(A,K)
(Q,K)
(A,Q)
(K,Q)
(K,A)
(Q,A)
(K,Q)
(A,Q)
(Q,A)
(K,A)
(Q,K)
(A,K)
win
win
win
win
win
win



(-,-)
q0
start
q1
q2
q3
q4
q5
q6
(A,K)
(A,Q)
(K,A)
(K,Q)
(Q,A)
(Q,K)
keep
trade
keep
trade
keep
trade
keep
trade
keep
trade
keep
trade
q7
q8
q9
q10
q11
q12
q13
q14
q15
q16
q17
q18
(A,K)
(Q,K)
(A,Q)
(K,Q)
(K,A)
(Q,A)
(K,Q)
(A,Q)
(Q,A)
(K,A)
(Q,K)
(A,K)
win
win
win
win
win
win

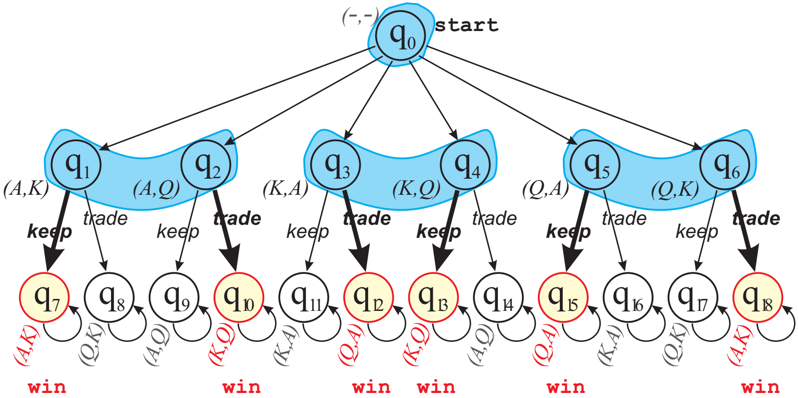


$$start \rightarrow \langle\langle a \rangle\rangle \Diamond win$$



$$start \rightarrow \langle\langle a \rangle\rangle \Diamond win$$

$$start \rightarrow K_a \langle\langle a \rangle\rangle \Diamond win$$



$$start \rightarrow \langle\langle a \rangle\rangle \Diamond win$$

$$start \rightarrow K_a \langle\langle a \rangle\rangle \Diamond win$$

Does it make sense?

Problem:

Strategic and epistemic abilities are not independent!

Problem:

Strategic and epistemic abilities are not independent!

$\langle\langle A \rangle\rangle \Phi$ = A can enforce Φ

Problem:

Strategic and epistemic abilities are not independent!

$\langle\langle A\rangle\rangle\Phi$ = A can enforce Φ

It should at least mean that A are able to identify and execute the right strategy!

Problem:

Strategic and epistemic abilities are not independent!

$\langle\langle A \rangle\rangle \Phi$ = A can enforce Φ

It should at least mean that A are able to identify and execute the right strategy!

Executable strategies = uniform strategies

Definition 5.1 (Uniform strategy)

Strategy s_a is uniform iff it specifies the same choices for indistinguishable situations:

- (no recall:) if $q \sim_a q'$ then $s_a(q) = s_a(q')$
- (perfect recall:) if $\lambda \approx_a \lambda'$ then $\Rightarrow s_a(\lambda) = s_a(\lambda)$, where $\lambda \approx_a \lambda'$ iff $\lambda[i] \sim_a \lambda'[i]$ for every i.

Definition 5.1 (Uniform strategy)

Strategy s_a is uniform iff it specifies the same choices for indistinguishable situations:

- (no recall:) if $q \sim_a q'$ then $s_a(q) = s_a(q')$
- (perfect recall:) if $\lambda \approx_a \lambda'$ then $\Rightarrow s_a(\lambda) = s_a(\lambda)$, where $\lambda \approx_a \lambda'$ iff $\lambda[i] \sim_a \lambda'[i]$ for every i.

A collective strategy is uniform iff it consists only of uniform individual strategies.

Note:

Having a successful strategy does not imply knowing that we have it!

Note:

Having a successful strategy does not imply knowing that we have it!

Knowing that a successful strategy exists does not imply knowing the strategy itself!

Levels of Strategic Ability

Our cases for $\langle\langle A \rangle\rangle \Phi$ under imperfect information:

1. There is σ (not necessarily executable!) such that, for every execution of σ, Φ holds

Levels of Strategic Ability

Our cases for $\langle\langle A \rangle\rangle \Phi$ under imperfect information:

1. There is σ (not necessarily executable!) such that, for every execution of σ, Φ holds
2. There is a uniform σ such that, for every execution of σ, Φ holds

Levels of Strategic Ability

Our cases for $\langle\langle A \rangle\rangle \Phi$ under imperfect information:

1. There is σ (not necessarily executable!) such that, for every execution of σ, Φ holds
2. There is a uniform σ such that, for every execution of σ, Φ holds
3. A know that there is a uniform σ such that, for every execution of σ, Φ holds

Levels of Strategic Ability

Our cases for $\langle\langle A \rangle\rangle \Phi$ under imperfect information:

1. There is σ (not necessarily executable!) such that, for every execution of σ, Φ holds
2. There is a uniform σ such that, for every execution of σ, Φ holds
3. A know that there is a uniform σ such that, for every execution of σ, Φ holds
4. There is a uniform σ such that A know that, for every execution of σ, Φ holds

Levels of Strategic Ability

Our cases for $\langle\langle A\rangle\rangle\Phi$ under imperfect information:

1. There is σ (not necessarily executable!) such that, for every execution of σ, Φ holds
2. There is a uniform σ such that, for every execution of σ, Φ holds
3. A know that there is a uniform σ such that, for every execution of σ, Φ holds
4. There is a uniform σ such that A know that, for every execution of σ, Φ holds

From now on, we restrict our discussion to uniform memoryless strategies (unless explicitly stated otherwise).

Levels of Strategic Ability

Our cases for $\langle\langle A\rangle\rangle\Phi$ under imperfect information:

1. There is σ (not necessarily executable!) such that, for every execution of σ, Φ holds
2. There is a uniform σ such that, for every execution of σ, Φ holds
3. A know that there is a uniform σ such that, for every execution of σ, Φ holds
4. There is a uniform σ such that A know that, for every execution of σ, Φ holds

From now on, we restrict our discussion to uniform memoryless strategies (unless explicitly stated otherwise).

Case [4]: knowing how to play

Case [4]: knowing how to play

- Single agent case: we take into account the paths starting from indistinguishable states (i.e., $\bigcup_{q' \in \mathrm{img}(q, \sim_a)} out(q, s_A)$)

Case [4]: knowing how to play

- Single agent case: we take into account the paths starting from indistinguishable states (i.e., $\bigcup_{q' \in \mathrm{img}(q,\sim_a)} out(q, s_A)$)

- What about coalitions?
- Question: in what sense should they know the strategy? Common knowledge (C_A), mutual knowledge (E_A), distributed knowledge (D_A)?

5.2 Economic Solution: ATL_{ir}

Schobbens' $\mathrm{ATL}_{\mathrm{ir}}$

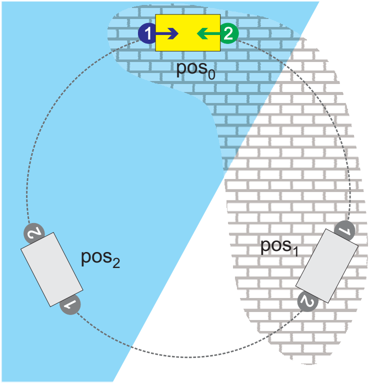
$\langle\langle A\rangle\rangle_{_{ir}}\gamma$: agents A know how to play in the sense of mutual knowledge (E_A)

Schobbens' ATL$_{ir}$

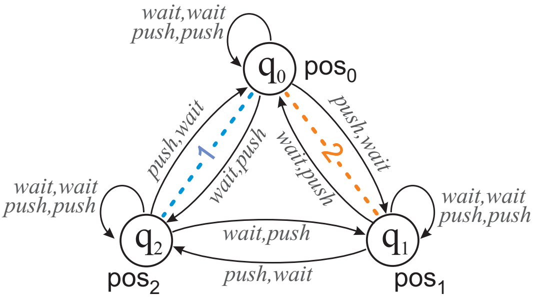
$\langle\langle A\rangle\rangle_{_{ir}}\gamma$: agents A know how to play in the sense of mutual knowledge (E_A)

$M, q \models \langle\langle A\rangle\rangle_{_{ir}}\gamma$ iff there is a collective uniform strategy s_A such that, for every path $\lambda \in \bigcup_{q' \sim_A q} out(q', s_A)$, we have $M, \Lambda \models \gamma$.

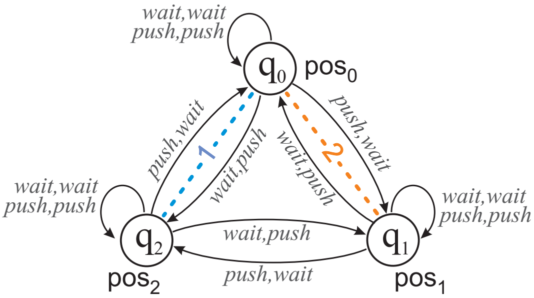
Example: Robots and Carriage



Example: Robots and Carriage



Example: Robots and Carriage



$$\neg(\mathsf{pos}_0 \to \langle\langle s \rangle\rangle_{ir} \Box \mathsf{pos}_0)$$

$$\mathsf{pos}_0 \to \langle\langle s \rangle\rangle_{ir} \Box \neg \mathsf{pos}_1$$

Schobbens' $\mathrm{ATL_{ir}}$

Interesting: $\langle\!\langle A\rangle\!\rangle_{ir}$ are not fixpoint operators any more!

Theorem 5.2

*The following formulae are **not** valid for $\mathrm{ATL_{ir}}$:*

- $\langle\!\langle A\rangle\!\rangle_{ir}\Box\varphi \quad\leftrightarrow\quad \varphi \wedge \langle\!\langle A\rangle\!\rangle_{ir}\bigcirc\langle\!\langle A\rangle\!\rangle_{ir}\Box\varphi$
- $\langle\!\langle A\rangle\!\rangle_{ir}\varphi_1\,\mathcal{U}\,\varphi_2 \quad\leftrightarrow\quad \varphi_2 \vee \varphi_1 \wedge \langle\!\langle A\rangle\!\rangle_{ir}\bigcirc\langle\!\langle A\rangle\!\rangle_{ir}\varphi_1\,\mathcal{U}\,\varphi_2.$

Schobbens' ATL$_{\text{ir}}$

Interesting: $\langle\!\langle A \rangle\!\rangle_{ir}$ are not fixpoint operators any more!

Theorem 5.2

*The following formulae are **not** valid for ATL$_{\text{ir}}$:*

- $\langle\!\langle A \rangle\!\rangle_{ir} \Box \varphi \quad \leftrightarrow \quad \varphi \wedge \langle\!\langle A \rangle\!\rangle_{ir} \bigcirc \langle\!\langle A \rangle\!\rangle_{ir} \Box \varphi$
- $\langle\!\langle A \rangle\!\rangle_{ir} \varphi_1 \mathcal{U} \varphi_2 \quad \leftrightarrow \quad \varphi_2 \vee \varphi_1 \wedge \langle\!\langle A \rangle\!\rangle_{ir} \bigcirc \langle\!\langle A \rangle\!\rangle_{ir} \varphi_1 \mathcal{U} \varphi_2.$

What is it about?

Schobbens' ATL_{ir}

Interesting: $\langle\!\langle A\rangle\!\rangle_{_{ir}}$ are not fixpoint operators any more!

Theorem 5.2

The following formulae are **not** *valid for* ATL_{ir}*:*

- $\langle\!\langle A\rangle\!\rangle_{_{ir}}\Box\varphi \quad\leftrightarrow\quad \varphi \wedge \langle\!\langle A\rangle\!\rangle_{_{ir}}\bigcirc\langle\!\langle A\rangle\!\rangle_{_{ir}}\Box\varphi$
- $\langle\!\langle A\rangle\!\rangle_{_{ir}}\varphi_1\,\mathcal{U}\,\varphi_2 \quad\leftrightarrow\quad \varphi_2 \vee \varphi_1 \wedge \langle\!\langle A\rangle\!\rangle_{_{ir}}\bigcirc\langle\!\langle A\rangle\!\rangle_{_{ir}}\varphi_1\,\mathcal{U}\,\varphi_2.$

What is it about? Forgetting!

Agents Can Forget...

Agents Can Forget... And Still Enforce Things

Schobbens' $\mathrm{ATL}_{\mathrm{ir}}$

Conjecture

Strategy for A cannot be synthesized incrementally.

Schobbens' ATL_{ir}

Conjecture

Strategy for A cannot be synthesized incrementally.

Indeed...

Schobbens' $\mathrm{ATL_{ir}}$

Conjecture

Strategy for A cannot be synthesized incrementally.

Indeed...

Theorem (Schobbens 2004; Jamroga & Dix 2006)

Model checking ATL_{ir} is Δ_2-complete in the number of transitions in the model and the length of the formula.

5.3 Constructive Strategic Logic

Knowing how to Play

- Single agent case: we take into account the paths starting from indistinguishable states $\rightsquigarrow$ ATL_{ir}

- What about coalitions? In what sense should they know the strategy? Common knowledge (C_A), mutual knowledge (E_A), distributed knowledge (D_A)...?
- ATL_{ir}: mutual knowledge
- But: other cases also make sense!

Given strategy σ, agents A can have:

- Common knowledge that σ is a winning strategy. This requires the least amount of additional communication (agents from A may agree upon a total order over their collective strategies at the beginning of the game and that they will always choose the maximal winning strategy with respect to this order)

Given strategy σ, agents A can have:

- Common knowledge that σ is a winning strategy. This requires the least amount of additional communication (agents from A may agree upon a total order over their collective strategies at the beginning of the game and that they will always choose the maximal winning strategy with respect to this order)
- Mutual knowledge that σ is a winning strategy: everybody in A knows that σ is winning

- Distributed knowledge that σ is a winning strategy: if the agents share their knowledge at the current state, they can identify the strategy as winning

- Distributed knowledge that σ is a winning strategy: if the agents share their knowledge at the current state, they can identify the strategy as winning
- "The leader": the strategy can be identified by agent $a \in A$

- Distributed knowledge that σ is a winning strategy: if the agents share their knowledge at the current state, they can identify the strategy as winning
- "The leader": the strategy can be identified by agent $a \in A$
- "Headquarters' committee": the strategy can be identified by subgroup $A' \subseteq A$

- Distributed knowledge that σ is a winning strategy: if the agents share their knowledge at the current state, they can identify the strategy as winning
- "The leader": the strategy can be identified by agent $a \in A$
- "Headquarters' committee": the strategy can be identified by subgroup $A' \subseteq A$
- "Consulting company": the strategy can be identified by some other group B

Many subtle cases...

Many subtle cases...

$\rightsquigarrow$ Solution: constructive knowledge operators

Constructive Strategic Logic (CSL)

- $\langle\langle A\rangle\rangle\Phi$: A have a uniform memoryless strategy to enforce Φ

Constructive Strategic Logic (CSL)

- $\langle\langle A\rangle\rangle\Phi$: A have a uniform memoryless strategy to enforce Φ
- $K_a\langle\langle a\rangle\rangle\Phi$: a has a strategy to enforce Φ, and knows that he has one
- For groups of agents: $C_A, E_A, D_A, \ldots$

Constructive Strategic Logic (CSL)

- $\langle\langle A\rangle\rangle\Phi$: A have a uniform memoryless strategy to enforce Φ
- $K_a\langle\langle a\rangle\rangle\Phi$: a has a strategy to enforce Φ, and knows that he has one
- For groups of agents: $C_A, E_A, D_A, \ldots$
- $\mathbb{K}_a\langle\langle a\rangle\rangle\Phi$: a has a strategy to enforce Φ, and knows that this is a winning strategy
- For groups of agents: $\mathbb{C}_A, \mathbb{E}_A, \mathbb{D}_A, \ldots$

Non-standard semantics:

- Formulae are evaluated in sets of states
- $M, Q \models \langle\langle A \rangle\rangle \gamma$: A have a single strategy to enforce γ from all states in Q

Non-standard semantics:

- Formulae are evaluated in sets of states
- $M, Q \models \langle\langle A\rangle\rangle\gamma$: A have a single strategy to enforce γ from all states in Q

Additionally:

- $out(Q, s_A) = \bigcup_{q \in Q} out(q, s_A)$
- $\mathrm{img}(Q, \mathcal{R}) = \bigcup_{q \in Q} \mathrm{img}(q, \mathcal{R})$

Non-standard semantics:

- Formulae are evaluated in sets of states
- $M, Q \models \langle\langle A \rangle\rangle \gamma$: A have a single strategy to enforce γ from all states in Q

Additionally:

- $out(Q, s_A) = \bigcup_{q \in Q} out(q, s_A)$
- $\mathrm{img}(Q, \mathcal{R}) = \bigcup_{q \in Q} \mathrm{img}(q, \mathcal{R})$
- $M, q \models \varphi$ iff $M, \{q\} \models \varphi$

Definition 5.3 (Semantics of CSL)

$M, Q \models p$ iff $p \in \pi(q)$ for every $q \in Q$;

Definition 5.3 (Semantics of CSL)

$M, Q \models p$ iff $p \in \pi(q)$ for every $q \in Q$;

$M, Q \models \neg\varphi$ iff not $M, Q \models \varphi$;

Definition 5.3 (Semantics of CSL)

$M, Q \models p$ iff $p \in \pi(q)$ for every $q \in Q$;

$M, Q \models \neg\varphi$ iff not $M, Q \models \varphi$;

$M, Q \models \varphi \wedge \psi$ iff $M, Q \models \varphi$ and $M, Q \models \psi$;

Definition 5.3 (Semantics of CSL)

$M, Q \models p$ iff $p \in \pi(q)$ for every $q \in Q$;

$M, Q \models \neg\varphi$ iff not $M, Q \models \varphi$;

$M, Q \models \varphi \wedge \psi$ iff $M, Q \models \varphi$ and $M, Q \models \psi$;

$M, Q \models \langle\!\langle A \rangle\!\rangle \gamma$ iff there exists s_A such that, for every $\lambda \in out(Q, s_A)$, we have that $M, \lambda \models \gamma$;

Definition 5.3 (Semantics of CSL)

$M, Q \models p$ iff $p \in \pi(q)$ for every $q \in Q$;

$M, Q \models \neg\varphi$ iff not $M, Q \models \varphi$;

$M, Q \models \varphi \wedge \psi$ iff $M, Q \models \varphi$ and $M, Q \models \psi$;

$M, Q \models \langle\langle A \rangle\rangle \gamma$ iff there exists s_A such that, for every $\lambda \in out(Q, s_A)$, we have that $M, \lambda \models \gamma$;

$M, Q \models \mathcal{K}_A \varphi$ iff $M, q \models \varphi$ for every $q \in \mathrm{img}(Q, \sim_A^{\mathcal{K}})$ (where $\mathcal{K} = C, E, D$);

Definition 5.3 (Semantics of CSL)

$M, Q \models p$ iff $p \in \pi(q)$ for every $q \in Q$;

$M, Q \models \neg\varphi$ iff not $M, Q \models \varphi$;

$M, Q \models \varphi \wedge \psi$ iff $M, Q \models \varphi$ and $M, Q \models \psi$;

$M, Q \models \langle\langle A \rangle\rangle \gamma$ iff there exists s_A such that, for every $\lambda \in out(Q, s_A)$, we have that $M, \lambda \models \gamma$;

$M, Q \models \mathcal{K}_A \varphi$ iff $M, q \models \varphi$ for every $q \in \mathrm{img}(Q, \sim_A^{\mathcal{K}})$ (where $\mathcal{K} = C, E, D$);

$M, Q \models \hat{\mathcal{K}}_A \varphi$ iff $M, \mathrm{img}(Q, \sim_A^{\mathcal{K}}) \models \varphi$ (where $\hat{\mathcal{K}} = \mathbb{C}, \mathbb{E}, \mathbb{D}$ and $\mathcal{K} = C, E, D$, respectively).

Validity in CSL

- Formula φ is valid iff $M, q \models \varphi$ for all models M and states q
- Formula φ is strongly valid iff for each M and every non-empty set of states Q it is the case that $M, Q \models \varphi$

Validity in CSL

- Formula φ is valid iff $M, q \models \varphi$ for all models M and states q
- Formula φ is strongly valid iff for each M and every non-empty set of states Q it is the case that $M, Q \models \varphi$

Theorem 5.4

1. *Strong validity implies validity.*
2. *Validity does not imply strong validity.*

Validity in CSL

- We are ultimately interested in simple validity

Validity in CSL

- We are ultimately interested in simple validity
- The importance of strong validity, on the other hand, lies in the fact that strong validity of $\varphi \leftrightarrow \psi$ makes φ and ψ completely interchangeable

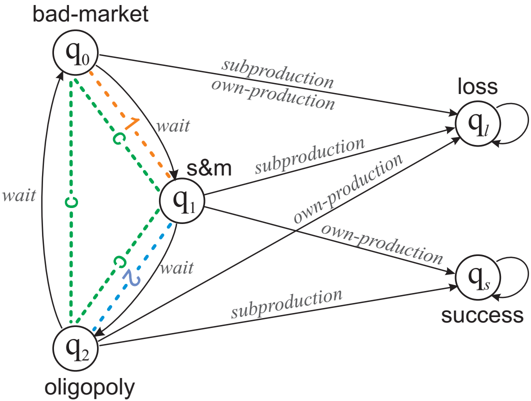
Validity in CSL

- We are ultimately interested in simple validity
- The importance of strong validity, on the other hand, lies in the fact that strong validity of $\varphi \leftrightarrow \psi$ makes φ and ψ completely interchangeable

Theorem 5.5

If $\varphi_1 \leftrightarrow \varphi_2$ is strongly valid, and ψ' is obtained from ψ through replacing an occurrence of φ_1 by φ_2, then $M, Q \models \psi$ iff $M, Q \models \psi'$.

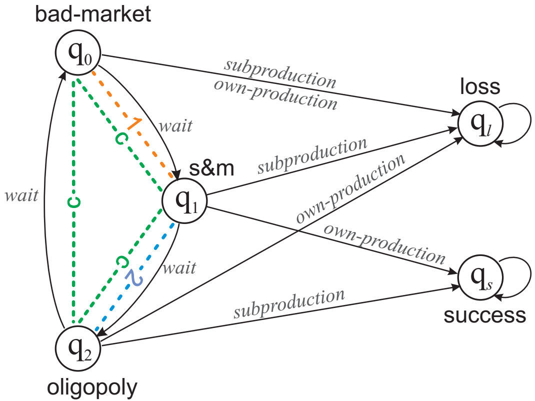
Example: Simple Market



$@\ q_1$:

$\neg \mathbb{K}_c \langle\langle c \rangle\rangle \Diamond$success

Example: Simple Market



$@\ q_1\ :$

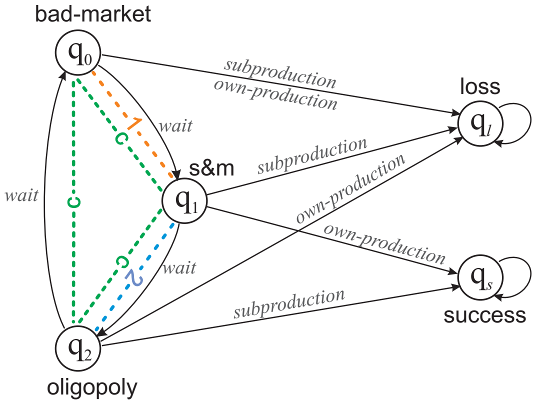
$\neg \mathbb{K}_c \langle\!\langle c \rangle\!\rangle \Diamond$success

$\neg \mathbb{E}_{\{1,2\}} \langle\!\langle c \rangle\!\rangle \Diamond$success

$\neg \mathbb{K}_1 \langle\!\langle c \rangle\!\rangle \Diamond$success

$\neg \mathbb{K}_2 \langle\!\langle c \rangle\!\rangle \Diamond$success

Example: Simple Market



$@\ q_1$:

$\neg \mathbb{K}_c \langle\langle c \rangle\rangle \Diamond$success

$\neg \mathbb{E}_{\{1,2\}} \langle\langle c \rangle\rangle \Diamond$success

$\neg \mathbb{K}_1 \langle\langle c \rangle\rangle \Diamond$success

$\neg \mathbb{K}_2 \langle\langle c \rangle\rangle \Diamond$success

$\mathbb{D}_{\{1,2\}} \langle\langle c \rangle\rangle \Diamond$success

Onion Soup Robbery

A virtual safe contains the recipe for the best onion soup in the world. The safe can only be opened by a k-digit binary code, where each digit c_i is sent from a prescribed location i $(1 \leq i \leq k)$. To open the safe and download the recipe it is enough that at least $n \leq k$ correct digits are sent at the same moment. However, if a wrong value is sent from one of the locations, or if an insufficient number (i.e., between 1 and $n-1$) of digits is submitted, then the safe locks up and activates an alarm. k agents are connected at the right locations; each of them can send 0, send 1, or do nothing (nop). Moreover, individual agents have only partial information about the code: agent i (connected to location i) knows the values of c_{i-1} XOR c_i and c_i XOR c_{i+1} (we take $c_0 = c_{k+1} = 0$). This implies that only agents 1 and k *know the values of "their" digits*. Still, *every agent knows whether his neighbors' digits are the same as his*.

Onion Soup Robbery: Some Properties

For OSR_k^n and the initial state, we have:

- $\neg\mathbb{E}_{\mathbb{A}\mathrm{gt}}\langle\langle\mathbb{A}\mathrm{gt}\rangle\rangle\Diamond\mathsf{open}$: the team cannot identify a winning strategy;

Onion Soup Robbery: Some Properties

For OSR_k^n and the initial state, we have:

- $\neg\mathbb{E}_{\mathbb{A}\mathrm{gt}}\langle\langle\mathbb{A}\mathrm{gt}\rangle\rangle\Diamond\mathsf{open}$: the team cannot identify a winning strategy;
- $\mathbb{D}_{\mathbb{A}\mathrm{gt}}\langle\langle\mathbb{A}\mathrm{gt}\rangle\rangle\Diamond\mathsf{open}$: if the agents share information they can recognize who should send what;

Onion Soup Robbery: Some Properties

For OSR_k^n and the initial state, we have:

- $\neg\mathbb{E}_{\mathbb{A}\mathrm{gt}}\langle\langle\mathbb{A}\mathrm{gt}\rangle\rangle\Diamond\mathsf{open}$: the team cannot identify a winning strategy;
- $\mathbb{D}_{\mathbb{A}\mathrm{gt}}\langle\langle\mathbb{A}\mathrm{gt}\rangle\rangle\Diamond\mathsf{open}$: if the agents share information they can recognize who should send what;
- $\mathbb{D}_{\{1,\dots,n-1\}}\langle\langle\mathbb{A}\mathrm{gt}\rangle\rangle\Diamond\mathsf{open}$: it is enough that the first $n-1$ agents devise the strategy. Note that the same holds for the last $n-1$ agents, i.e., the subteam $\{k-n+2,\dots,k\}$.

Theorem 5.6 (Expressivity)

CSL is strictly more expressive than ATL_{ir}.

Theorem 5.6 (Expressivity)

CSL is strictly more expressive than ATL_{ir}.

Theorem 5.7 (Verification complexity)

The complexity of model checking CSL is the same as for ATL_{ir}.

5.4 Constructive Knowledge

Properties of Constructive Knowledge

Non-standard semantics raises some natural questions:

- Is constructive knowledge... em, well, **knowledge**?

Properties of Constructive Knowledge

Non-standard semantics raises some natural questions:

- Is constructive knowledge... em, well, **knowledge**?
  $\leadsto$ semantic vs. syntactic analysis

Properties of Constructive Knowledge

Non-standard semantics raises some natural questions:

- Is constructive knowledge... em, well, **knowledge**?
  $\rightsquigarrow$ semantic vs. syntactic analysis
- Is constructive knowledge a special kind of standard knowledge? Or the other way around?

Properties of Constructive Knowledge

Non-standard semantics raises some natural questions:

- Is constructive knowledge... em, well, **knowledge**?
  $\rightsquigarrow$ semantic vs. syntactic analysis
- Is constructive knowledge a special kind of standard knowledge? Or the other way around?
- Is there a relevant subset of the language for whom a more standard semantics can be given?

Is $\mathbb{K}_a$ an Epistemic Operator?

Theorem 5.8

Below, we list the constructive knowledge versions of some of the S5 properties for individual agents. "Yes" means that the schema is strongly valid; "No" means that it is not even weakly valid (incidentally, none of the properties turns out to be weakly but not strongly valid).

$\mathbf{K}$	$\mathbb{K}_a(\varphi \to \psi) \to (\mathbb{K}_a\varphi \to \mathbb{K}_a\psi)$	*Yes*
$\mathbf{D}$	$\neg\mathbb{K}_a\bot$	*Yes*
$\mathbf{T}$	$\mathbb{K}_a\varphi \to \varphi$	*No*
$\mathbf{4}$	$\mathbb{K}_a\varphi \to \mathbb{K}_a\mathbb{K}_a\varphi$	*Yes*
$\mathbf{4}^+$	$\mathbb{K}_a\varphi \leftrightarrow \mathbb{K}_a\mathbb{K}_a\varphi$	*Yes*
$\mathbf{5}$	$\neg\mathbb{K}_a\varphi \to \mathbb{K}_a\neg\mathbb{K}_a\varphi$	*Yes*
$\mathbf{5}^+$	$\neg\mathbb{K}_a\varphi \leftrightarrow \mathbb{K}_a\neg\mathbb{K}_a\varphi$	*Yes*
$\mathbf{B}$	$\varphi \to \mathbb{K}_a\neg\mathbb{K}_a\neg\varphi$	*No*

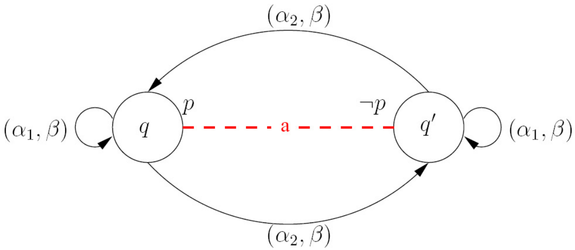
Is $\mathbb{K}_a$ an Epistemic Operator?

Theorem 5.8

Below, we list the constructive knowledge versions of some of the S5 properties for individual agents. "Yes" means that the schema is strongly valid; "No" means that it is not even weakly valid (incidentally, none of the properties turns out to be weakly but not strongly valid).

K	$\mathbb{K}_a(\varphi \to \psi) \to (\mathbb{K}_a\varphi \to \mathbb{K}_a\psi)$	*Yes*
D	$\neg\mathbb{K}_a\bot$	*Yes*
T	$\mathbb{K}_a\varphi \to \varphi$	***No***
4	$\mathbb{K}_a\varphi \to \mathbb{K}_a\mathbb{K}_a\varphi$	*Yes*
$\mathbf{4}^+$	$\mathbb{K}_a\varphi \leftrightarrow \mathbb{K}_a\mathbb{K}_a\varphi$	*Yes*
5	$\neg\mathbb{K}_a\varphi \to \mathbb{K}_a\neg\mathbb{K}_a\varphi$	*Yes*
$\mathbf{5}^+$	$\neg\mathbb{K}_a\varphi \leftrightarrow \mathbb{K}_a\neg\mathbb{K}_a\varphi$	*Yes*
B	$\varphi \to \mathbb{K}_a\neg\mathbb{K}_a\neg\varphi$	*No*

Invalidity of Axiom T



Let M be as above

Now, $M, q \models \mathbb{K}_a \neg p$, but $M, q \not\models \neg p$

In Quest for the Truth Axiom

- $\mathbb{K}_a$ is not S5: axioms $\mathbf{K}, \mathbf{D}, \mathbf{4}, \mathbf{5}$ hold, but $\mathbf{T}$ does not
- However, if we slightly restrict the language, then the corresponding $\mathbf{T}$ axiom becomes strongly valid

In Quest for the Truth Axiom

- $\mathbb{K}_a$ is not S5: axioms $\mathbf{K}, \mathbf{D}, \mathbf{4}, \mathbf{5}$ hold, but $\mathbf{T}$ does not
- However, if we slightly restrict the language, then the corresponding $\mathbf{T}$ axiom becomes strongly valid

- Let CSL$^-$ be the subset of CSL in which, between every occurrence of constructive knowledge ($\mathbb{C}_A, \mathbb{E}_A, \mathbb{D}_A$) and negation, there is always at least one operator other than conjunction
- In particular, the requirement is met when $\mathbb{C}_A, \mathbb{E}_A, \mathbb{D}_A$ are never immediately followed by $\neg$ or $\wedge$

In Quest for the Truth Axiom

- $\mathbb{K}_a$ is not S5: axioms $\mathbf{K}, \mathbf{D}, \mathbf{4}, \mathbf{5}$ hold, but $\mathbf{T}$ does not
- However, if we slightly restrict the language, then the corresponding $\mathbf{T}$ axiom becomes strongly valid

- Let CSL^{-} be the subset of CSL in which, between every occurrence of constructive knowledge ($\mathbb{C}_A, \mathbb{E}_A, \mathbb{D}_A$) and negation, there is always at least one operator other than conjunction
- In particular, the requirement is met when $\mathbb{C}_A, \mathbb{E}_A, \mathbb{D}_A$ are never immediately followed by $\neg$ or $\wedge$

Theorem 5.9

Every CSL^{-} instance of $\mathbf{T}$ (i.e., $\mathbb{K}_a\psi \rightarrow \psi$) is strongly valid.

In Quest for the Truth Axiom

Is then the constructive knowledge in CSL^{-} S5?

In Quest for the Truth Axiom

Is then the constructive knowledge in CSL^- S5?

Not really

In Quest for the Truth Axiom

Is then the constructive knowledge in CSL^- S5?
Not really

- The extension of schema $\mathbf{T}$ is different in CSL and CSL^-
- More importantly, in CSL^- schemata $\mathbf{K}$ and $\mathbf{5}$ are not valid, but they are not invalid either – they are simply not formulae at all
- Finally, CSL^- lacks the S5 principle of uniform substitution

Properties of Collective Constructive Knowledge

Theorem 5.10

Below, we list some of the S5 properties for collective constructive knowledge operators. "Yes" means that the schema is strongly valid; "No" means that it is not even weakly valid.

	$\mathbb{C}_A$	$\mathbb{E}_A$	$\mathbb{D}_A$
K	*Yes*	*Yes*	*Yes*
D	*Yes*	*Yes*	*Yes*
T	*No*	*No*	*No*
4	*Yes*	*No*	*Yes*
$\mathbf{4}^+$	*Yes*	*No*	*Yes*
5	*Yes*	*No*	*Yes*
$\mathbf{5}^+$	*Yes*	*No*	*Yes*
B	*No*	*No*	*No*

Properties of Collective Constructive Knowledge

Theorem 5.11

Every CSL$^-$ instance of schema $\mathbf{T}$ for collective constructive knowledge operators $\mathbb{C}_A, \mathbb{E}_A, \mathbb{D}_A$ is strongly valid.

Normal Form and State-Based Semantics

Constructive Normal Form

A CSL formula is in constructive normal form (CSNF) if every subformula starting with a $\hat{\mathcal{K}}_A$ operator is of the form $\hat{\mathcal{K}}_{A_1} \dots \hat{\mathcal{K}}_{A_n} \psi$ where ψ starts with a cooperation modality.

Normal Form and State-Based Semantics

Constructive Normal Form

A CSL formula is in constructive normal form (CSNF) if every subformula starting with a $\hat{\mathcal{K}}_A$ operator is of the form $\hat{\mathcal{K}}_{A_1} \dots \hat{\mathcal{K}}_{A_n} \psi$ where ψ starts with a cooperation modality.

Proposition

Every CSL formula is strongly equivalent to a formula in constructive normal form.

Note: equivalent does not mean the same!

Normal Form CSL

Observation

The "normal form CSL" can be given semantics entirely in terms of models and states.

Normal Form CSL

Observation

The "normal form CSL" can be given semantics entirely in terms of models and states.

$M, q \models \hat{\mathcal{K}}^1_{A_1} \dots \hat{\mathcal{K}}^n_{A_n} \langle\langle A \rangle\rangle \gamma$ iff there exists S_A such that, for every $\lambda \in out(\mathrm{img}(q, rel(\hat{\mathcal{K}}^1_{A_1} \dots \hat{\mathcal{K}}^n_{A_n}), S_A)$, we have that $M, \lambda \models \gamma$,

where $rel(\hat{\mathcal{K}}^1_{A_1} \dots \hat{\mathcal{K}}^n_{A_n}) = \sim^{\mathcal{K}^1}_{A_1} \circ \dots \circ \sim^{\mathcal{K}^n}_{A_n}$.

Normal Form CSL vs. Onion Soup

- $\neg\mathbb{E}_{\mathbb{A}\mathrm{gt}}\langle\langle\mathbb{A}\mathrm{gt}\rangle\rangle\Diamond\mathsf{open}$
- $\mathbb{D}_{\mathbb{A}\mathrm{gt}}\langle\langle\mathbb{A}\mathrm{gt}\rangle\rangle\Diamond\mathsf{open}$
- $\mathbb{D}_{\{1,\dots,n-1\}}\langle\langle\mathbb{A}\mathrm{gt}\rangle\rangle\Diamond\mathsf{open}$

These are normal form formulae!

5.5 Between Perception and Recall

Strategies for Different Settings

Four variants of ability: IR, Ir, iR, ir (Schobbens 2004)

- I/i: perfect/imperfect information
- R/r: perfect/imperfect recall

Strategies for Different Settings

Four variants of ability: IR, Ir, iR, ir (Schobbens 2004)

- I/i: perfect/imperfect information
- R/r: perfect/imperfect recall

- r: $s_a : St \rightarrow Act$ (memoryless strategies)
- R: $s_a : St^+ \rightarrow Act$ (perfect recall strategies)

Strategies for Different Settings

Four variants of ability: IR, Ir, iR, ir (Schobbens 2004)

- I/i: perfect/imperfect information
- R/r: perfect/imperfect recall

- r: $s_a : St \rightarrow Act$ (memoryless strategies)
- R: $s_a : St^+ \rightarrow Act$ (perfect recall strategies)

- i: only uniform strategies,
- I: no restrictions

Strategies for Different Settings

Four variants of ability: IR, Ir, iR, ir (Schobbens 2004)

- I/i: perfect/imperfect information
- R/r: perfect/imperfect recall

- r: $s_a : St \to Act$ (memoryless strategies)
- R: $s_a : St^+ \to Act$ (perfect recall strategies)

- i: only uniform strategies,
- I: no restrictions

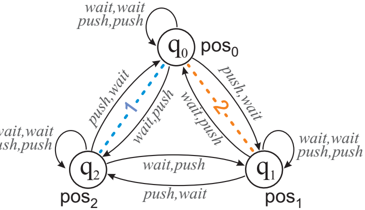
- r: s_a is uniform iff $q \sim_a q' \Rightarrow s_a(q) = s_a(q')$
- R: s_a is uniform iff $\lambda \approx_a \lambda' \Rightarrow s_a(\lambda) = s_a(\lambda')$
- $\lambda \approx_a \lambda'$ iff $\forall_i \lambda[i] \sim_a \lambda'[i]$

Model Checking Complexity

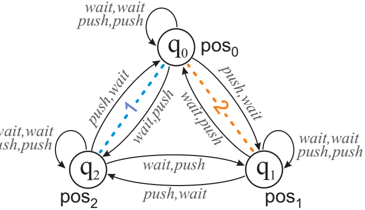
logic	ir	iR	Ir	IR
$\langle\langle\Gamma\rangle\rangle - ATL$	NP	U [11]	$n\,l$ [2]	$n\,l$ [2]
ATL	$\Delta_2 P$	U [11]	$n\,l$ [2]	$n\,l$ [2]
ATL^+	$\Delta_3 P$	U [11]	$\Delta_3 P$	$\Delta_3 P$
ATL^*	$PSPACE$	U [11]	$PSPACE$	$DEXP$ [9]

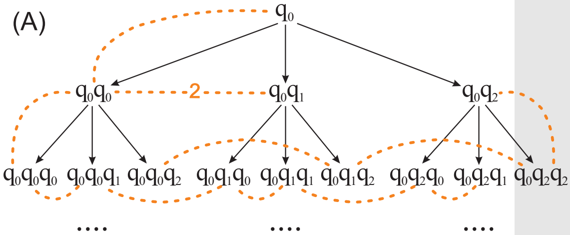
NP	complete for nondeterministic polynomial time
$\Delta_2 P = P^{NP}$	complete for polynomial calls to an NP oracle
$\Delta_3 P = P^{NP^{NP}}$	complete for polynomial calls to a $\Sigma_2 P$ oracle
EXP	complete for deterministic exponential time
DEXP	complete for deterministic doubly exponential time
U	undecidable
l	size of the formula
n	size of the model

Perfect vs. Imperfect Recall

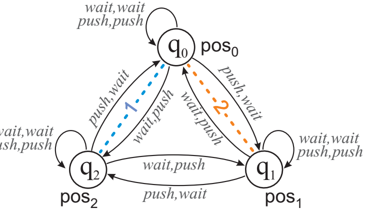


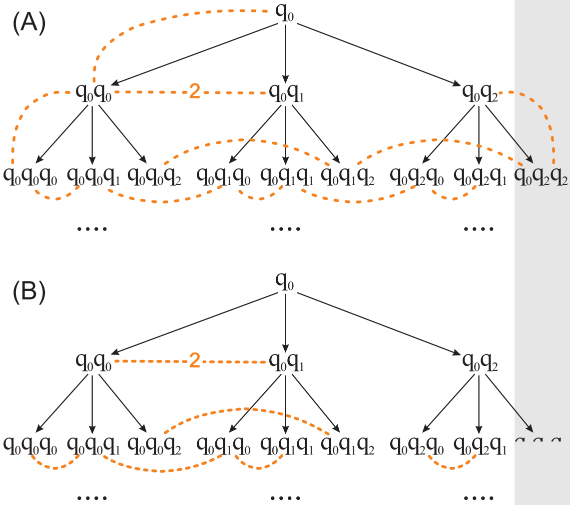
Perfect vs. Imperfect Recall



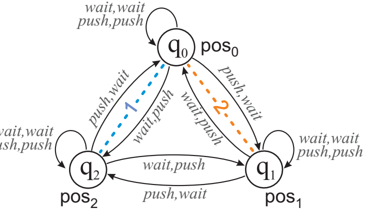


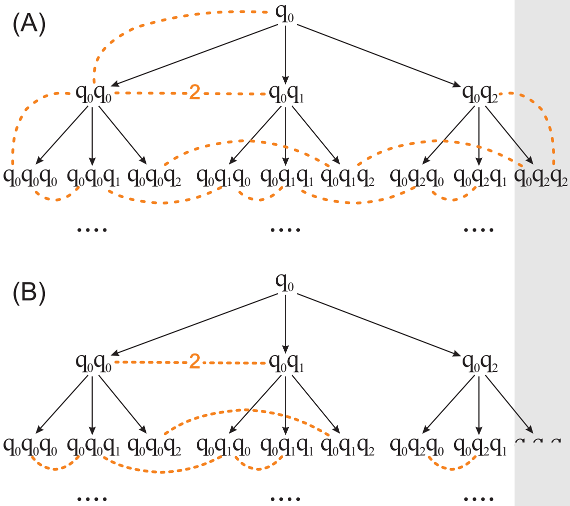
Perfect vs. Imperfect Recall





Perfect vs. Imperfect Recall





Advice: the restrictions on strategies and the semantics of epistemic operators should match!

5.6 References

[Schobbens 2004] P. Y. Schobbens.
Alternating-time logic with imperfect recall.
Electronic Notes in Theoretical Computer Science, 85(2), 2004.

[Jamroga and Ågotnes 2007] W. Jamroga and T. Ågotnes.
Constructive knowledge: What agents can achieve under incomplete information.
Journal of Applied Non-Classical Logics, 17(4):423–475, 2007.

[van der Hoek and Wooldridge 2003] W. van der Hoek and M. Wooldridge.
Cooperation, knowledge and time: Alternating-time Temporal Epistemic Logic and its applications.
Studia Logica, 75(1):125–157, 2003.

Strat. Solution Concepts

Introduction

- Let us look at how we can logically characterise solution concepts for strategic games

Introduction

- Let us look at how we can logically characterise solution concepts for strategic games
- Modal logic characterisations of solution concepts have been studied by many authors, e.g.
  - Bonanno: both strategic and extensive games
  - Harrenstein et al.: extensive form games; modalities for preferences (see our Friday lectures)
- Here: we will take Coalition Logic/ATL as a starting point

- Strategic game: $G = (N, \{\Sigma_i : i \in N\}, o, S, \{\succeq_i : i \in N\})$
- Solution concepts describe outcomes of games

Nash Equilibrium

Informally: given a game, a strategy profile is a Nash equilibrium iff every strategy is a best response (for that agent) to the other strategies.

Nash Equilibrium

Informally: given a game, a strategy profile is a Nash equilibrium iff every strategy is a best response (for that agent) to the other strategies.

Formally:

$$G = (N, \{\Sigma_i : i \in N\}, o, S, \{\succeq_i : i \in N\})$$

Definition 6.1 (Nash Equilibrium)

A strategy profile σ_N is a (pure strategy) Nash equilibrium of G iff for every $i \in N$ and σ_i'

$$o(\sigma_i, \sigma_{-i}) \succeq_i o(\sigma_i', \sigma{-}i)$$

Nash Equilibrium: example: Prisoner's dilemma

		Bill	
		Cooperate	**Defect**
Ann	**Cooperate**	Ann:-1, Bill:-1	Ann:-4, Bill: 0
	Defect	Ann:0, Bill:-4	Ann:-3, Bill: -3

Nash Equilibrium: example: Prisoner's dilemma

		Bill	
		Cooperate	**Defect**
Ann	**Cooperate**	Ann:-1, Bill:-1	Ann:-4, Bill: 0
	Defect	Ann:0, Bill:-4	Ann:-3, Bill: -3

One Nash equilibrium: (Defect,Defect)

Nash Equilibrium: example: Bach or Stravinsky

Bill

Ann		**B**	**S**
	B	Ann:2, Bill:1	Ann:0, Bill: 0
	S	Ann:0, Bill:0	Ann:1, Bill: 2

Nash Equilibrium: example: Bach or Stravinsky

		Bill	
		B	**S**
Ann	**B**	Ann:2, Bill:1	Ann:0, Bill: 0
	S	Ann:0, Bill:0	Ann:1, Bill: 2

Two Nash equilibria: (B,B) and (S,S)

Weakly Dominant Strategies

Informally: a strategy is weakly dominant if it is as least as good as any other strategy no matters what the other agents do.

Weakly Dominant Strategies

Informally: a strategy is weakly dominant if it is as least as good as any other strategy no matters what the other agents do.

Formally:

Definition 6.2 (Weak Dominance)

A strategy σ_i weakly dominates strategy σ_i' iff for all σ_{-i}

$$o(\sigma_i, \sigma_{-i}) \succeq_i o(\sigma_i', \sigma_{-i})$$

A strategy is weakly dominant for i iff it weakly dominates all other strategies for i.

Dominance: example: Prisoner's dilemma

		Bill	
		Cooperate	**Defect**
Ann	**Cooperate**	Ann:-1, Bill:-1	Ann:-4, Bill: 0
	Defect	Ann:0, Bill:-4	Ann:-3, Bill: -3

Dominance: example: Prisoner's dilemma

Ann \ Bill	Cooperate	Defect
Cooperate	Ann:-1, Bill:-1	Ann:-4, Bill: 0
Defect	Ann:0, Bill:-4	Ann:-3, Bill: -3

Defect is dominant for Ann

Dominance: example: Prisoner's dilemma

		Bill	
		Cooperate	**Defect**
Ann	**Cooperate**	Ann:-1, Bill:-1	Ann:-4, Bill: 0
	Defect	Ann:0, Bill:-4	Ann:-3, Bill: -3

Defect is dominant for Ann
Defect is dominant for Bill

6.1 Logical Characterisations

Adding preferences

- ATL/CL can express properties about (sequences of) game forms
- In order to reason about solution concepts, we need to add preferences to the picture
- Can be done in several ways
- In lectures 9 and 10, we use preference modalities
- Here we choose a simple solution: “primitive” utility propositions

Utility propositions

Let U be a finite set of utilities. We assume that the primitive propositions Π includes a proposition

$$u_i \geq v$$

for each agent i and $v \in U$.

Utility propositions

Let U be a finite set of utilities. We assume that the primitive propositions Π includes a proposition

$$u_i \geq v$$

for each agent i and $v \in U$.

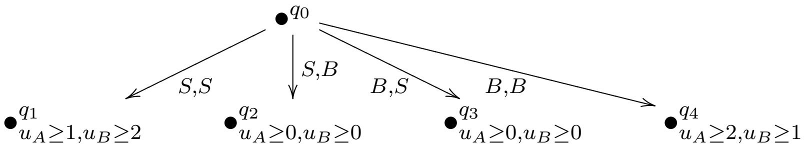
It is now straightforward to identify a strategic game in each state (where the outcomes are new states). We use $\Gamma(\mathcal{M}, s)$ to denote the game played in state s of structure $\mathcal{M}$.

Example (BoS)

$\Gamma(\mathcal{M}, q_0)$:

		Bill	
		B	**S**
Ann	**B**	Ann:2, Bill:1	Ann:0, Bill: 0
	S	Ann:0, Bill:0	Ann:1, Bill: 2

$\mathcal{M}$:



- We can now express properties of games: $\mathcal{M}, s \models \phi$ means that the game $\Gamma(\mathcal{M}, s)$ has the property decribed by ϕ

- We can now express properties of games: $\mathcal{M}, s \models \phi$ means that the game $\Gamma(\mathcal{M}, s)$ has the property decribed by ϕ
- Game properties of special interest: solution concepts
  - From now on: simplifying assumption: $N = 2$

- We can now express properties of games: $\mathcal{M}, s \models \phi$ means that the game $\Gamma(\mathcal{M}, s)$ has the property decribed by ϕ
- Game properties of special interest: solution concepts
  - From now on: simplifying assumption: $N = 2$
- Can we express, e.g., Nash equilibrium using the key construct $\langle\langle C \rangle\rangle$?

- We can now express properties of games: $\mathcal{M}, s \models \phi$ means that the game $\Gamma(\mathcal{M}, s)$ has the property decribed by ϕ
- Game properties of special interest: solution concepts
  - From now on: simplifying assumption: $N = 2$
- Can we express, e.g., Nash equilibrium using the key construct $\langle\langle C \rangle\rangle$?
- Turns out to be difficult. Main reasons:

- We can now express properties of games: $\mathcal{M}, s \models \phi$ means that the game $\Gamma(\mathcal{M}, s)$ has the property decribed by ϕ
- Game properties of special interest: solution concepts
  - From now on: simplifying assumption: $N = 2$
- Can we express, e.g., Nash equilibrium using the key construct $\langle\langle C \rangle\rangle$?
- Turns out to be difficult. Main reasons:
  - Solution concepts such as Nash equilibrium are properties of strategies, but we cannot refer directly to strategies in the language

- We can now express properties of games: $\mathcal{M}, s \models \phi$ means that the game $\Gamma(\mathcal{M}, s)$ has the property decribed by ϕ
- Game properties of special interest: solution concepts
  - From now on: simplifying assumption: $N = 2$
- Can we express, e.g., Nash equilibrium using the key construct $\langle\langle C \rangle\rangle$?
- Turns out to be difficult. Main reasons:
  - Solution concepts such as Nash equilibrium are properties of strategies, but we cannot refer directly to strategies in the language
  - We often need to reason in the context of a fixed strategy for one or more agents: "If my opponent cooperates, then..". This requires an "irrevocable" interpretation of strategies.

- We can now express properties of games: $\mathcal{M}, s \models \phi$ means that the game $\Gamma(\mathcal{M}, s)$ has the property decribed by ϕ
- Game properties of special interest: solution concepts
  - From now on: simplifying assumption: $N = 2$
- Can we express, e.g., Nash equilibrium using the key construct $\langle\langle C \rangle\rangle$?
- Turns out to be difficult. Main reasons:
  - Solution concepts such as Nash equilibrium are properties of strategies, but we cannot refer directly to strategies in the language
  - We often need to reason in the context of a fixed strategy for one or more agents: "If my opponent cooperates, then..". This requires an "irrevocable" interpretation of strategies.
  - Reasoning about solution concepts involve counterfactual arguments such as "Suppose my opponent cooperates. Then I better defect. If he defects, however, I should defect as well."
- We will thus make another addition to the language, in addition to utility propositions

Counterfactuals

- Example: "Suppose my opponent cooperates. Then I better defect. If he defects, however, I should defect as well."
- Counterfactuals are not logical implications (otherwise one of the claims above would be trivially true)
- Counterfactuals have been analysed by philosophers (Stalnaker, Lewis):
  - "if counterfactually ϕ then ψ": if we adjust the world minimally so that ϕ, then ψ

A Counterfactual Operator

Extend the language of ATL (or Coalition Logic) with a counterfactual operator

$$C_i(\sigma_i, \varphi)$$

where

- i is an agent
- σ_i is a strategy term. We assume a set of strategy terms Υ_i for each agent i.
- φ is a formula

with the intended meaning that if i played strategy σ_i, then φ would be true

A Counterfactual Operator

Extend the language of ATL (or Coalition Logic) with a counterfactual operator

$$C_i(\sigma_i, \varphi)$$

where

- i is an agent
- σ_i is a strategy term. We assume a set of strategy terms Υ_i for each agent i.
- φ is a formula

with the intended meaning that if i played strategy σ_i, then φ would be true

Restriction: no occurrence of a term in Υ_i inside φ

Interpretation

Extend the semantic structures (CGSs) with an interpretation function $[\![\cdot]\!]_{\mathcal{M}}$ mapping a strategy term $\sigma_i \in \Upsilon_i$ to a strategy

$$[\![\sigma_i]\!]_{\mathcal{M}}$$

for agent i.

Interpretation

Extend the semantic structures (CGSs) with an interpretation function $[\![\cdot]\!]_{\mathcal{M}}$ mapping a strategy term $\sigma_i \in \Upsilon_i$ to a strategy

$$[\![\sigma_i]\!]_{\mathcal{M}}$$

for agent i.

Interpretation:

$$\mathcal{M}, q \models C_i(\sigma_i, \varphi) \Leftrightarrow (\mathcal{M} \dagger [\![\sigma_i]\!], q \models \varphi)$$

Interpretation

Extend the semantic structures (CGSs) with an interpretation function $[\![\cdot]\!]_{\mathcal{M}}$ mapping a strategy term $\sigma_i \in \Upsilon_i$ to a strategy

$$[\![\sigma_i]\!]_{\mathcal{M}}$$

for agent i.

Interpretation:

$$\mathcal{M}, q \models C_i(\sigma_i, \varphi) \Leftrightarrow (\mathcal{M} \dagger [\![\sigma_i]\!], q \models \varphi)$$

Assumption: there is a term for every possible strategy.

Note that

$$\not\models \langle\langle i\rangle\rangle\varphi \leftrightarrow \bigvee_{\sigma_i\in\Upsilon_i} C_i(\sigma_i,\varphi)$$

Note that

$$\not\models \langle\!\langle i \rangle\!\rangle \varphi \leftrightarrow \bigvee_{\sigma_i \in \Upsilon_i} C_i(\sigma_i, \varphi)$$

... similarly to the fact that $\langle\!\langle i \rangle\!\rangle$ is different with the standard and the irrevocable semantics: the update semantics rules out any possible future choices

Characterising Weak Dominance

Find a formula $WD_i(\alpha)$ such that

$$\mathcal{M}, q \models WD_i(\alpha) \Leftrightarrow \alpha \text{ is weakly dominant in } \Gamma(\mathcal{M}, q)$$

Characterising Weak Dominance

Find a formula $WD_i(\alpha)$ such that

$$\mathcal{M}, q \models WD_i(\alpha) \Leftrightarrow \ \alpha \text{ is weakly dominant in } \Gamma(\mathcal{M}, q)$$

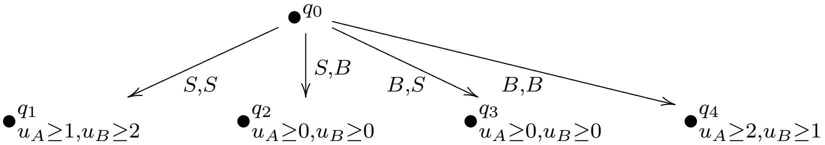
Consider this:

$$wd_i(\alpha) \ \equiv \ \bigwedge_{v \in U} (\langle\langle i \rangle\rangle \bigcirc (u_i \geq v) \rightarrow C_i(\alpha, \langle\langle\rangle\rangle \bigcirc (u_i \geq v)))$$

$$wd_i(\alpha) \;\equiv\; \bigwedge_{v \in U} (\langle\!\langle i \rangle\!\rangle \bigcirc (u_i \geq v) \rightarrow C_i(\alpha, \langle\!\langle \rangle\!\rangle \bigcirc (u_i \geq v)))$$

$$wd_i(\alpha) \equiv \bigwedge_{v \in U} (\langle\langle i \rangle\rangle \bigcirc (u_i \geq v) \rightarrow C_i(\alpha, \langle\langle\rangle\rangle \bigcirc (u_i \geq v)))$$

$\mathcal{M}$:

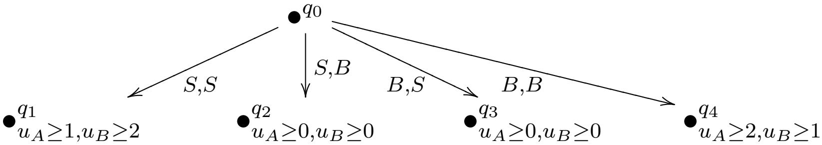


$\Gamma(\mathcal{M}, q_0)$:

		Bill	
		B	**S**
Ann	**B**	Ann:2, Bill:1	Ann:0, Bill: 0
	S	Ann:0, Bill:0	Ann:1, Bill: 2

$$wd_i(\alpha) \equiv \bigwedge_{v \in U} (\langle\langle i \rangle\rangle \bigcirc (u_i \geq v) \rightarrow C_i(\alpha, \langle\langle\rangle\rangle \bigcirc (u_i \geq v)))$$

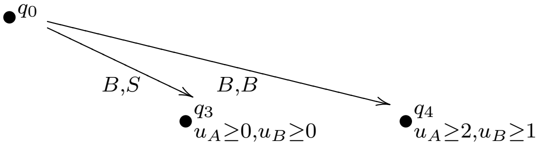
$\mathcal{M}$:



- If $\mathcal{M}, q_0 \models \langle\langle A \rangle\rangle \bigcirc u_A \geq v$, then $v = 0$

$$wd_i(\alpha) \equiv \bigwedge_{v \in U} (\langle\langle i \rangle\rangle \bigcirc (u_i \geq v) \rightarrow C_i(\alpha, \langle\langle\rangle\rangle \bigcirc (u_i \geq v)))$$

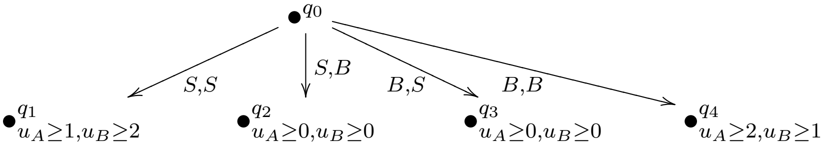
$\mathcal{M} \dagger B_A$:



- If $\mathcal{M}, q_0 \models \langle\langle A \rangle\rangle \bigcirc u_A \geq v$, then $v = 0$
- But $\mathcal{M} \dagger B_A, q_0 \models \langle\langle\rangle\rangle \bigcirc u_A \geq 0$ (if Ann plays B, she will get at least 0)

$$wd_i(\alpha) \equiv \bigwedge_{v \in U} (\langle\langle i \rangle\rangle \bigcirc (u_i \geq v) \rightarrow C_i(\alpha, \langle\langle\rangle\rangle \bigcirc (u_i \geq v)))$$

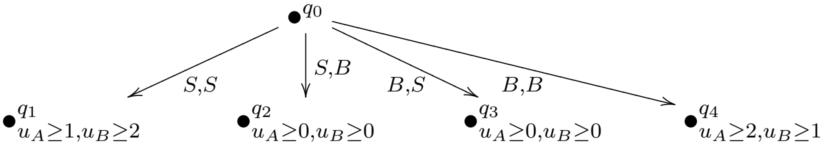
$\mathcal{M}$:



- If $\mathcal{M}, q_0 \models \langle\langle A \rangle\rangle \bigcirc u_A \geq v$, then $v = 0$
- But $\mathcal{M} \dagger B_A, q_0 \models \langle\langle\rangle\rangle \bigcirc u_A \geq 0$ (if Ann plays B, she will get at least 0)
- So $\mathcal{M}, q_0 \models C_A(B, \langle\langle\rangle\rangle \bigcirc u_A \geq 0)$ as well

$$wd_i(\alpha) \equiv \bigwedge_{v \in U} (\langle\langle i \rangle\rangle \bigcirc (u_i \geq v) \rightarrow C_i(\alpha, \langle\langle\rangle\rangle \bigcirc (u_i \geq v)))$$

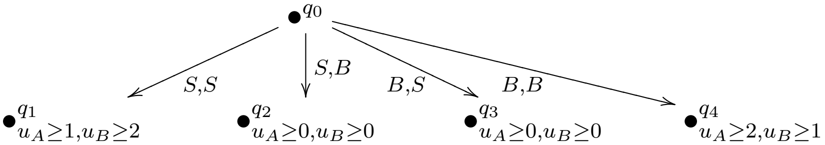
$\mathcal{M}$:



- If $\mathcal{M}, q_0 \models \langle\langle A \rangle\rangle \bigcirc u_A \geq v$, then $v = 0$
- But $\mathcal{M} \dagger B_A, q_0 \models \langle\langle\rangle\rangle \bigcirc u_A \geq 0$ (if Ann plays B, she will get at least 0)
- So $\mathcal{M}, q_0 \models C_A(B, \langle\langle\rangle\rangle \bigcirc u_A \geq 0)$ as well
- Thus, $\mathcal{M}, q_0 \models wd_i(B)$

$$wd_i(\alpha) \equiv \bigwedge_{v \in U} (\langle\langle i \rangle\rangle \bigcirc (u_i \geq v) \rightarrow C_i(\alpha, \langle\langle\rangle\rangle \bigcirc (u_i \geq v)))$$

$\mathcal{M}$:



- If $\mathcal{M}, q_0 \models \langle\langle A \rangle\rangle \bigcirc u_A \geq v$, then $v = 0$
- But $\mathcal{M} \dagger B_A, q_0 \models \langle\langle\rangle\rangle \bigcirc u_A \geq 0$ (if Ann plays B, she will get at least 0)
- So $\mathcal{M}, q_0 \models C_A(B, \langle\langle\rangle\rangle \bigcirc u_A \geq 0)$ as well
- Thus, $\mathcal{M}, q_0 \models wd_i(B)$
- But B is not a dominant strategy for Ann! (BoS has no dominant strategies)

Characterising Weak Dominance

Find a formula $WD_i(\alpha)$ such that

$$\mathcal{M}, q \models WD_i(\alpha) \Leftrightarrow \ \alpha \text{ is weakly dominant in } \Gamma(\mathcal{M}, q)$$

$$wd_i(\alpha) \ \equiv \ \bigwedge_{v \in U} (\langle\langle i \rangle\rangle \bigcirc (u_i \geq v) \rightarrow C_i(\alpha, \langle\langle\rangle\rangle \bigcirc (u_i \geq v)))$$

Characterising Weak Dominance

Find a formula $WD_i(\alpha)$ such that

$$\mathcal{M}, q \models WD_i(\alpha) \Leftrightarrow \ \alpha \text{ is weakly dominant in } \Gamma(\mathcal{M}, q)$$

$$wd_i(\alpha) \ \equiv \ \bigwedge_{v \in U} (\langle\langle i \rangle\rangle \bigcirc (u_i \geq v) \rightarrow C_i(\alpha, \langle\langle\rangle\rangle \bigcirc (u_i \geq v)))$$

Solution – for games with finitely many strategies:

$$WD_i(\alpha) \ \equiv \ \bigwedge_{\beta \in \Upsilon_j} C_j(\beta, wd_i(\alpha))$$

Characterising Nash Equilibrium

Find a formula $NE_i(\alpha_1, \alpha_2)$ such that

$\mathcal{M}, q \models NE(\alpha_1, \alpha_2) \Leftrightarrow (\alpha_1, \alpha_2)$ is a Nash equilibrium of $\Gamma(\mathcal{M}, q)$

Characterising Nash Equilibrium

Find a formula $NE_i(\alpha_1, \alpha_2)$ such that

$\mathcal{M}, q \models NE(\alpha_1, \alpha_2) \Leftrightarrow (\alpha_1, \alpha_2)$ is a Nash equilibrium of $\Gamma(\mathcal{M}, q)$

Best response:

$$BR_i(\alpha_k, \alpha_i) \equiv C_k(\alpha_k, \bigwedge_{v \in U} ((\langle\langle i \rangle\rangle \bigcirc (u_i \geq v)) \rightarrow C_i(\alpha_i, \langle\langle\rangle\rangle \bigcirc (u_i \geq v)))$$

Characterising Nash Equilibrium

Find a formula $NE_i(\alpha_1, \alpha_2)$ such that

$\mathcal{M}, q \models NE(\alpha_1, \alpha_2) \Leftrightarrow (\alpha_1, \alpha_2)$ is a Nash equilibrium of $\Gamma(\mathcal{M}, q)$

Best response:

$$BR_i(\alpha_k, \alpha_i) \equiv C_k(\alpha_k, \bigwedge_{v \in U}((\langle\langle i \rangle\rangle \bigcirc (u_i \geq v)) \rightarrow C_i(\alpha_i, \langle\langle\rangle\rangle \bigcirc (u_i \geq v)))$$

$$NE(\alpha_1, \alpha_2) \equiv BR_1(\alpha_2, \alpha_1) \wedge BR_2(\alpha_1, \alpha_2)$$

6.2 References

[1] M. J. Osborne and A. Rubinstein.
A Course in Game Theory.
The MIT Press: Cambridge, MA, 1994.

[2] W. van der Hoek, W. Jamroga, and M. Wooldridge.
A logic for strategic reasoning.
In *Proceedings of AAMAS'05*, pages 157–164, 2005.

[3] G. Bonanno.
Modal logic and game theory: Two alternative approaches.
Risk Decision and Policy, 7(3):309–324, 2002.

[4] B.P. Harrenstein, W. van der Hoek, J.-J. Meyer, and C. Witteveen.
A modal characterization of Nash equilibrium.
Fundamenta Informaticae, 57(2–4):281–321, 2003.

Reasoning about Rational Play

Game-Theoretical Analysis of Games

- Solution concepts define rationality of players

Game-Theoretical Analysis of Games

- Solution concepts define rationality of players
  - maxmin
  - Nash equilibrium
  - subgame-perfect Nash
  - undominated strategies
  - Pareto optimality

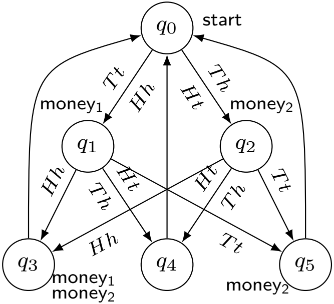
Game-Theoretical Analysis of Games

- Solution concepts define rationality of players
  - maxmin
  - Nash equilibrium
  - subgame-perfect Nash
  - undominated strategies
  - Pareto optimality
- Then: we assume that players play rationally
- ...and we ask about the outcome of the game under this assumption

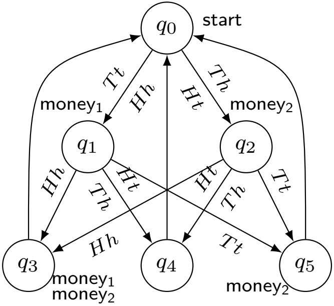
Game-Theoretical Analysis of Games

- Solution concepts define rationality of players
  - maxmin
  - Nash equilibrium
  - subgame-perfect Nash
  - undominated strategies
  - Pareto optimality
- Then: we assume that players play rationally
- ...and we ask about the outcome of the game under this assumption
- Role of rationality criteria: constrain the possible game moves to “sensible” ones

Example: Pennies Game

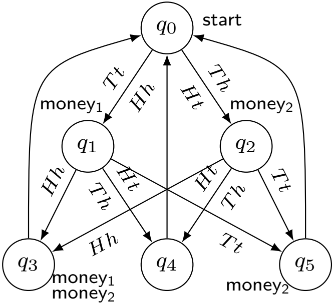


Example: Pennies Game



$$\mathsf{start} \rightarrow \neg\langle\langle 1\rangle\rangle\Diamond\mathsf{money}_1$$

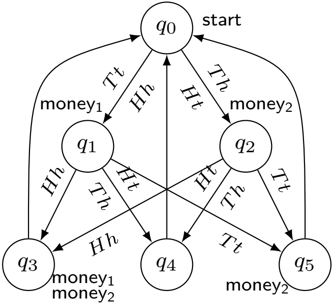
Example: Pennies Game



$$\mathsf{start} \rightarrow \neg\langle\langle 1\rangle\rangle\Diamond\mathsf{money}_1$$

$$\mathsf{start} \rightarrow \neg\langle\langle 2\rangle\rangle\Diamond\mathsf{money}_2$$

Example: Pennies Game



$$\text{start} \rightarrow \neg\langle\langle 1\rangle\rangle\Diamond \text{money}_1$$

$$\text{start} \rightarrow \neg\langle\langle 2\rangle\rangle\Diamond \text{money}_2$$

Game-Theoretical Analysis of Games

Two points of focus:

- characterization of rationality
  $\rightsquigarrow$ research in game theory

Game-Theoretical Analysis of Games

Two points of focus:

- characterization of rationality
  $\rightsquigarrow$ research in game theory
- using solution concepts to predict outcomes in a given game
  $\rightsquigarrow$ applications of game theory

Motivation

We would like to ...

... reason about the outcome of rational play

Motivation

We would like to ...

... reason about the outcome of rational play

... have a logic that embed any solution concept

... compare different game theoretical solution concepts wrt their outcomes

Motivation

We would like to ...

- ... reason about the outcome of rational play
- ... have a logic that embed any solution concept
- ... compare different game theoretical solution concepts wrt their outcomes

So ...

- ... we extend ATL with a notion of rationality/plausibility
- ... reason about what rational agents can achieve

Inspiration:

- Game Logics with Preferences (van Otterloo, van der Hoek & Wooldridge): Nash equilibria, subgame perfect strategies
- Epistemic Temporal Strategic Logic (van Otterloo & Jonker): undominated strategies

7.1 ATL + Plausibility

ATL with Plausibility

ATL: reasoning about *all* possible behaviors.

$\langle\langle A\rangle\rangle\varphi$: agents A have some collective strategy to enforce φ against any response of their opponents.

ATL with Plausibility

ATL: reasoning about *all* possible behaviors.

$\langle\langle A \rangle\rangle\varphi$: agents A have some collective strategy to enforce φ against any response of their opponents.

ATLP: reasoning about *plausible* behaviors.

$\mathbf{Pl}\,\langle\langle A \rangle\rangle\varphi$: agents A have a plausible collective strategy to enforce φ against any plausible response of their opponents.

ATL with Plausibility

ATL: reasoning about *all* possible behaviors.

$\langle\langle A\rangle\rangle\varphi$: agents A have some collective strategy to enforce φ against any response of their opponents.

ATLP: reasoning about *plausible* behaviors.

$\mathbf{Pl}\,\langle\langle A\rangle\rangle\varphi$: agents A have a plausible collective strategy to enforce φ against any plausible response of their opponents.

Important

The possible strategies of both A and $\mathbb{A}\mathrm{gt}\backslash A$ are restricted.

ATL with Plausibility

Syntax of ATLP

$$\varphi ::= p \mid \neg\varphi \mid \varphi \land \varphi \mid \langle\langle A\rangle\rangle \bigcirc \varphi \mid \langle\langle A\rangle\rangle \Box \varphi \mid \langle\langle A\rangle\rangle \varphi \mathcal{U} \varphi \mid$$
$$(\textbf{set-pl}\ \omega)\varphi$$

ATL with Plausibility

Syntax of ATLP

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\langle A\rangle\rangle \bigcirc \varphi \mid \langle\langle A\rangle\rangle \Box \varphi \mid \langle\langle A\rangle\rangle \varphi \mathcal{U} \varphi \mid (\textbf{set-pl}\ \omega)\varphi$$

New in ATLP:

$(\textbf{set-pl}\ \omega)$: the set of plausible profiles is set/reset to the strategies described by ω.
Only plausible strategy profiles are considered!

ATL with Plausibility

Syntax of ATLP

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\langle A\rangle\rangle \bigcirc \varphi \mid \langle\langle A\rangle\rangle \Box \varphi \mid \langle\langle A\rangle\rangle \varphi \mathcal{U} \varphi \mid (\textbf{set-pl}\ \omega)\varphi$$

New in ATLP:

(**set-pl** ω) : the set of plausible profiles is set/reset to the strategies described by ω.
Only plausible strategy profiles are considered!

Example: $(\textbf{set-pl}\ greedy_1)\langle\langle 2\rangle\rangle\Diamond \mathsf{money}_2$

Concurrent Game Structures with Plausibility

$$M = (\mathbb{A}\mathrm{gt}, St, \Pi, \pi, Act, d, \delta, \Upsilon, \Omega, \|\cdot\|)$$

Concurrent Game Structures with Plausibility

$M = (\mathbb{A}\mathrm{gt}, St, \Pi, \pi, Act, d, \delta, \Upsilon, \Omega, \|\cdot\|)$

- $\Upsilon \subseteq \Sigma$: set of plausible strategy profiles

	Deny	Confess
Deny	$\boxed{-2, -2}$	$-5, -1$
Confess	$-1, -5$	$\boxed{-4, -4}$

Υ

Concurrent Game Structures with Plausibility

$M = (\mathbb{A}\mathrm{gt}, St, \Pi, \pi, Act, d, \delta, \Upsilon, \Omega, \|\cdot\|)$

- $\Upsilon \subseteq \Sigma$: set of plausible strategy profiles

	Deny	Confess
Deny	$\boxed{-2, -2}$	$-5, -1$
Confess	$-1, -5$	$\boxed{-4, -4}$

Υ

- $\Omega = \{\omega_1, \omega_2, \ldots\}$: set of plausibility terms

Example: ω_{NE} may stand for all Nash equilibria

Concurrent Game Structures with Plausibility

$M = (\mathbb{A}\mathrm{gt}, St, \Pi, \pi, Act, d, \delta, \Upsilon, \Omega, \|\cdot\|)$

- $\Upsilon \subseteq \Sigma$: set of plausible strategy profiles

	Deny	Confess
Deny	$\boxed{-2, -2}$	$-5, -1$
Confess	$-1, -5$	$\boxed{-4, -4}$

Υ

- $\Omega = \{\omega_1, \omega_2, \ldots\}$: set of plausibility terms

  Example: ω_{NE} may stand for all Nash equilibria

- $\|\cdot\| : St \to (\Omega \to 2^{\Sigma})$: plausibility mapping, assigns set of strategy profiles to each state and plausibility term

  Example: $\|\omega_{NE}\|_q = \{(\text{confess}, \text{confess})\}$

Semantics of ATLP

$\Sigma_A(\Upsilon)$: collective strategies of A that are consistent with Υ

Restricting A's strategies

$$\Sigma_A(\Upsilon) = \{s_A \in \Sigma_A \mid \exists t \in \Upsilon \quad (t[A] = s_A)\}$$

Semantics of ATLP

$\Sigma_A(\Upsilon)$: collective strategies of A that are consistent with Υ

Restricting A's strategies

$$\Sigma_A(\Upsilon) = \{s_A \in \Sigma_A \mid \exists t \in \Upsilon \quad (t[A] = s_A)\}$$

We also restrict the opponents' responses to s_A
$\Upsilon(s_A)$: plausible strategy profiles of $\mathbb{A}\mathrm{gt}$ that agree on s_A

Restricting A's opponents strategies

$$\Upsilon(s_A) = \{t \in \Upsilon \mid t[A] = s_A\}$$

Restricting Strategies

	Deny	Confess
Deny	$\boxed{-2, -2}$	$-5, -1$
Confess	$-1, -5$	$\boxed{-4, -4}$

Υ

$$\Upsilon = \{(confess_1, confess_2), (deny_1, deny_2)\}$$

Restricting Strategies

	Deny	Confess
Deny	$\boxed{-2, -2}$	$-5, -1$
Confess	$-1, -5$	$\boxed{-4, -4}$

Υ

$$\Upsilon = \{(confess_1, confess_2), (deny_1, deny_2)\}$$
$$\Sigma_1(\Upsilon) = \{confess_1, deny_1\}$$

Restricting Strategies

	Deny	Confess
Deny	$\boxed{-2,-2}$	$-5,-1$
Confess	$-1,-5$	$\boxed{-4,-4}$

Υ

$$\Upsilon = \{(\mathit{confess}_1, \mathit{confess}_2), (\mathit{deny}_1, \mathit{deny}_2)\}$$
$$\Sigma_1(\Upsilon) = \{\mathit{confess}_1, \mathit{deny}_1\}$$
$$P(\mathit{confess}_1) = \{(\mathit{confess}_1, \mathit{confess}_2)\}.$$

Outcome of a Strategy

Outcome = Paths that may occur when agents A perform s_A

Outcome of a Strategy

Outcome = Paths that may occur when agents A perform s_A and only plausible strategy profiles are played

Outcome of a Strategy

Outcome = Paths that may occur when agents A perform s_A and only plausible strategy profiles are played

$$out_\Upsilon(q, s_A) =$$

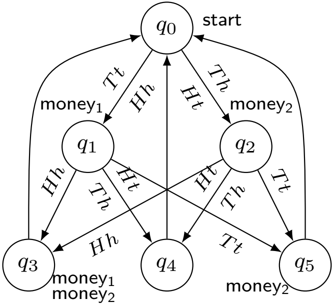
$$\{\lambda \in St^+ \mid \exists t \in \Upsilon(s_A)\ \forall i \in \mathbb{N}\ \big(\lambda[i+1] = \delta(\lambda[i], t(\lambda[i]))\big)\}$$

Outcome of a Strategy

Outcome = Paths that may occur when agents A perform s_A and only plausible strategy profiles are played

$$out_\Upsilon(q, s_A) =$$

$$\{\lambda \in St^+ \mid \exists t \in \Upsilon(s_A)\ \forall i \in \mathbb{N}\ \big(\lambda[i+1] = \delta(\lambda[i], t(\lambda[i]))\big)\}$$

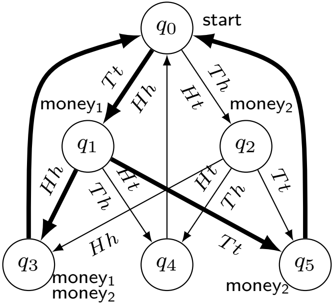


Outcome of a Strategy

Outcome = Paths that may occur when agents A perform s_A and only plausible strategy profiles are played

$$out_\Upsilon(q, s_A) =$$

$$\{\lambda \in St^+ \mid \exists t \in \Upsilon(s_A)\ \forall i \in \mathbb{N}\ \big(\lambda[i+1] = \delta(\lambda[i], t(\lambda[i]))\big)\}$$



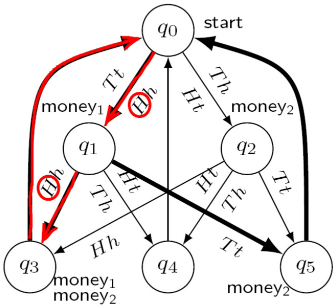
P: the players always show same sides of their coins

Outcome of a Strategy

Outcome = Paths that may occur when agents A perform s_A and only plausible strategy profiles are played

$$out_{\Upsilon}(q, s_A) =$$

$$\{\lambda \in St^{+} \mid \exists t \in \Upsilon(s_A)\ \forall i \in \mathbb{N}\ \big(\lambda[i+1] = \delta(\lambda[i], t(\lambda[i]))\big)\}$$



P: the players always show same sides of their coins

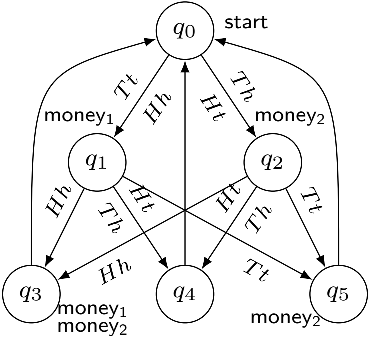
s_1: always show "heads"

Semantics of ATLP

$M, q \models \langle\langle A \rangle\rangle \gamma$ iff there is a strategy s_A consistent with Υ such that $M, \lambda \models \gamma$ for all $\lambda \in out_\Upsilon(q, s_A)$

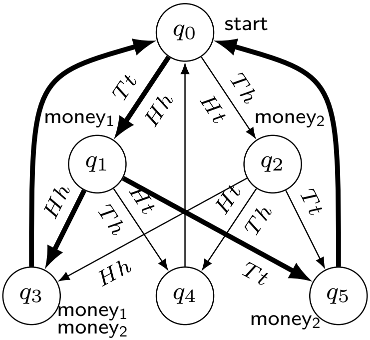
$M, q \models (\textbf{set-pl}\ \omega)\varphi$ iff $M^\omega, q \models \varphi$ where the new model M^ω is equal to M but the new set Υ^ω of plausible strategy profiles is set to $\|\omega\|_q$.

Example: Pennies Game



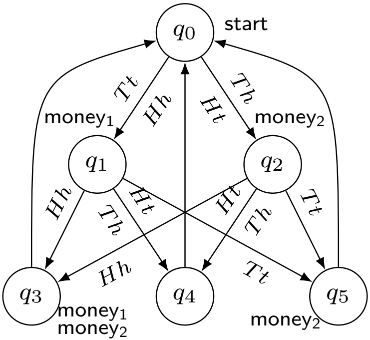
$$M, q_0 \models (\textbf{set-pl}\ \text{sameside})\langle\langle\emptyset\rangle\rangle\bigcirc \text{money}_1$$

Example: Pennies Game



$M, q_0 \models$ (**set-pl** sameside)$\langle\langle\emptyset\rangle\rangle\bigcirc$money$_1$

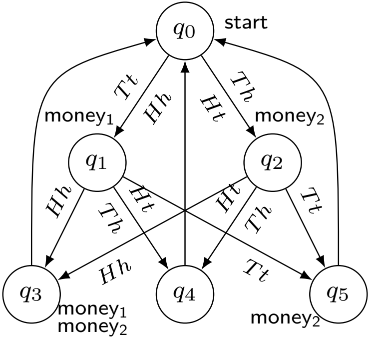
Example: Pennies Game



$M, q_0 \models$ (**set-pl** sameside)$\langle\langle\emptyset\rangle\rangle\bigcirc$money$_1$

$M, q_0 \models$ (**set-pl** ω_{NE})$\langle\langle 2\rangle\rangle\Diamond$money$_2$

Example: Pennies Game



$M, q_0 \models$ (**set-pl** sameside)$\langle\langle\emptyset\rangle\rangle\bigcirc$money$_1$

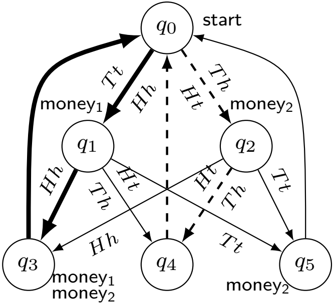
$M, q_0 \models$ (**set-pl** ω_{NE})$\langle\langle 2\rangle\rangle\Diamond$money$_2$

What is a Nash equilibrium in this game?

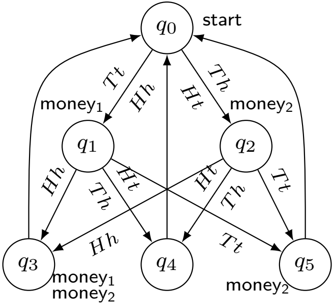
We need some kind of winning criteria!

Agent 1 “wins”, if $\gamma_1 \equiv \Box(\neg \mathsf{start} \rightarrow \mathsf{money}_1)$ is satisfied.
Agent 2 “wins”, if $\gamma_2 \equiv \Diamond \mathsf{money}_2$ is satisfied.

Agent 1 "wins", if $\gamma_1 \equiv \Box(\neg \mathsf{start} \rightarrow \mathsf{money}_1)$ is satisfied.
Agent 2 "wins", if $\gamma_2 \equiv \Diamond \mathsf{money}_2$ is satisfied.

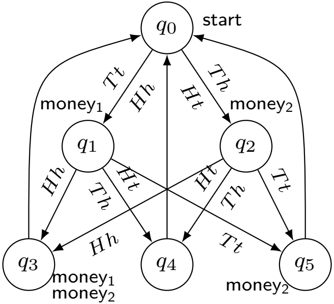


Agent 1 "wins", if $\gamma_1 \equiv \Box(\neg \mathsf{start} \rightarrow \mathsf{money}_1)$ is satisfied.
Agent 2 "wins", if $\gamma_2 \equiv \Diamond \mathsf{money}_2$ is satisfied.



$\gamma_1 \backslash \gamma_2$	hh	ht	th	tt
HH	$\mathbf{1, 1}$	$0, 0$	$0, 1$	$0, 1$
HT	$0, 0$	$0, 1$	$0, 1$	$0, 1$
TH	$0, 1$	$0, 1$	$\mathbf{1, 1}$	$0, 0$
TT	$0, 1$	$0, 1$	$0, 0$	$0, 1$

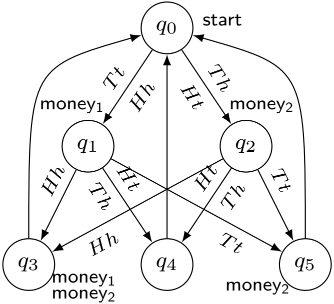
Agent 1 "wins", if $\gamma_1 \equiv \Box(\neg\mathsf{start} \rightarrow \mathsf{money}_1)$ is satisfied.
Agent 2 "wins", if $\gamma_2 \equiv \Diamond\mathsf{money}_2$ is satisfied.



$\gamma_1 \backslash \gamma_2$	hh	ht	th	tt
HH	**1, 1**	0, 0	0, 1	0, 1
HT	0, 0	0, 1	0, 1	0, 1
TH	0, 1	0, 1	**1, 1**	0, 0
TT	0, 1	0, 1	0, 0	0, 1

Now we have a qualitative notion of success.

Agent 1 "wins", if $\gamma_1 \equiv \Box(\neg\mathsf{start} \rightarrow \mathsf{money}_1)$ is satisfied.
Agent 2 "wins", if $\gamma_2 \equiv \Diamond\mathsf{money}_2$ is satisfied.



$\gamma_1 \backslash \gamma_2$	hh	ht	th	tt
HH	**1, 1**	0, 0	0, 1	0, 1
HT	0, 0	0, 1	0, 1	0, 1
TH	0, 1	0, 1	**1, 1**	0, 0
TT	0, 1	0, 1	0, 0	0, 1

Now we have a qualitative notion of success.

$$M, q_0 \models (\textbf{set-pl}\ \omega_{NE})\langle\langle 2\rangle\rangle\Box(\neg\mathsf{start} \rightarrow \mathsf{money}_1)$$

where $\|\omega_{NE}\|_{q_0}$ = "all profiles belonging to grey cells".

What about games with non-binary payoffs?

What about games with non-binary payoffs?

- Option 1: instead of a single winning condition, we use a list of conditions to encode preferences over outcomes

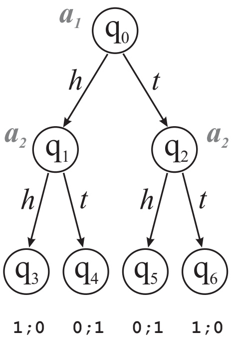
What about games with non-binary payoffs?

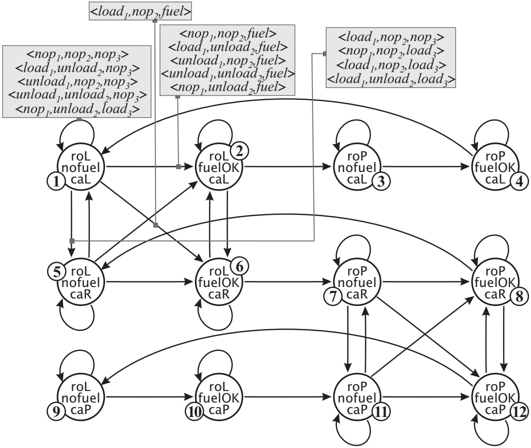
- Option 1: instead of a single winning condition, we use a list of conditions to encode preferences over outcomes
- Option 2: we use the construction by Baltag to embed utilities in CGS, and then refer to temporal patterns of utilities

What about games with non-binary payoffs?

- Option 1: instead of a single winning condition, we use a list of conditions to encode preferences over outcomes
- Option 2: we use the construction by Baltag to embed utilities in CGS, and then refer to temporal patterns of utilities
- Simplest characteristic of such patterns: the utility obtained eventually at the end of the game

Extensive Games as Concurrent Game Structures





The Construction

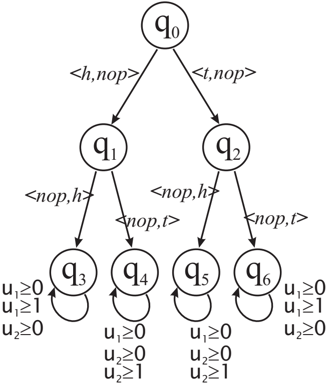
- Model terminal nodes as "sink" states
- Emulate utilities with propositions
- $M, q \models u_a \geq v$: "a gets at least v in state q"

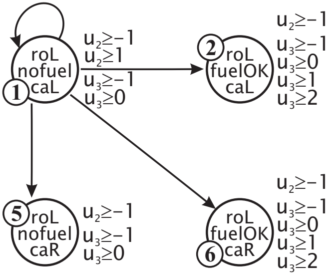
The Construction

- Model terminal nodes as "sink" states
- Emulate utilities with propositions
- $M, q \models u_a \geq v$: "a gets at least v in state q"

Now: CGS are a generalization of extensive games

Extensive Games as Concurrent Game Structures





Temporalized Solution Concepts

- Outcome of a game:
- in an extensive game: single utility value
- in CGS: infinite temporal path

Temporalized Solution Concepts

- Outcome of a game:
- in an extensive game: single utility value
- in CGS: infinite temporal path

  $\rightsquigarrow$ temporal pattern of utilities

Temporalized Solution Concepts

- Outcome of a game:
- in an extensive game: single utility value
- in CGS: infinite temporal path
  $\rightsquigarrow$ temporal pattern of utilities
- (In extensive games, paths are identical to states – in CGS not!)

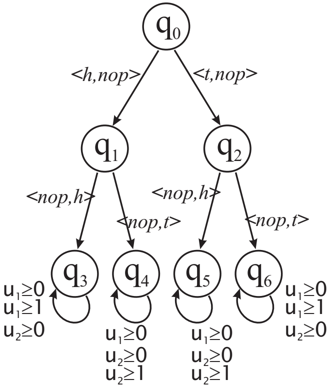
Temporalized Solution Concepts

- Outcome of a game:
- in an extensive game: single utility value
- in CGS: infinite temporal path
  $\rightsquigarrow$ temporal pattern of utilities
- (In extensive games, paths are identical to states – in CGS not!)
- We need to define the payoff for agent a of path λ
- Qualitative approach: see previous slides
- Quantitative approach: guaranteed utility ($\rightsquigarrow$ a gets always at least u), achievable utility ($\rightsquigarrow$ a gets eventually at least u)...?
- $\Box u_a \geq 1$, $\Diamond u_a \geq 1$, $\bigcirc u_a \geq 1$, ...

Temporalized Solution Concepts

- Outcome of a game:
- in an extensive game: single utility value
- in CGS: infinite temporal path
  $\rightsquigarrow$ temporal pattern of utilities
- (In extensive games, paths are identical to states – in CGS not!)
- We need to define the payoff for agent a of path λ
- Qualitative approach: see previous slides
- Quantitative approach: guaranteed utility ($\rightsquigarrow$ a gets always at least u), achievable utility ($\rightsquigarrow$ a gets eventually at least u)...?
- $\Box u_a \geq 1$, $\Diamond u_a \geq 1$, $\bigcirc u_a \geq 1$, ...
- ...Temporalized solution concepts (parameterized with temporal operators)

Temporalized Solution Concepts



$$M, q_0 \models (\textbf{set-pl}\ \omega_{NE})\langle\langle 2\rangle\rangle\Diamond(u_2 \geq 1)$$

7.2 Plausibility Specifications

How to Obtain Plausibility Terms?

Plausibility terms: abstract labels, no structure!

How to Obtain Plausibility Terms?

Plausibility terms: abstract labels, no structure!

Idea

Formulae that describe plausible strategies!

$(\textbf{set-pl}\ \sigma.\theta)\varphi$: "suppose that θ characterizes rational strategy profiles, then φ holds".

How to Obtain Plausibility Terms?

Plausibility terms: abstract labels, no structure!

Idea

Formulae that describe plausible strategies!

$(\textbf{set-pl}\ \sigma.\theta)\varphi$: "suppose that θ characterizes rational strategy profiles, then φ holds".

We need to "plug in" logical characterizations of rationality assumptions $\rightsquigarrow$ CATL

How to Obtain Plausibility Terms?

Plausibility terms: abstract labels, no structure!

Idea

Formulae that describe plausible strategies!

$(\textbf{set-pl}\ \sigma.\theta)\varphi$: "suppose that θ characterizes rational strategy profiles, then φ holds".

We need to "plug in" logical characterizations of rationality assumptions $\rightsquigarrow$ CATL

But: in fact, we can use ATLP instead!

How to Obtain Plausibility Terms?

Plausibility terms: abstract labels, no structure!

Idea

Formulae that describe plausible strategies!

$(\textbf{set-pl}\ \sigma.\theta)\varphi$: "suppose that θ characterizes rational strategy profiles, then φ holds".

We need to "plug in" logical characterizations of rationality assumptions $\rightsquigarrow$ CATL

But: in fact, we can use ATLP instead!
$\rightsquigarrow$ The same language for characterizing rationality and reasoning about the outcome of rational play

Sometimes quantifiers are needed...

E.g.: $(\textbf{set-pl}\ \sigma.\ \forall\sigma'\ dominates(\sigma,\sigma'))$

ATLP: Extending the Syntax

Definition 7.1 (Logics $\mathcal{L}^k_{ATLP}$)

Let Ω be a set of primitive plausibility terms, and Var be a set of strategic variables (with typical element ω). $\mathcal{L}^k_{ATLP}(\mathbb{A}\mathrm{gt}, \Pi, Var, \Omega)$ are defined recursively:

- $\mathcal{L}^0_{ATLP}(\mathbb{A}\mathrm{gt}, \Pi, Var, \Omega) = \mathcal{L}^{base}_{ATLP}(\mathbb{A}\mathrm{gt}, \Pi, \Omega_0)$
  where $\Omega_0 = \mathcal{T}(\Omega)$;
- $\mathcal{L}^k_{ATLP}(\mathbb{A}\mathrm{gt}, \Pi, Var, \Omega) = \mathcal{L}^{base}_{ATLP}(\mathbb{A}\mathrm{gt}, \Pi, \Omega_k)$, where:
  - $\Omega_k := \mathcal{T}(\Omega_{k-1} \cup \Omega^k)$,
  - $\Omega^k := \{\sigma_1.(Q_2\sigma_2)\dots(Q_n\sigma_n)\varphi \mid n \in \mathbb{N}, \forall i\ (1 \le i \le n \Rightarrow \sigma_i \in Var,\ Q_i \in \{\forall, \exists\}, \varphi \in \mathcal{L}^{base}_{ATLP}(\mathbb{A}\mathrm{gt}, \Pi, \mathcal{T}(\Omega_{k-1} \cup \{\sigma_1, \dots, \sigma_n\})))\}$.

ATLP: Extending the Syntax

Definition 7.1 (Logics $\mathcal{L}^k_{ATLP}$)

Let Ω be a set of primitive plausibility terms, and Var be a set of strategic variables (with typical element ω). $\mathcal{L}^k_{ATLP}(\mathbb{A}\mathrm{gt}, \Pi, Var, \Omega)$ are defined recursively:

- $\mathcal{L}^0_{ATLP}(\mathbb{A}\mathrm{gt}, \Pi, Var, \Omega) = \mathcal{L}^{base}_{ATLP}(\mathbb{A}\mathrm{gt}, \Pi, \Omega_0)$ where $\Omega_0 = \mathcal{T}(\Omega)$;
- $\mathcal{L}^k_{ATLP}(\mathbb{A}\mathrm{gt}, \Pi, Var, \Omega) = \mathcal{L}^{base}_{ATLP}(\mathbb{A}\mathrm{gt}, \Pi, \Omega_k)$, where:
  - $\Omega_k := \mathcal{T}(\Omega_{k-1} \cup \Omega^k)$,
  - $\Omega^k := \{\sigma_1.(Q_2\sigma_2)\dots(Q_n\sigma_n)\varphi \mid n \in \mathbb{N}, \forall i\,(1 \le i \le n \Rightarrow \sigma_i \in Var,\ Q_i \in \{\forall, \exists\}, \varphi \in \mathcal{L}^{base}_{ATLP}(\mathbb{A}\mathrm{gt}, \Pi, \mathcal{T}(\Omega_{k-1} \cup \{\sigma_1, \dots, \sigma_n\})))\,\}$.

The set of ATLP formulae with arbitrary finite nesting of plausibility terms is defined by $\mathcal{L}^\infty_{ATLP}$

Formal Semantics...

Formal Semantics...

Formal Semantics... let's jump over it

7.3 Characterizations

Qualitative Characterization of Nash Equilibrium

σ_a is a's best response to σ (wrt $\overrightarrow{\gamma}$):

$$BR_a^{\overrightarrow{\gamma}}(\sigma) \equiv (\textbf{set-pl } \sigma[\mathbb{A}\mathrm{gt} \setminus \{a\}])\big(\langle\langle a \rangle\rangle \gamma_a \rightarrow (\textbf{set-pl } \sigma)\langle\langle \emptyset \rangle\rangle \gamma_a\big)$$

Qualitative Characterization of Nash Equilibrium

σ_a is a's best response to σ (wrt $\overrightarrow{\gamma}$):

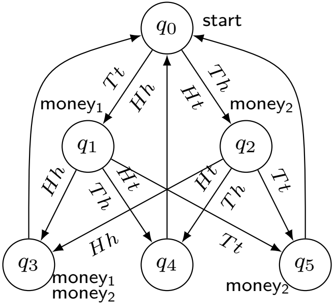
$$BR_a^{\overrightarrow{\gamma}}(\sigma) \equiv (\textbf{set-pl } \sigma[\mathbb{A}\mathrm{gt} \setminus \{a\}])\big(\langle\langle a \rangle\rangle \gamma_a \rightarrow (\textbf{set-pl } \sigma)\langle\langle \emptyset \rangle\rangle \gamma_a\big)$$

σ is a Nash equilibrium:

$$NE^{\overrightarrow{\gamma}}(\sigma) \equiv \bigwedge_{a \in \mathbb{A}\mathrm{gt}} BR_a^{\overrightarrow{\gamma}}(\sigma)$$

Example: Pennies Game revisited

$\gamma_1 \equiv \Box(\neg\text{start} \rightarrow \text{money}_1)$; $\gamma_2 \equiv \Diamond\text{money}_2$



$\gamma_1 \backslash \gamma_2$	hh	ht	th	tt
HH	**1, 1**	0, 0	0, 1	0, 1
HT	0, 0	0, 1	0, 1	0, 1
TH	0, 1	0, 1	**1, 1**	0, 0
TT	0, 1	0, 1	0, 0	0, 1

$$M_1, q_0 \models (\textbf{set-pl}\ \sigma.NE^{\gamma_1,\gamma_2}(\sigma))\langle\langle 2\rangle\rangle\Box(\neg\text{start} \rightarrow \text{money}_1)$$

...where $NE^{\gamma_1,\gamma_2}(\sigma)$ is defined as on the last slide

Characterizations of Other Solution Concepts

σ is a subgame perfect Nash equilibrium:

$$SPN^{\vec{\gamma}}(\sigma) \equiv \langle\langle\emptyset\rangle\rangle\Box NE^{\vec{\gamma}}(\sigma)$$

σ is Pareto optimal:

$$PO^{\vec{\gamma}}(\sigma) \equiv \forall\sigma'\Big(\bigwedge_{a\in\mathbb{A}\mathrm{gt}}((\textbf{set-pl }\sigma')\langle\langle\emptyset\rangle\rangle\gamma_a \rightarrow (\textbf{set-pl }\sigma)\langle\langle\emptyset\rangle\rangle\gamma_a) \vee \bigvee_{a\in\mathbb{A}\mathrm{gt}}((\textbf{set-pl }\sigma)\langle\langle\emptyset\rangle\rangle\gamma_a \wedge \neg(\textbf{set-pl }\sigma')\langle\langle\emptyset\rangle\rangle\gamma_a\Big).$$

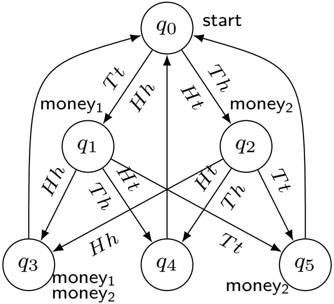
Characterizations of Other Solution Concepts

σ is undominated:

$$
\begin{aligned}
UNDOM^{\vec{\gamma}}(\sigma) \quad \equiv \quad & \forall\sigma_1\forall\sigma_2\exists\sigma_3 \\
& \Big(\big((\textbf{set-pl } \langle\sigma_1^{\{a\}}, \sigma_2^{\mathbb{A}\mathrm{gt}\setminus\{a\}}\rangle)\langle\!\langle\emptyset\rangle\!\rangle\gamma_a \rightarrow \\
& \qquad (\textbf{set-pl } \langle\sigma^{\{a\}}, \sigma_2^{\mathbb{A}\mathrm{gt}\setminus\{a\}}\rangle)\langle\!\langle\emptyset\rangle\!\rangle\gamma_a\big) \\
& \vee\big((\textbf{set-pl } \langle\sigma^{\{a\}}, \sigma_3^{\mathbb{A}\mathrm{gt}\setminus\{a\}}\rangle)\langle\!\langle\emptyset\rangle\!\rangle\gamma_a \wedge \\
& \qquad \neg(\textbf{set-pl } \langle\sigma_1^{\{a\}}, \sigma_3^{\mathbb{A}\mathrm{gt}\setminus\{a\}}\rangle)\langle\!\langle\emptyset\rangle\!\rangle\gamma_a\big)\Big).
\end{aligned}
$$

Example: Pennies Game again

$\gamma_1 \equiv \Box(\neg \mathsf{start} \rightarrow \mathsf{money}_1)$; $\gamma_2 \equiv \Diamond \mathsf{money}_2$.



$\gamma_1 \backslash \gamma_2$	hh	ht	th	tt
HH	**1, 1**	0, 0	0, 1	0, 1
HT	0, 0	0, 1	0, 1	0, 1
TH	0, 1	0, 1	**1, 1**	0, 0
TT	0, 1	0, 1	0, 0	0, 1

$$M, q_0 \models (\textbf{set-pl}\ \sigma.PO^{\gamma_1,\gamma_2}(\sigma))\langle\langle\emptyset\rangle\rangle\Diamond(\mathsf{money}_1 \wedge \mathsf{money}_2)$$

Theorem 7.2

Let M be a CGSP, q a state in M, and $\overrightarrow{\eta} = \langle \eta_1, \ldots, \eta_k \rangle$ a vector of path formulae (winning conditions). Moreover, let $\Gamma(M, q, \overrightarrow{\eta})$ be the strategic game obtained from M, q by assigning strategy profiles with binary payoffs according to $\overrightarrow{\eta}$. Then the following holds:

1. $\| \sigma.NE^{\eta}(\sigma) \|_{M,q}$ *denotes the set of Nash equilibria in* $\Gamma(M, q, \overrightarrow{\eta})$*;*
2. $\| \sigma.PO^{\eta}(\sigma) \|_{M,q}$ *denotes the set of Pareto optimal strategy profiles in* $\Gamma(M, q, \overrightarrow{\eta})$*;*
3. $\| \sigma.UNDOM^{\eta}(\sigma) \|_{M,q}$ *denotes the set of undominated strategies in* $\Gamma(M, q, \overrightarrow{\eta})$*.*
4. $\| \sigma.SPN^{\eta}(\sigma) \|_{M,q}$ *denotes the set of strategy profiles that are in Nash equilibrium for every* $\Gamma(M, q', \overrightarrow{\eta})$ *(for all reachable* q'*)*

Temporalized Solution Concepts

Nash Equilibrium:

$$BR_a^T(\sigma) \equiv (\mathbf{str}_{\mathbb{A}\mathrm{gt}\setminus A}\,\sigma[\mathbb{A}\mathrm{gt}\setminus\{a\}])$$
$$(\bigwedge_{v\in U}(\langle\langle a\rangle\rangle T(u_a \geq v)) \rightarrow (\mathbf{str}_a\,\sigma[a])\langle\langle\emptyset\rangle\rangle T(u_a \geq v))$$

Temporalized Solution Concepts

Nash Equilibrium:

$$BR_a^T(\sigma) \equiv (\mathbf{str}_{\mathbb{A}\mathrm{gt}\setminus A}\,\sigma[\mathbb{A}\mathrm{gt}\setminus\{a\}]) \\ (\bigwedge_{v\in U}(\langle\!\langle a\rangle\!\rangle T(u_a \geq v)) \rightarrow (\mathbf{str}_a\,\sigma[a])\langle\!\langle\emptyset\rangle\!\rangle T(u_a \geq v))$$

$$NE^T(\sigma) \equiv \bigwedge_{a\in\mathbb{A}\mathrm{gt}} BR_a^T(\sigma)$$

Temporalized Solution Concepts

Nash Equilibrium:

$$BR_a^T(\sigma) \equiv (\mathbf{str}_{\mathbb{A}\mathrm{gt}\setminus A}\, \sigma[\mathbb{A}\mathrm{gt} \setminus \{a\}])$$
$$\Big(\bigwedge_{v \in U} (\langle\langle a \rangle\rangle T(u_a \geq v)) \rightarrow (\mathbf{str}_a\, \sigma[a])\langle\langle \emptyset \rangle\rangle T(u_a \geq v)\Big)$$

$$NE^T(\sigma) \equiv \bigwedge_{a \in \mathbb{A}\mathrm{gt}} BR_a^T(\sigma)$$

$$SPN^T(\sigma) \equiv \langle\langle \emptyset \rangle\rangle \Box NE^T(\sigma)$$

Temporalized Solution Concepts

$$PO^T(\sigma) \quad \equiv \quad \forall\sigma'\Big(\bigwedge_{a\in\mathbb{A}\mathrm{gt}}\bigwedge_{v\in U}((\textbf{set-pl } \sigma')\mathbf{Pl}\,\langle\langle\emptyset\rangle\rangle T(u_a \geq v) \rightarrow$$
$$(\textbf{set-pl } \sigma)\mathbf{Pl}\,\langle\langle\emptyset\rangle\rangle T(u_a \geq v)) \vee$$
$$\bigvee_{a\in\mathbb{A}\mathrm{gt}}\bigvee_{v\in U}((\textbf{set-pl } \sigma)\mathbf{Pl}\,\langle\langle\emptyset\rangle\rangle T(u_a \geq v) \wedge$$
$$\neg(\textbf{set-pl } \sigma')\mathbf{Pl}\,\langle\langle\emptyset\rangle\rangle T(u_a \geq v))\Big).$$

Temporalized Solution Concepts

$$UNDOM^T(\sigma) \equiv \quad \forall\sigma_1\forall\sigma_2\exists\sigma_3$$

$$\Big(\bigwedge_{v\in U}\big((\textbf{set-pl}\ \langle\sigma_1^{\{a\}},\sigma_2^{\mathbb{A}\mathrm{gt}\setminus\{a\}}\rangle)\mathbf{Pl}\,\langle\!\langle\emptyset\rangle\!\rangle T(u_a \geq v) \rightarrow$$

$$(\textbf{set-pl}\ \langle\sigma^{\{a\}},\sigma_2^{\mathbb{A}\mathrm{gt}\setminus\{a\}}\rangle)\mathbf{Pl}\,\langle\!\langle\emptyset\rangle\!\rangle T(u_a \geq v)\big)$$

$$\vee \bigvee_{v\in U}\big((\textbf{set-pl}\ \langle\sigma^{\{a\}},\sigma_3^{\mathbb{A}\mathrm{gt}\setminus\{a\}}\rangle)\mathbf{Pl}\,\langle\!\langle\emptyset\rangle\!\rangle T(u_a \geq v) \wedge$$

$$\neg(\textbf{set-pl}\ \langle\sigma_1^{\{a\}},\sigma_3^{\mathbb{A}\mathrm{gt}\setminus\{a\}}\rangle)\mathbf{Pl}\,\langle\!\langle\emptyset\rangle\!\rangle T(u_a \geq v)\big)\Big).$$

Temporalized Solution Concepts

Theorem 7.3

Let Γ be an extensive game with a finite set of utilities. Then the following holds:

1. $s \in \|\sigma.NE^{\Diamond}(\sigma)\|_{M(\Gamma),\emptyset}$ *iff* s *is a Nash equilibrium in* Γ*;*
2. $s \in \|\sigma.SPN^{\Diamond}(\sigma)\|_{M(\Gamma),\emptyset}$ *iff* s *is a subgame perfect Nash equilibrium in* Γ*;*
3. $s \in \|\sigma.PO^{\Diamond}(\sigma)\|_{M(\Gamma),\emptyset}$ *iff* s *is Pareto optimal in* Γ*;*
4. $s \in \|\sigma.UNDOM^{\Diamond}(\sigma)\|_{M(\Gamma),\emptyset}$ *iff* s *is undominated in* Γ*.*

7.4 Model Checking

Solving Games through Model Checking ATLP

- Concurrent game structure = generalized extensive game
- Plausibility specification $\rightsquigarrow$ solution concept

Solving Games through Model Checking ATLP

- Concurrent game structure = generalized extensive game
- Plausibility specification $\rightsquigarrow$ solution concept
- $\langle\langle A\rangle\rangle\gamma$ defines a game where A want to achieve γ

Solving Games through Model Checking ATLP

- Concurrent game structure = generalized extensive game
- Plausibility specification $\rightsquigarrow$ solution concept
- $\langle\langle A \rangle\rangle \gamma$ defines a game where A want to achieve γ
  2-player, binary, zero-sum game

Solving Games through Model Checking ATLP

- Concurrent game structure = generalized extensive game
- Plausibility specification $\rightsquigarrow$ solution concept
- $\langle\langle A\rangle\rangle\gamma$ defines a game where A want to achieve γ
  2-player, binary, zero-sum game
  $\rightsquigarrow$ players, payoffs

Solving Games through Model Checking ATLP

- Concurrent game structure = generalized extensive game
- Plausibility specification $\rightsquigarrow$ solution concept
- $\langle\langle A\rangle\rangle\gamma$ defines a game where A want to achieve γ
  2-player, binary, zero-sum game
  $\rightsquigarrow$ players, payoffs
- Model checking formulae of ATLP $\rightsquigarrow$ solving games

Model checking complexity of ATLP

	0	1	$\dots$	i	$\dots$	∞
$\mathcal{L}^{basic}_{ATLP}$	$\mathbf{P}$	-	$\dots$	-	$\dots$	-
$\mathcal{L}^{0}_{ATLP}$	$\mathbf{P}$	-	$\dots$	-	$\dots$	-
$\mathcal{L}^{1}_{ATLP}$	$\mathbf{\Delta^P_3}$	$\mathbf{\Delta^P_4}$	$\dots$	$\mathbf{\Delta^P_{i+3}}$	$\dots$	PSPACE
$\mathcal{L}^{2}_{ATLP}$	$\mathbf{\Delta^P_4}$	$\mathbf{\Delta^P_6}$	$\dots$	$\mathbf{\Delta^P_{5+i-\max\{0,1-i\}}}$	$\dots$	PSPACE
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$\mathcal{L}^{k}_{ATLP}$ $i > k+1$	$\mathbf{\Delta^P_{k+2}}$	$\mathbf{\Delta^P_{k+4}}$	$\dots$	$\mathbf{\Delta^P_{i+2k+1-\max\{0,k-i-1\}}}$	$\dots$	PSPACE

Model checking complexity of ATLP

	0	1	...	i	...	∞
$\mathcal{L}_{ATLP}^{basic}$	$\mathbf{P}$	-	...	-	...	-
$\mathcal{L}_{ATLP}^{0}$	$\mathbf{P}$	-	...	-	...	-
$\mathcal{L}_{ATLP}^{1}$	$\mathbf{\Delta_3^P}$	$\mathbf{\Delta_4^P}$	...	$\mathbf{\Delta_{i+3}^P}$	...	PSPACE
$\mathcal{L}_{ATLP}^{2}$	$\mathbf{\Delta_4^P}$	$\mathbf{\Delta_6^P}$	...	$\mathbf{\Delta_{5+i-\max\{0,1-i\}}^P}$	...	PSPACE
...	...	...	...	...	...	...
$\mathcal{L}_{ATLP}^{k}$ $i > k+1$	$\mathbf{\Delta_{k+2}^P}$	$\mathbf{\Delta_{k+4}^P}$	...	$\mathbf{\Delta_{i+2k+1-\max\{0,k-i-1\}}^P}$	...	PSPACE

SAT/mechanism design complexity: open!

7.5 References

[Bulling *et al.*2008] N. Bulling, W.Jamroga, and J. Dix.
Reasoning about rational agents in ATLP.
Technical Report IfI-08-03, Clausthal University of Technology, 2008.

[van der Hoek et al. 2004] W. van der Hoek, S. van Otterloo, and M. Wooldridge.
Preferences in Game Logics.
Proceedings of AAMAS-04, 152–159, 2004.

Model checking

8.1 Model Checking Time and Strategies

Model Checking

Model checking: Does φ hold in model $\mathcal{M}$ and state q?

Model Checking

Model checking: Does φ hold in model $\mathcal{M}$ and state q?

Two perspectives to model checking MAS:

Model Checking

Model checking: Does φ hold in model $\mathcal{M}$ and state q?

Two perspectives to model checking MAS:

- Model represents the view of an objective observer
- Formula: specification to be met

Model Checking

Model checking: Does φ hold in model $\mathcal{M}$ and state q?

Two perspectives to model checking MAS:

Verification

- Model represents the view of an objective observer
- Formula: specification to be met

Model Checking

Model checking: Does φ hold in model $\mathcal{M}$ and state q?

Two perspectives to model checking MAS:

Verification

- Model represents the view of an objective observer
- Formula: specification to be met

- Model represents the subjective view of an agent
- Formula: goal to be achieved

Model Checking

Model checking: Does φ hold in model $\mathcal{M}$ and state q?

Two perspectives to model checking MAS:

Verification

- Model represents the view of an objective observer
- Formula: specification to be met

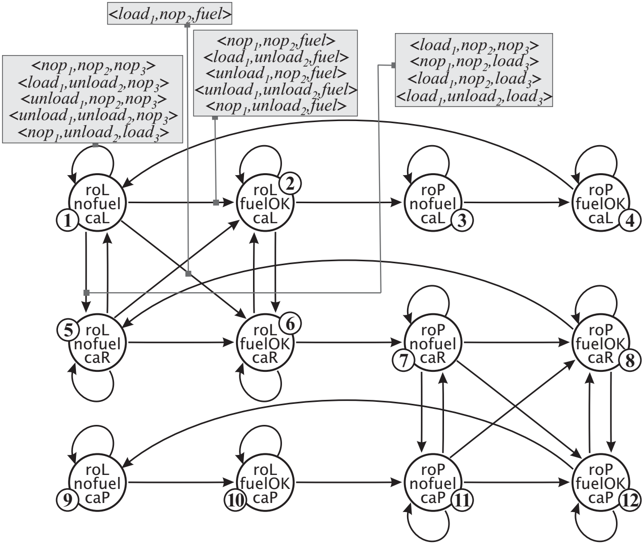
Planning

- Model represents the subjective view of an agent
- Formula: goal to be achieved

function $mcheck(\mathcal{M}, \varphi)$.
Model checking formulae of ATL.
Returns the exact subset of St for which formula φ holds.

case $\varphi \equiv p$ **:** return $\{q \in St \mid p \in \pi(q)\}$
case $\varphi \equiv \neg\psi$ **:** return $St \setminus mcheck(\mathcal{M}, \psi)$
case $\varphi \equiv \psi_1 \wedge \psi_2$ **:** return $mcheck(\mathcal{M}, \psi_1) \cap mcheck(\mathcal{M}, \psi_2)$
case $\varphi \equiv \langle\langle A \rangle\rangle \bigcirc \psi$ **:** return $pre(A, mcheck(\mathcal{M}, \psi))$
case $\varphi \equiv \langle\langle A \rangle\rangle \Box \psi$ **:**
$Q_1 := Q$; $Q_2 := Q_3 := mcheck(\mathcal{M}, \psi)$;
while $Q_1 \not\subseteq Q_2$ **do** $Q_1 := Q_1 \cap Q_2$; $Q_2 := pre(A, Q_1) \cap Q_3$ **od;**
return Q_1
case $\varphi \equiv \langle\langle A \rangle\rangle \psi_1 \, \mathcal{U} \psi_2$ **:**
$Q_1 := \emptyset$; $Q_2 := mcheck(\mathcal{M}, \psi_2)$; $Q_3 := mcheck(\mathcal{M}, \psi_1)$;
while $Q_2 \not\subseteq Q_1$ **do** $Q_1 := Q_1 \cup Q_2$; $Q_2 := pre(A, Q_1) \cap Q_3$ **od;**
return Q_1
end case

Example: Simple Rocket Domain



Example: Simple Rocket Domain

- Verification example: we want to make sure that agents 1 and 3 can move the cargo to the other location.
- $\mathsf{caL} \rightarrow \langle\langle 1, 3 \rangle\rangle \Diamond \mathsf{caP}$

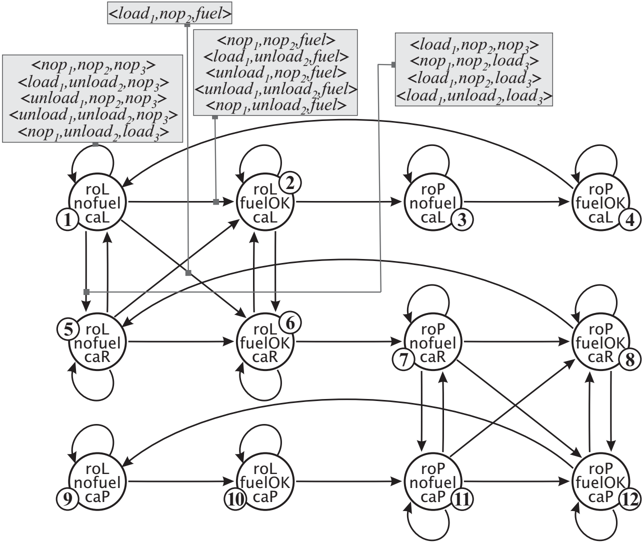
Example: Simple Rocket Domain

- Verification example: we want to make sure that agents 1 and 3 can move the cargo to the other location.
- $\mathsf{caL} \rightarrow \langle\langle 1,3\rangle\rangle\Diamond\mathsf{caP} \;\wedge\; \mathsf{caP} \rightarrow \langle\langle 1,3\rangle\rangle\Diamond\mathsf{caL}$

Example: Simple Rocket Domain

- Verification example: we want to make sure that agents 1 and 3 can move the cargo to the other location.
- $\mathsf{caL} \rightarrow \langle\langle 1,3\rangle\rangle\Diamond\mathsf{caP} \;\wedge\; \mathsf{caP} \rightarrow \langle\langle 1,3\rangle\rangle\Diamond\mathsf{caL}$
  $\wedge\; \mathsf{caR} \rightarrow (\langle\langle 1,3\rangle\rangle\Diamond\mathsf{caL} \wedge \langle\langle 1,3\rangle\rangle\Diamond\mathsf{caP})$

Example: Simple Rocket Domain



Nice results: model checking CTL and ATL is tractable!

Nice results: model checking CTL and ATL is tractable!

Theorem (Clarke, Emerson & Sistla 1986)

CTL model checking is P-complete, and can be done in time linear in the size of the model and the length of the formula.

Nice results: model checking CTL and ATL is tractable!

Theorem (Clarke, Emerson & Sistla 1986)

CTL model checking is P-complete, and can be done in time linear in the size of the model and the length of the formula.

Theorem (Alur, Kupferman & Henzinger 1998)

ATL model checking is P-complete, and can be done in time linear in the size of the model and the length of the formula.

So... Let's model-check!

So... Let's model-check!

Not as easy as it seems...

Some Complexity Classes

- P: problems solvable in polynomial time by a deterministic Turing machine

Some Complexity Classes

- P: problems solvable in polynomial time by a deterministic Turing machine
- NP: problems solvable in polynomial time by a non-deterministic Turing machine

Some Complexity Classes

- P: problems solvable in polynomial time by a deterministic Turing machine
- NP: problems solvable in polynomial time by a non-deterministic Turing machine
- $\mathbf{\Sigma^P_n}/\mathbf{\Pi^P_n}/\mathbf{\Delta^P_n}$: problems solvable in polynomial time with use of adaptive queries to an n-level oracle

Some Complexity Classes

- P: problems solvable in polynomial time by a deterministic Turing machine
- NP: problems solvable in polynomial time by a non-deterministic Turing machine
- $\mathbf{\Sigma^P_n}/\mathbf{\Pi^P_n}/\mathbf{\Delta^P_n}$: problems solvable in polynomial time with use of adaptive queries to an n-level oracle
- PSPACE: problems solvable by queries to a multilevel oracle with unbounded "height"

Some Complexity Classes

- P: problems solvable in polynomial time by a deterministic Turing machine
- NP: problems solvable in polynomial time by a non-deterministic Turing machine
- $\mathbf{\Sigma_n^P}/\mathbf{\Pi_n^P}/\mathbf{\Delta_n^P}$: problems solvable in polynomial time with use of adaptive queries to an n-level oracle
- PSPACE: problems solvable by queries to a multilevel oracle with unbounded “height”
- EXPTIME: problems solvable in exponential time

Some Complexity Classes

- P: problems solvable in polynomial time by a deterministic Turing machine
- NP: problems solvable in polynomial time by a non-deterministic Turing machine
- $\mathbf{\Sigma^P_n}/\mathbf{\Pi^P_n}/\mathbf{\Delta^P_n}$: problems solvable in polynomial time with use of adaptive queries to an n-level oracle
- PSPACE: problems solvable by queries to a multilevel oracle with unbounded "height"
- EXPTIME: problems solvable in exponential time

What is this about?

Some Complexity Classes

- P: problems solvable in polynomial time by a deterministic Turing machine
- NP: problems solvable in polynomial time by a non-deterministic Turing machine
- $\Sigma_n^P / \Pi_n^P / \Delta_n^P$: problems solvable in polynomial time with use of adaptive queries to an n-level oracle
- PSPACE: problems solvable by queries to a multilevel oracle with unbounded "height"
- EXPTIME: problems solvable in exponential time

What is this about?

Scalability!

Complexity of Model Checking Temporal and Strategic Logics

	m, l
CTL	P-complete
LTL	PSPACE-complete
CTL*	PSPACE-complete
ATL	P-complete
ATL*	PSPACE-complete

Complexity of Model Checking Temporal and Strategic Logics

	m, l
CTL	P-complete
LTL	PSPACE-complete
CTL*	PSPACE-complete
ATL	P-complete
ATL*	PSPACE-complete

For strategies with perfect recall:

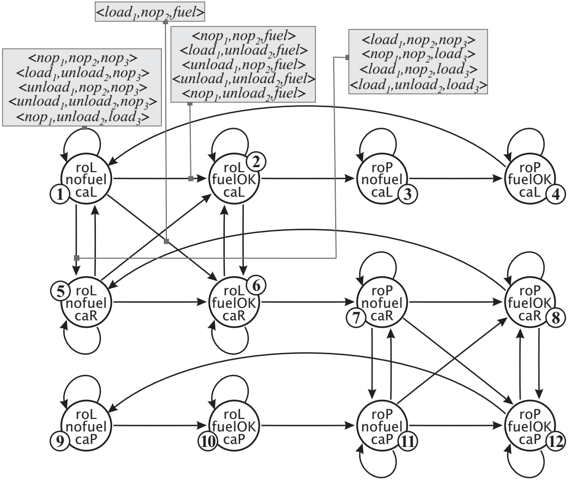
	m, l
ATL	P-complete
ATL*	2EXPTIME-complete

- Nice results: model checking CTL and ATL is tractable.

- Nice results: model checking CTL and ATL is tractable.
- But: the result is relative to the size of the model and the formula

- Nice results: model checking CTL and ATL is tractable.
- But: the result is relative to the size of the model and the formula
- Well known catch (CTL): size of models is exponential wrt a higher-level description

3 agents . . . 12 states



Do Agents Make Model Checking Explode?

Do Agents Make Model Checking Explode?

	m, l	n_{local}, l
CTL	P-complete	
LTL	PSPACE-complete	
CTL*	PSPACE-complete	
ATL	P-complete	
ATL*	PSPACE-complete	

Do Agents Make Model Checking Explode?

	m, l	n_{local}, l
CTL	P-complete	PSPACE-complete
LTL	PSPACE-complete	PSPACE-complete
CTL*	PSPACE-complete	PSPACE-complete
ATL	P-complete	
ATL*	PSPACE-complete	

Do Agents Make Model Checking Explode?

	m, l	n_{local}, l
CTL	P-complete	PSPACE-complete
LTL	PSPACE-complete	PSPACE-complete
CTL*	PSPACE-complete	PSPACE-complete
ATL	P-complete	EXPTIME-complete
ATL*	PSPACE-complete	EXPTIME-complete

Further Problems

- How is the size of a model defined?

Further Problems

- How is the size of a model defined?

  Size of M = number of transitions in M

- What if we define it as the number of states?

Further Problems

- How is the size of a model defined?

  Size of M = number of transitions in M

- What if we define it as the number of states?
- For CTL: $m = O(n^2) \leadsto$ no problem

Further Problems

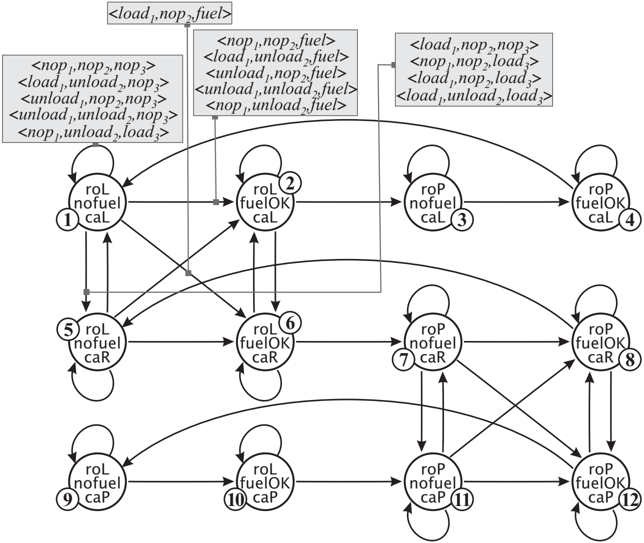
- How is the size of a model defined?

  Size of M = number of transitions in M

- What if we define it as the number of states?

- For CTL: $m = O(n^2) \leadsto$ no problem
- For ATL: transitions are labeled
- m is not bound by n^2!

3 agents . . . 12 states, 216 transitions



Do Agents Make Model Checking Explode?

- Observation: the number of transitions can be exponential in the number of agents
- $m = O(nd^k)$
- m: transitions, n: states, d: actions (decisions), k: agents

Do Agents Make Model Checking Explode?

- Observation: the number of transitions can be exponential in the number of agents
- $m = O(nd^k)$
- m: transitions, n: states, d: actions (decisions), k: agents
- What about model checking?

Do Agents Make Model Checking Explode?

- Observation: the number of transitions can be exponential in the number of agents
- $m = O(nd^k)$
- m: transitions, n: states, d: actions (decisions), k: agents
- What about model checking?

Theorem (Jamroga & Dix 2005; Laroussinie, Markey & Oreiby 2006)

ATL model checking is $\Delta_2^{\mathbf{P}}$-complete with respect to the number of states and agents.

Summary of Complexity Results

	m, l	n, k, l	n_{local}, l
CTL			
LTL			
CTL*			
ATL			
ATL*			

Summary of Complexity Results

	m, l	n, k, l	n_{local}, l
CTL	P		
LTL	PSPACE		
CTL*	PSPACE		
ATL	P		
ATL*	PSPACE		

Summary of Complexity Results

	m, l	n, k, l	n_{local}, l
CTL	P		PSPACE
LTL	PSPACE		PSPACE
CTL*	PSPACE		PSPACE
ATL	P		EXPTIME
ATL*	PSPACE		EXPTIME

Summary of Complexity Results

	m, l	n, k, l	n_{local}, l
CTL	P	P	PSPACE
LTL	PSPACE	PSPACE	PSPACE
CTL*	PSPACE	PSPACE	PSPACE
ATL	P	Δ_2^P	EXPTIME
ATL*	PSPACE	EXPTIME	EXPTIME

Looking for Moral

Main message:

- Complexity is very sensitive to the context!

Looking for Moral

Main message:

- Complexity is very sensitive to the context!
- In particular, the way we define the parameters, and measure their size, is crucial.

Even if model checking appears very easy, it can be very hard.

Even if model checking appears very easy, it can be very hard.

Still, people do automatic model checking!

Even if model checking appears very easy, it can be very hard.

Still, people do automatic model checking!

- LTL: SPIN

Even if model checking appears very easy, it can be very hard.

Still, people do automatic model checking!

- LTL: SPIN
- CTL/ATL: MOCHA, MCMAS, VeriCS

Even if model checking appears very easy, it can be very hard.

Still, people do automatic model checking!

- LTL: SPIN
- CTL/ATL: MOCHA, MCMAS, VeriCS

Even if model checking is theoretically hard, it can be feasible in practice.

8.2 Imperfect Information

Model Checking Imperfect Information Games

Recall: $\langle\!\langle A \rangle\!\rangle_{ir}$ are not fixpoint operators any more

Conjecture

Strategy for A cannot be synthesized incrementally.

Model Checking Imperfect Information Games

Recall: $\langle\langle A\rangle\rangle_{ir}$ are not fixpoint operators any more

Conjecture

Strategy for A cannot be synthesized incrementally.

Indeed...

Model Checking Imperfect Information Games

Recall: $\langle\langle A\rangle\rangle_{ir}$ are not fixpoint operators any more

Conjecture

Strategy for A cannot be synthesized incrementally.

Indeed...

Theorem (Schobbens 2004; Jamroga & Dix 2006)

Model checking ATL$_{ir}$ is Δ_2-complete in the number of transitions in the model and the length of the formula.

Proof Idea: Inclusion in Δ_2

Let $mctl(\varphi, M)$ be a CTL model checker that returns the set of all states that satisfy φ in M

$\underline{mcheck(M, q, \langle\langle A\rangle\rangle\Box\psi)}$:

1. Run $mcheck(\psi, M, q)$ for every $q \in St$, and label the states in which the answer was "yes" with an additional proposition yes (not used elsewhere).
2. Guess the best strategy of A, and "trim" model M by removing all the transitions inconsistent with the strategy (yielding a sparser model M').
3. Return "yes" if $\mathrm{img}(q, \sim_A^E) \subseteq mctl(\mathsf{A}\bigcirc\mathsf{yes}, M')$, and "no" otherwise.

Other cases: analogous

Proof Idea: Hardness (by reduction of SNSAT)

Definition (SNSAT)

Input:

$$\begin{aligned} z_1 &\equiv \exists X_1\ \varphi_1(z_1, X_1) \\ z_2 &\equiv \exists X_2\ \varphi_2(z_1, z_2, X_2) \\ &\ldots.. \\ z_p &\equiv \exists X_p\ \varphi_p(z_1, \ldots, z_{p-1}, X_p). \end{aligned}$$

Output: The truth value of z_p.

Proof Idea: Hardness (by reduction of SNSAT)

Definition (SNSAT)

Input:
$$\begin{aligned} z_1 &\equiv \exists X_1\, \varphi_1(z_1, X_1) \\ z_2 &\equiv \exists X_2\, \varphi_2(z_1, z_2, X_2) \\ &\quad \ldots\ldots \\ z_p &\equiv \exists X_p\, \varphi_p(z_1, \ldots, z_{p-1}, X_p). \end{aligned}$$

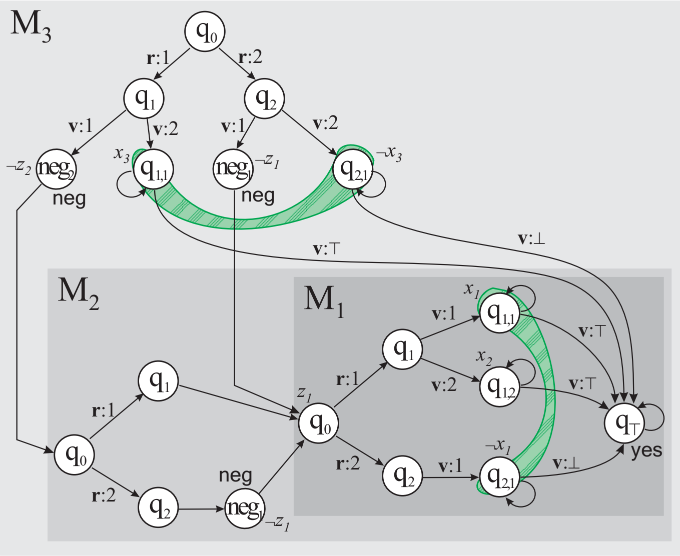
Output: The truth value of z_p.

Lemma 8.1

Let
$$\begin{aligned} \Phi_1 &\equiv \langle\langle \mathbf{v} \rangle\rangle_{ir} (\neg \mathsf{neg})\, \mathcal{U}\, \mathsf{yes}, \\ \Phi_i &\equiv \langle\langle \mathbf{v} \rangle\rangle_{ir} (\neg \mathsf{neg})\, \mathcal{U}\, (\mathsf{yes} \vee (\mathsf{neg} \wedge \mathrm{A} \bigcirc \neg \Phi_{i-1})). \end{aligned}$$

Now, for all r: z_r is true iff $M_r, q_0^r \models \Phi_r$.

Proof Idea: Hardness



Model Checking Imperfect Information Games

Corollary

Imperfect information strategies cannot be synthesized incrementally: we cannot do better than guess the whole strategy and check if it succeeds.

Model Checking Imperfect Information Games

Corollary

Imperfect information strategies cannot be synthesized incrementally: we cannot do better than guess the whole strategy and check if it succeeds.

Imperfect information makes model checking harder!

Model Checking Imperfect Information Games

Corollary

Imperfect information strategies cannot be synthesized incrementally: we cannot do better than guess the whole strategy and check if it succeeds.

Imperfect information makes model checking harder!
Or...?

Summary of Model Checking Results

	m, l	n, k, l	n_{local}, k, l
CTL			
ATL			
$\mathrm{ATL_{ir}}$/CSL			

Summary of Model Checking Results

	m, l	n, k, l	n_{local}, k, l
CTL	P		
ATL	P		
ATL_{ir}/CSL	Δ_2^P		

Summary of Model Checking Results

	m, l	n, k, l	n_{local}, k, l
CTL	P	P	
ATL	P	Δ_3^P	
ATL$_{ir}$/CSL	Δ_2^P	Δ_3^P	

Summary of Model Checking Results

	m, l	n, k, l	n_{local}, k, l
CTL	P	P	PSPACE
ATL	P	Δ_3^P	EXPTIME
ATL_{ir}/CSL	Δ_2^P	Δ_3^P	PSPACE

Summary of Model Checking Results

	m, l	n, k, l	n_{local}, k, l
CTL	P [1]	P [1]	PSPACE [2]
ATL	P [3]	Δ_3^P [5,8]	EXPTIME [6,7]
ATL_{ir}/CSL	Δ_2^P [4,9]	Δ_3^P [9]	PSPACE [7]

[1] Clarke, Emerson & Sistla (1986).
[2] Kupferman, Vardi & Wolper (2000).
[3] Alur, Henzinger & Kupferman (2002).
[4] Schobbens (2004).
[5] Jamroga & Dix (2005).
[6] Hoek, Lomuscio & Wooldridge (2006).
[7] Jamroga & Ågotnes (2007).
[8] Laroussinie, Markey & Oreiby (2007).
[9] Jamroga & Dix (2008).

The Message Again...

- Complexity is **very** sensitive to the context!
- In particular, **the way we define the input, and measure its size,** is crucial.

8.3 The Phantom Result

Between Perception and Recall

logic	ir	iR	Ir	IR
$\langle\langle\Gamma\rangle\rangle - ATL$	NP	U [11]	nl [2]	nl [2]
ATL	$\Delta_2 P$	U [11]	nl [2]	nl [2]
ATL^+	$\Delta_3 P$	U [11]	$\Delta_3 P$	$\Delta_3 P$
ATL^*	$PSPACE$	U [11]	$PSPACE$	$DEXP$ [9]

NP	complete for nondeterministic polynomial time
$\Delta_2 P = P^{NP}$	complete for polynomial calls to an NP oracle
$\Delta_3 P = P^{NP^{NP}}$	complete for polynomial calls to a $\Sigma_2 P$ oracle
EXP	complete for deterministic exponential time
DEXP	complete for deterministic doubly exponential time
U	undecidable
l	size of the formula
n	size of the model

The Undecidability “Result”

- Most cite it from (Alur et al., 1997–2002)
- Alur et al. cite Yannakakis ("Synchronous multi-player games with incomplete information are undecidable", 1997)
- **Personal communication!**
- No proof has been published (nor has the result been formally stated)


91 BUYL

Relevant Existing Results

- (Peterson & Reif, 1979):

Relevant Existing Results

- (Peterson & Reif, 1979):
- Solving games with imperfect information and perfect recall is decidable in the case of a single proponent
- Solving games with imperfect information and perfect recall is undecidable in the case of a team of proponents

Relevant Existing Results

- (Peterson & Reif, 1979):
- Solving games with imperfect information and perfect recall is decidable in the case of a single proponent
- Solving games with imperfect information and perfect recall is undecidable in the case of a team of proponents
- But: their games are defined via Turing machines, while in "our" games are close to finite automata

Relevant Existing Results

- (Pnueli & Rosner, 1990):

Relevant Existing Results

- (Pnueli & Rosner, 1990):
- Realizability problem for distributed systems is undecidable
- The setting very close to ours. Difference: “winning conditions” are defined via LTL specifications, so we have winning paths rather than states. In particular, the reduction of the halting problem for deterministic Turing Machines to the realizability problem (that proves undecidability of the problem) employs LTL formulae that are not expressible in CTL

Relevant Existing Results

- (Van der Meyden and Shilov, 1999):

Relevant Existing Results

- (Van der Meyden and Shilov, 1999):
- Model checking **LTL+K** with perfect recall is decidable (with a nonelementary lower bound)
- Model checking **LTL+K+C** with perfect recall is undecidable

Relevant Existing Results

- (Van der Meyden and Shilov, 1999):
- Model checking **LTL+K** with perfect recall is decidable (with a nonelementary lower bound)
- Model checking **LTL+K+C** with perfect recall is undecidable

- (Garanina, Kalinina and Shilov, 2004):

Relevant Existing Results

- (Van der Meyden and Shilov, 1999):
- Model checking **LTL+K** with perfect recall is decidable (with a nonelementary lower bound)
- Model checking **LTL+K+C** with perfect recall is undecidable

- (Garanina, Kalinina and Shilov, 2004):
- Model checking **CTL+K** with perfect recall is decidable (with a nonelementary lower bound)
- Model checking **CTL+K+C** with perfect recall is undecidable

8.4 References

[Jamroga and Dix 2008] W. Jamroga and J. Dix.
Model checking abilities of agents: A closer look.
Theory of Computing Systems, 42(3):366–410, 2008.

[Schobbens 2004] P. Y. Schobbens.
Alternating-time logic with imperfect recall.
Electronic Notes in Theoretical Computer Science, 85(2), 2004.

Axiom. of Coal. Games

9.1 Coalitional Games

Coalitional Games

The difference between non-cooperative games and coalitional games is that the former takes possible actions of individual players as primary, while the latter takes possible actions of coalitions as primary.

Non-coop. Game: Individual Actions Primary

		Bill	
		Cooperate	**Defect**
Ann	**Cooperate**	Ann:-1, Bill:-1	Ann:-4, Bill: 0
	Defect	Ann:0, Bill:-4	Ann:-3, Bill: -3

Non-coop. Game: Individual Actions Primary

		Bill	
		Cooperate	**Defect**
Ann	**Cooperate**	Ann:-1, Bill:-1	Ann:-4, Bill: 0
	Defect	Ann:0, Bill:-4	Ann:-3, Bill: -3

ATL/CL:

$$\langle\langle\{Ann\}\rangle\rangle B \geq 3$$

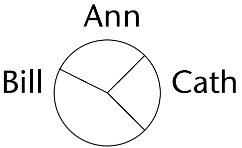
We can derive possible actions of coalitions, and thus coalitional power, from the individual actions:

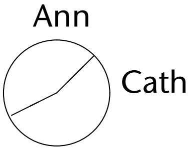
$$\langle\langle\{Ann, Bill\}\rangle\rangle(A = 1 \wedge B = 1)$$

Coalitional Game: Coalitional Actions Primary

Example: Three-player majority game:

- Three persons
- One cake
- Any majority group (two or three) controls the division of the cake to the members of the group
- Each person cares (only) about how much cake he gets





Definition 9.1 (Coalitional Game (with Transferable Payoff))

A coalitional game (with transferable payoff) is a tuple $(N, \Omega, V, \{\sqsupseteq_i\}_{i \in N})$:

- N is the set of players
- Ω is the set of outcomes
- V assigns a set of choices $V(C) \subseteq \Omega$ to each non-empty coalition $C \subseteq N$
- For each i, $\sqsupseteq_i$ is a a preference relation over the outcomes
  - Usually assumed to be reflexive, transitive and complete
  - We write $\sqsupset_i$ for the strict variant
  - Is often described by a utility function u_i for each player i over the outcomes: $\omega \sqsupseteq_i \omega'$ iff $u_i(\omega) \geq u_i(\omega')$

Coalitional Game: Example

The cake game:

- $N = \{Ann, Bill, Cath\}$
- Ω: the collection of possible ways to divide a cake between Ann and Bill, between Ann and Cath, between Bill and Cath, and between Ann and Bill and Cath
  - $\Omega = \{(A = 10\%, B = 90\%), (B = 50\%, C = 50\%), (A = 20\%, B = 30\%, C = 50\%), \ldots\}$
- $V(C) = \begin{cases} \text{all divisions of the cake among } C & |C| \geq 2 \\ \emptyset & \text{otherwise} \end{cases}$
- $\omega_1 \sqsupseteq_{Ann} \omega_2$ iff Ann gets at least as much cake in ω_1 as in ω_2, etc.

Coalitional Games with Transferable Payoff

Definition 9.2

Coalitional Game with Transferable Payoff A coalitional game with transferable payoff is a pair (N, v):

- N is the set of players
- v assigns a real number $v(C)$ to each non-empty coalition $C \subseteq N$; the worth of C

Coalitional Games with Transferable Payoff

Definition 9.2

Coalitional Game with Transferable Payoff A coalitional game with transferable payoff is a pair (N, v):

- N is the set of players
- v assigns a real number $v(C)$ to each non-empty coalition $C \subseteq N$; the worth of C

- Games WTP can be seen as a special class of games WOTP

Coalitional Games with Transferable Payoff

Definition 9.2

Coalitional Game with Transferable Payoff A coalitional game with transferable payoff is a pair (N, v):

- N is the set of players
- v assigns a real number $v(C)$ to each non-empty coalition $C \subseteq N$; the worth of C

- Games WTP can be seen as a special class of games WOTP
- But games WOTP are more general

Coalitional Games with Transferable Payoff

Definition 9.2

Coalitional Game with Transferable Payoff A coalitional game with transferable payoff is a pair (N, v):

- N is the set of players
- v assigns a real number $v(C)$ to each non-empty coalition $C \subseteq N$; the worth of C

- Games WTP can be seen as a special class of games WOTP
- But games WOTP are more general
- We will use games WOTP

Coalitional Games with Transferable Payoff

Definition 9.2

Coalitional Game with Transferable Payoff A coalitional game with transferable payoff is a pair (N, v):

- N is the set of players
- v assigns a real number $v(C)$ to each non-empty coalition $C \subseteq N$; the worth of C

- Games WTP can be seen as a special class of games WOTP
- But games WOTP are more general
- We will use games WOTP
- Henceforth: by "coalitional game" we mean "coalitional game WOTP".

Solution Concepts for Coalitional Games

- A solution concept assigns a set of outcomes to each game
- General idea: like in non-cooperative games: what are the stable outcomes?
- Stability: no coalition can profit from deviating

Solution Concepts for Coalitional Games

- A solution concept assigns a set of outcomes to each game
- General idea: like in non-cooperative games: what are the stable outcomes?
- Stability: no coalition can profit from deviating
- Some important concepts:
  - The core
  - Stable sets
  - The bargaining set

The Core

Definition 9.3 (The Core)

The core of a coalitional game is the set of outcomes $\omega \in V(N)$ for which there is no coalition C with an outcome $\omega' \in V(C)$ such that $\omega' \succ_i \omega$ for all $i \in C$.

The Core

Definition 9.3 (The Core)

The core of a coalitional game is the set of outcomes $\omega \in V(N)$ for which there is no coalition C with an outcome $\omega' \in V(C)$ such that $\omega' \succ_i \omega$ for all $i \in C$.

What is the core of the cake game?

The Core

Definition 9.3 (The Core)

The core of a coalitional game is the set of outcomes $\omega \in V(N)$ for which there is no coalition C with an outcome $\omega' \in V(C)$ such that $\omega' \succ_i \omega$ for all $i \in C$.

What is the core of the cake game?

Key property of a coalitional game: is the core empty?

Stable Sets

- Idea: an outcome is stable if no (sub)coalition has an incentive to deviate and form a stable coalition (recursive!)
- From von Neumann and Morgenstern, 1944
- A stable set is a set of outcomes
- A game may have more than one stable set
- .. but must not have any
- Characterised by imputations and objections

Imputation

Definition 9.4 (Imputation)

An imputation is an outcome $\omega \in V(N)$ that for each agent i is as least as good as any outcome the singleton coalition $\{i\}$ can choose on his own.

Objection

Definition 9.5 (Objection)

An imputation ω is a C-objection to an imputation ω' if every agent in C prefers ω over ω' and the coalition C can choose an outcome which for every agent in C is as least as good as ω. ω is an objection to ω' if ω is a C-objection to ω' for some coalition C.

Stable Set

Definition 9.6 (Stable Set)

A set of imputations Y is a stable set if it satisfies:

Internal stability If $\omega \in Y$, there is no objection to ω in Y.

External stability If $\omega \notin Y$, there is an objection to ω in Y.

The Bargaining Set

- A set of imputations
- Unique
- Always exists
- Defined in terms of objections and counterobjections – but the concept of objection is different from the stable sets case
- Will introduce it formally later

9.2 Coalitional Game Logic

Goal

- We want to be able to reason about coalitional games in a formal logic
- In particular: characterise solution concepts

Coalitional Game Logic

- We have already used the modality $\langle\langle C \rangle\rangle$ to reason about coalitional ability in non-cooperative games
- It is natural and straightforward to interpret this modality by the V function in coalitional games
- Additional assumptions on propositions in the language:
  - ω, where ω is (the name of) an outcome in Ω: meaning that the current outcome is ω
  - $\omega \succeq_i \omega'$: meaning that agent i weakly prefers outcome ω over ω'

Coalitional Game Logic

- We have already used the modality $\langle\langle C \rangle\rangle$ to reason about coalitional ability in non-cooperative games
- It is natural and straightforward to interpret this modality by the V function in coalitional games
- Additional assumptions on propositions in the language:
  - ω, where ω is (the name of) an outcome in Ω: meaning that the current outcome is ω
  - $\omega \succeq_i \omega'$: meaning that agent i weakly prefers outcome ω over ω'

Let Γ be a coalitional game.

$$
\begin{array}{lcl}
\Gamma \models \omega \succeq_i \omega' & \Leftrightarrow & \omega \sqsupseteq_i \omega' \\
\Gamma \models \langle C \rangle \phi & \Leftrightarrow & \exists \omega \in V(C), \omega \models \phi \\
\omega \models \omega' & \Leftrightarrow & \omega = \omega'
\end{array}
$$

The Core

The core of a coalitional game is the set of outcomes $\omega \in V(N)$ for which there is no coalition C with an outcome $\omega' \in V(C)$ such that $\omega' \succ_i \omega$ for all $i \in C$.

ω is a member of the core (assuming finite Ω):

$$CM(\omega) \equiv \langle N \rangle \omega \wedge \neg \left[\bigvee_{C \subseteq N} \bigvee_{\omega' \in \Omega} (\langle C \rangle \omega') \wedge \bigwedge_{i \in C} (\omega' \succ_i \omega) \right]$$

The Core

The core of a coalitional game is the set of outcomes $\omega \in V(N)$ for which there is no coalition C with an outcome $\omega' \in V(C)$ such that $\omega' \succ_i \omega$ for all $i \in C$.

ω is a member of the core (assuming finite Ω):

$$CM(\omega) \equiv \langle N \rangle \omega \wedge \neg \left[\bigvee_{C \subseteq N} \bigvee_{\omega' \in \Omega} (\langle C \rangle \omega') \wedge \bigwedge_{i \in C} (\omega' \succ_i \omega) \right]$$

The core is non-empty:

$$CNE \equiv \bigvee_{\omega \in \Omega} CM(\omega)$$

Imputation

An imputation is an outcome $\omega \in V(N)$ that for each agent i is as least as good as any outcome the singleton coalition $\{i\}$ can choose on his own.

$$IMP(\omega) \equiv \langle N \rangle \omega \wedge \bigwedge_{\omega' \in \Omega} \bigwedge_{i \in N} (\langle \{i\} \rangle \omega' \rightarrow \omega \succeq_i \omega')$$

Objection

An imputation ω is a C-objection to an imputation ω' if every agent in C prefers ω over ω' and the coalition C can choose an outcome which for every agent in C is as least as good as ω. ω is an objection to ω' if ω is a C-objection to ω' for some coalition C.

$$OBJ(\omega, \omega', C) \equiv (\bigwedge_{i \in C} \omega \succ_i \omega') \wedge \bigvee_{\omega'' \in \Omega} (\langle C \rangle \omega'' \wedge \bigwedge_{i \in C} \omega'' \succeq_i \omega)$$

Stable Set

A set of imputations Y is a stable set if it satisfies:

Internal stability If $\omega \in Y$, there is no objection to ω in Y.

External stability If $\omega \notin Y$, there is an objection to ω in Y.

$$
\begin{aligned}
&STABLE(Y) \equiv \\
&\quad \bigwedge_{\omega \in Y} IMP(\omega) \\
&\quad \wedge \left(\bigwedge_{\omega \in Y} \bigwedge_{C \subseteq N} \bigwedge_{\omega' \in Y} \neg OBJ(\omega', \omega, C)\right) \\
&\quad \wedge \left(\bigwedge_{\omega \in \Omega \setminus Y} IMP(\omega) \rightarrow \left(\bigvee_{C \subseteq N} \bigvee_{\omega' \in Y} OBJ(\omega', \omega, C)\right)\right)
\end{aligned}
$$

The Bargaining Set

ω' is an objection of C to ω:

$$OBJB(\omega', C, \omega) \equiv \langle C \rangle \omega' \wedge \bigwedge_{k \in C} \omega' \succ_k \omega$$

There exists a counterobjection to the objection ω' of C to ω, where $i \in C$ and $j \notin C$:

$$COUNTER(\omega', C, i, j, \omega) \equiv \bigvee_{v \in \Omega} \bigvee_{D' \subseteq N \setminus \{i\}} (\langle D' \cup \{j\} \rangle v \\ \wedge \left(\left(\bigwedge_{k \in (D' \cup \{j\}) \setminus C} v \succeq_k \omega \right) \wedge \left(\bigwedge_{k \in (D' \cup \{j\}) \cap C} v \succeq_k \omega' \right) \right))$$

The Bargaining Set

Outcome ω is in the bargaining set:

$$INBARG(\omega) \equiv IMP(\omega) \wedge \bigwedge_{C \subseteq N} \bigwedge_{i \in C} \bigwedge_{j \in N \setminus C} \bigwedge_{\omega' \in \Omega} [OBJB(\omega', C, \omega) \rightarrow COUNTER(\omega', C, i, j, \omega)]$$

$$BS(Y) = \bigwedge_{\omega \in Y} INBARG(\omega) \wedge \bigwedge_{\omega \in \Omega \setminus Y} \neg INBARG(\omega)$$

This Coalitional Game Logic:

- is very expressive (for finite games)
- can characterise solution concepts

This Coalitional Game Logic:

- is very expressive (for finite games)
- can characterise solution concepts

However, the characterisations:

- do not work for infinite games (games with infinitely many outcomes)
- are not very succinct
- depend on Ω and are thus different for games with different sets of outcomes

9.3 Modal Coalitional Game Logic

- Recall the def. of a coalitional game:

$$(N, \Omega, V, \{\sqsupseteq_i\}_{i \in N})$$

- Recall the def. of a coalitional game:

$$(N, \Omega, V, \{\sqsupseteq_i\}_{i \in N})$$

- In Coalitional Game Logic we used V to interpret $\langle C \rangle$, and atomic propositions for $\sqsupseteq_i$

- Recall the def. of a coalitional game:

$$(N, \Omega, V, \{\sqsupseteq_i\}_{i \in N})$$

- In Coalitional Game Logic we used V to interpret $\langle C \rangle$, and atomic propositions for $\sqsupseteq_i$
- Observation: there is "more structure" in $\sqsupseteq_i$!

- Recall the def. of a coalitional game:

$$(N, \Omega, V, \{\sqsupseteq_i\}_{i \in N})$$

- In Coalitional Game Logic we used V to interpret $\langle C \rangle$, and atomic propositions for $\sqsupseteq_i$
- Observation: there is "more structure" in $\sqsupseteq_i$!
- Modal Coalitional Game Logic: we will use $\sqsupseteq_i$ to interpret $\langle C \rangle$, and atomic propositions for V.

Modal Coalitional Game Logic (MCGL)

Main constructs ($C \subseteq N$):

$$\langle C \rangle \varphi$$

meaning: (all agents in) C prefers φ

Modal Coalitional Game Logic (MCGL)

Main constructs ($C \subseteq N$):

$$\langle C \rangle \varphi$$

meaning: (all agents in) C prefers φ
and

$$p_C$$

meaning: C can choose the current outcome

Modal Coalitional Game Logic (MCGL)

Main constructs ($C \subseteq N$):

$$\langle C \rangle \varphi$$

meaning: (all agents in) C prefers φ
and

$$p_C$$

meaning: C can choose the current outcome

Henceforth: use

$$\mathcal{C} = 2^N \setminus \emptyset$$

to denote the set of coalitions

Formal Language

Let

$$\Theta = \Theta' \cup \{p_C : C \subseteq N\}$$

where Θ' is a countably infinite set of atomic propositions

Formal Language

Let

$$\Theta = \Theta' \cup \{p_C : C \subseteq N\}$$

where Θ' is a countably infinite set of atomic propositions

The MCGL language (will add more later):

$$\phi ::= p \mid \langle C \rangle \phi \mid \langle C^s \rangle \phi \mid \neg \phi \mid \phi_1 \wedge \phi_2 \mid \cdots$$

where $p \in \Theta$, $C \subseteq N$, $i \in N$. Derived: $[\cdot], [\cdot^s]$ are the duals of $\langle \cdot \rangle, \langle \cdot^s \rangle$, respectively.

Interpretation

Let $\Gamma = (N, \Omega, V, \sqsupseteq_1, \ldots, \sqsupseteq_m)$ be a coalitional game, let π be a valuation of Θ' in Ω, and let $w \in \Omega$.

- $\Gamma, \pi, w \models p_C$ iff $w \in V(C)$
- $\Gamma, \pi, w \models p$ iff $w \in \pi(p)$, when $p \in \Theta'$
- $\Gamma, \pi, w \models \langle C \rangle \phi$ iff there is a v such that for every $i \in C$, $v \sqsupseteq_i w$, and $\Gamma, \pi, v \models \phi$

Interpretation

Let $\Gamma = (N, \Omega, V, \sqsupseteq_1, \ldots, \sqsupseteq_m)$ be a coalitional game, let π be a valuation of Θ' in Ω, and let $w \in \Omega$.

- $\Gamma, \pi, w \models p_C$ iff $w \in V(C)$
- $\Gamma, \pi, w \models p$ iff $w \in \pi(p)$, when $p \in \Theta'$
- $\Gamma, \pi, w \models \langle C \rangle \phi$ iff there is a v such that for every $i \in C$, $v \sqsupseteq_i w$, and $\Gamma, \pi, v \models \phi$
- $\Gamma, \pi, w \models \langle C^s \rangle \phi$ iff there is a v such that for every $i \in C$, $v \sqsupseteq_i w$ and not $w \sqsupseteq_i v$, and $\Gamma, \pi, v \models \phi$

Interpretation

Let $\Gamma = (N, \Omega, V, \sqsupseteq_1, \ldots, \sqsupseteq_m)$ be a coalitional game, let π be a valuation of Θ' in Ω, and let $w \in \Omega$.

- $\Gamma, \pi, w \models p_C$ iff $w \in V(C)$
- $\Gamma, \pi, w \models p$ iff $w \in \pi(p)$, when $p \in \Theta'$
- $\Gamma, \pi, w \models \langle C \rangle \phi$ iff there is a v such that for every $i \in C$, $v \sqsupseteq_i w$, and $\Gamma, \pi, v \models \phi$
- $\Gamma, \pi, w \models \langle C^s \rangle \phi$ iff there is a v such that for every $i \in C$, $v \sqsupseteq_i w$ and not $w \sqsupseteq_i v$, and $\Gamma, \pi, v \models \phi$

Let us write:

$$\begin{array}{ll} \Gamma, w \models \phi & \text{iff } \Gamma, \pi, w \models \phi \text{ for all } \pi \\ \Gamma \models \phi & \text{iff } \Gamma, w \models \phi \text{ for all } w \end{array}$$

Characterising the Core

The core of a coalitional game is the set of outcomes $\omega \in V(N)$ for which there is no coalition C with an outcome $\omega' \in V(C)$ such that $\omega' \succ_i \omega$ for all $i \in C$.

Characterising the Core

The core of a coalitional game is the set of outcomes $\omega \in V(N)$ for which there is no coalition C with an outcome $\omega' \in V(C)$ such that $\omega' \succ_i \omega$ for all $i \in C$.

$$MCM \equiv p_N \wedge \bigwedge_{C \subseteq N} [C^s]\neg p_C$$

Theorem 9.7

$\Gamma, \omega \models MCM$ *iff* ω *is in the core of* Γ

Characterising Imputations

An imputation is an outcome $\omega \in V(N)$ that for each agent i is as least as good as any outcome the singleton coalition $\{i\}$ can choose on his own.

Characterising Imputations

An imputation is an outcome $\omega \in V(N)$ that for each agent i is as least as good as any outcome the singleton coalition $\{i\}$ can choose on his own.

$$MIMP \equiv p_N \wedge \bigwedge_{i \in N} [C^s] \neg p_i$$

Theorem 9.8

$\Gamma, \omega \models MIMP$ *iff* ω *is an imputation in* Γ

Stable Sets

- Difficult to characterise stable sets and the bargaining set in MCGL
- How to refer to sets of outcomes? Formulae are interpreted in single outcomes, and we can't refer directly to outcomes in the formula (unlike in CGL).
- Here is a way: the extension

$$\phi^{\Gamma} = \{\omega : \Gamma, \omega \models \phi\}$$

  is a set.
- Example: $MCM\Gamma$ is the core of Γ

Stable Sets

Let

$$MOBJ(C,\alpha) \equiv MIMP \wedge \langle C^s\rangle(MIMP \wedge \alpha \wedge \langle C\rangle p_C)$$

meaning: $\Gamma,\omega \models MOBJ(C,\alpha)$ iff ω is an imputation and there exists a C-objection ω' to ω such that $\Gamma,\omega' \models \alpha$

Stable Sets

Let

$$MOBJ(C,\alpha) \equiv MIMP \wedge \langle C^s\rangle(MIMP \wedge \alpha \wedge \langle C\rangle p_C)$$

meaning: $\Gamma,\omega \models MOBJ(C,\alpha)$ iff ω is an imputation and there exists a C-objection ω' to ω such that $\Gamma,\omega' \models \alpha$

Theorem 9.9

Let γ be a formula.

$$\begin{aligned}\Gamma \models \quad & (\gamma \to MIMP) \wedge (\gamma \to \neg \bigvee_{C\subseteq N} MOBJ(C,\gamma)) \wedge \\ & (\neg\gamma \to \bigvee_{C\subseteq N} MOBJ(C,\gamma))\end{aligned}$$

iff γ^Γ is a stable set in Γ.

MCGL: advantages and disadvantages

- Note that we can characterise, e.g., the core also for infinite games
- The characterisation is the same for all games over the same set of agents
- But: not as expressive as CGL (for finite games)

9.4 Axiomatisation

Let us try to view this as a normal modal logic – and be very explicit. Hencforth assume a fixed set of agents N.

$$\phi ::= p \mid \langle C \rangle \phi \mid \langle C^s \rangle \phi \mid \neg \phi \mid \phi_1 \wedge \phi_2 \ \mid \cdots$$

A model would be a tuple :

$$M = (W, \{R_C : C \in \mathcal{C}\}, \{R^s_C : C \in \mathcal{C}\}, \pi)$$

where π is a valuation of $\Theta = \Theta' \cup \{p_C : C \in \mathcal{C}\}$

Let us try to view this as a normal modal logic – and be very explicit. Hencforth assume a fixed set of agents N.

$$\phi ::= p \mid \langle C \rangle \phi \mid \langle C^s \rangle \phi \mid \neg \phi \mid \phi_1 \wedge \phi_2 \ \mid \cdots$$

A model would be a tuple :

$$M = (W, \{R_C : C \in \mathcal{C}\}, \{R^s_C : C \in \mathcal{C}\}, \pi)$$

where π is a valuation of $\Theta = \Theta' \cup \{p_C : C \in \mathcal{C}\}$

And to get correspondence with coalitional games:

- REFL $\forall_{i \in N} R_i$ is reflexive
- TRANS $\forall_{i \in N} R_i$ is transitive
- COMPL $\forall_{i \in N} R_i$ is complete
- STRICT $\forall_{i \in N} R^s_i wu$ iff both $R_i wu$ and not $R_i uw$
- INTERSECTION $\forall_{C \in \mathcal{C}} R_C = \bigcap_{i \in C} R_i$
- INTERSECTION-STRICT $\forall_{C \in \mathcal{C}} R^s_C = \bigcap_{i \in C} R^s_i$

where we write R_i for $R_{\{i\}}$.

Axioms

REFL, TRANS:

T	$[i]p \rightarrow p$	
4	$[i]p \rightarrow [i][i]p$	

Axioms

REFL, TRANS:

T	$[i]p \rightarrow p$	
4	$[i]p \rightarrow [i][i]p$	

.. but several of the other properties do not have canonical formulae

Common approach when we have a property P, such as INTERSECTION, which are neither modally definable nor has a canonical formula:

- Construct the canonical model

Common approach when we have a property P, such as INTERSECTION, which are neither modally definable nor has a canonical formula:

- Construct the canonical model
- Transform it into a model which satisfies the same formulae, but which has the property P

Common approach when we have a property P, such as INTERSECTION, which are neither modally definable nor has a canonical formula:

- Construct the canonical model
- Transform it into a model which satisfies the same formulae, but which has the property P
- However: transformation may be difficult when there are several properties that must be achieved/maintained at the same time – as in our case

Common approach when we have a property P, such as INTERSECTION, which are neither modally definable nor has a canonical formula:

- Construct the canonical model
- Transform it into a model which satisfies the same formulae, but which has the property P
- However: transformation may be difficult when there are several properties that must be achieved/maintained at the same time – as in our case
- Many alternative approaches have been used and studied in detail

Common approach when we have a property P, such as INTERSECTION, which are neither modally definable nor has a canonical formula:

- Construct the canonical model
- Transform it into a model which satisfies the same formulae, but which has the property P
- However: transformation may be difficult when there are several properties that must be achieved/maintained at the same time – as in our case
- Many alternative approaches have been used and studied in detail
- Here: we will use standard techniques combining the difference modality with a step-by-step method using converse modalities
  - See Modal Logic by Blackburn et al.; more references at the end

The Difference Modality

The difference modality is a diamond $\langle D \rangle$, where $\langle D \rangle \phi$ means that ϕ is true somewhere else.

The Difference Modality

The difference modality is a diamond $\langle D \rangle$, where $\langle D \rangle \phi$ means that ϕ is true somewhere else.

It has a fixed interpretation in a model M:

$$M, w \models \langle D \rangle \phi \Leftrightarrow \exists v \neq w \; M, v \models \phi$$

The Difference Modality

The difference modality is a diamond $\langle D \rangle$, where $\langle D \rangle \phi$ means that ϕ is true somewhere else.

It has a fixed interpretation in a model M:

$$M, w \models \langle D \rangle \phi \Leftrightarrow \exists v \neq w \ M, v \models \phi$$

- We add the difference modality to the language
- Not only to be able to axiomatise the logic
- .. but also because it is useful for reasoning about games. E.g. the core is not empty:

$$MCNE \equiv MCM \vee \langle D \rangle MCM$$

Converse Modalities

Let $\langle i^c \rangle$ denote the converse of the diamond $\langle i \rangle$:

- $\langle i \rangle \phi$: there is an outcome which is preferred by i over the current one, in which ϕ is true
- $\langle i^c \rangle \phi$: there is an outcome over which the current outcome is preferred by i, in which ϕ is true

Converse Modalities

Let $\langle i^c \rangle$ denote the converse of the diamond $\langle i \rangle$:

- $\langle i \rangle \phi$: there is an outcome which is preferred by i over the current one, in which ϕ is true
- $\langle i^c \rangle \phi$: there is an outcome over which the current outcome is preferred by i, in which ϕ is true
- We include converses for all the diamonds
- Converses make the step-by-step model construction technique we are going to use possible
- Also useful for reasoning about games

MCGL: Full language and explicit models

$$\phi ::= \; p \mid \langle C\rangle\phi \mid \langle C^s\rangle\phi \mid \langle D\rangle\phi \mid \langle C^c\rangle\phi \mid \langle C^{sc}\rangle\phi \mid \neg\phi \mid \phi_1 \wedge \phi_2$$

MCGL: Full language and explicit models

$$\phi ::= \ p \mid \langle C\rangle\phi \mid \langle C^s\rangle\phi \mid \langle D\rangle\phi \mid \langle C^c\rangle\phi \mid \langle C^{sc}\rangle\phi \mid \neg\phi \mid \phi_1 \wedge \phi_2$$

$$M = (W, \ \{R_C : C \in \mathcal{C}\}, \{R_C^s : C \in \mathcal{C}\}, D, \{R_C^c : C \in \mathcal{C}\}, \{R_C^{sc} : C \in \mathcal{C}\}, \pi)$$

REFL $\forall_{i\in N} R_i$ is reflexive

TRANS $\forall_{i\in N} R_i$ is transitive

COMPL $\forall_{i\in N} R_i$ is complete

STRICT $\forall_{i\in N} R_i^s wu$ iff both $R_i wu$ and not $R_i uw$

DIFF $D = \{(w,u) : w \neq u\}$

INTERSECTION $\forall_{C\in\mathcal{C}} R_C = \bigcap_{i\in C} R_i$

INTERSECTION-STRICT $\forall_{C\in\mathcal{C}} R_C^s = \bigcap_{i\in C} R_i^s$

CONVERSE Rwv iff $R^c vw$, for $R \in \{R_i, R_i^s, R_C, R_C^s, D\}$

Explicit models vs. games

Recall the interpretation in coalitional games:

$$\Gamma, \pi, \omega \models \phi$$

- Explicit models are just another representation of (Γ, π) pairs
- An axiomatisation of models will be an axiomatisation of games as well

Axioms: overview

- Normality: $ModusPonens$, $Usub$, $Prop$, as well as K and Nec for all the boxes
- $T, 4$ for individual preference relations
- Axioms and rules for the difference modality
- Axioms for completeness of individual preferences
- Converse axioms
- Strictness axioms
- Intersection axioms

(see the paper p. 34 for a summary)

Axioms: difference modality

D_1	$p \to [D]\langle D\rangle p$	*symmetry*
D_2	$\Diamond_1 \cdots \Diamond_k p \to (p \vee \langle D\rangle p)$	$\Diamond_i \in Diamonds$
$D - rule$	$\vdash (p \wedge \neg\langle D\rangle p) \to \theta \Rightarrow \vdash \theta$	p not in θ

Relatively standard, see Blackburn et al., Modal Logic.

Axioms: completeness/totality

Trichotomy	$(p \wedge [i]q) \rightarrow [D](q \vee p \vee \langle i \rangle p)$	

Axioms: converses

$Converse_1(\xi)$	$p \rightarrow [\xi]\langle\xi^c\rangle p$	$\xi \in \Xi$
$Converse_2(\xi)$	$p \rightarrow [\xi^c]\langle\xi\rangle p$	$\xi \in \Xi$

where:

$$\Xi = \{C, C^s, C^c, C^{sc}\}$$

Axioms: strictness

$Strict_1$	$p \wedge \langle i \rangle (q \wedge [i]\neg p) \rightarrow \langle i^s \rangle q$
$Strict_2$	$(p \wedge [D]\neg p \wedge \langle i^s \rangle q) \rightarrow \langle i \rangle (q \wedge \neg \langle i \rangle p)$
$Strict_3$	$\langle i^s \rangle p \rightarrow \langle D \rangle p$

Axioms: intersection

$Intersect_1$	$((p \wedge [D]\neg p) \vee \langle D\rangle(p \wedge [D]\neg p)) \to (\bigwedge_{i \in C}\langle i\rangle p \to \langle C\rangle p)$	
$Intersect_2$	$((p \wedge [D]\neg p) \vee \langle D\rangle(p \wedge [D]\neg p)) \to (\bigwedge_{i \in C}\langle i^s\rangle p \to \langle C^s\rangle p)$	
$Intersect_3$	$\langle C\rangle p \to \langle i\rangle p$	$i \in C$
$Intersect_4$	$\langle C^s\rangle p \to \langle i^s\rangle p$	$i \in C$

9.5 Completeness

Outline

We will use a step-by-step method:

Outline

We will use a step-by-step method:

- The result will be a submodel of the canonical model

Outline

We will use a step-by-step method:

- The result will be a submodel of the canonical model
- We will build a network, which has much of the information needed for a proper model

Outline

We will use a step-by-step method:

- The result will be a submodel of the canonical model
- We will build a network, which has much of the information needed for a proper model
- We build it step-by-step, by starting with a small network and repairing its defects by extending it

Outline

We will use a step-by-step method:

- The result will be a submodel of the canonical model
- We will build a network, which has much of the information needed for a proper model
- We build it step-by-step, by starting with a small network and repairing its defects by extending it
- The converse operators make it possible to go back and forth, and to describe a finite network using formulae

Outline

We will use a step-by-step method:

- The result will be a submodel of the canonical model
- We will build a network, which has much of the information needed for a proper model
- We build it step-by-step, by starting with a small network and repairing its defects by extending it
- The converse operators make it possible to go back and forth, and to describe a finite network using formulae
- The D operator lets us
  - Define the needed model properties
  - Construct a named model by requiring that a formula of the form $p \wedge \neg\langle D\rangle p$ holds in a state

Outline

We will use a step-by-step method:

- The result will be a submodel of the canonical model
- We will build a network, which has much of the information needed for a proper model
- We build it step-by-step, by starting with a small network and repairing its defects by extending it
- The converse operators make it possible to go back and forth, and to describe a finite network using formulae
- The D operator lets us
  - Define the needed model properties
  - Construct a named model by requiring that a formula of the form $p \wedge \neg\langle D\rangle p$ holds in a state

Definition 9.10 (Network)

A network is a tuple

$$\mathcal{N} = (N, E, d, r, \Lambda)$$

- (N, E) is a finite, undirected, connected and acyclic graph
- d maps each edge $\{s, t\} \in E$ to a relation in the set $\{R_C, R_C^s, D : C \in \mathcal{C}\}$
- r maps each edge $\{s, t\} \in E$ to either s or t
- Λ labels each node in N with a finite set of formulae

We can describe a network with formulae:
Let $E(s)$ denote the set of nodes adjacent to s, and let

$$\langle st\rangle = \begin{cases} \langle i\rangle & d(\{s,t\}) = R_i \text{ and } r(\{s,t\}) = s \\ \langle i^c\rangle & d(\{s,t\}) = R_i \text{ and } r(\{s,t\}) = t \\ \langle i^s\rangle & d(\{s,t\}) = R_i^s \text{ and } r(\{s,t\}) = s \\ \langle i^{sc}\rangle & d(\{s,t\}) = R_i^s \text{ and } r(\{s,t\}) = t \\ \langle C\rangle & d(\{s,t\}) = R_C \text{ and } r(\{s,t\}) = s \\ \langle C^c\rangle & d(\{s,t\}) = R_C \text{ and } r(\{s,t\}) = t \\ \langle C^s\rangle & d(\{s,t\}) = R_C^s \text{ and } r(\{s,t\}) = s \\ \langle C^{sc}\rangle & d(\{s,t\}) = R_C^s \text{ and } r(\{s,t\}) = t \\ \langle D\rangle & d(\{s,t\}) = D \text{ and } r(\{s,t\}) = s \\ \langle D\rangle & d(\{s,t\}) = D \text{ and } r(\{s,t\}) = t \end{cases}$$

$$\Delta(\mathcal{N}, s) = \bigwedge \Lambda(s) \wedge \bigwedge_{v\in E(s)} \langle sv\rangle \Phi(\mathcal{N}, v, s)$$
$$\Phi(\mathcal{N}, t, s) = \bigwedge \Lambda(t) \wedge \bigwedge_{s\neq v\in E(t)} \langle tv\rangle \Phi(\mathcal{N}, v, t)$$

We can describe a network with formulae:
Let $E(s)$ denote the set of nodes adjacent to s, and let

$$\langle st \rangle = \begin{cases} \langle i \rangle & d(\{s,t\}) = R_i \text{ and } r(\{s,t\}) = s \\ \langle i^c \rangle & d(\{s,t\}) = R_i \text{ and } r(\{s,t\}) = t \\ \langle i^s \rangle & d(\{s,t\}) = R_i^s \text{ and } r(\{s,t\}) = s \\ \langle i^{sc} \rangle & d(\{s,t\}) = R_i^s \text{ and } r(\{s,t\}) = t \\ \langle C \rangle & d(\{s,t\}) = R_C \text{ and } r(\{s,t\}) = s \\ \langle C^c \rangle & d(\{s,t\}) = R_C \text{ and } r(\{s,t\}) = t \\ \langle C^s \rangle & d(\{s,t\}) = R_C^s \text{ and } r(\{s,t\}) = s \\ \langle C^{sc} \rangle & d(\{s,t\}) = R_C^s \text{ and } r(\{s,t\}) = t \\ \langle D \rangle & d(\{s,t\}) = D \text{ and } r(\{s,t\}) = s \\ \langle D \rangle & d(\{s,t\}) = D \text{ and } r(\{s,t\}) = t \end{cases}$$

$$\Delta(\mathcal{N}, s) = \bigwedge \Lambda(s) \wedge \bigwedge_{v \in E(s)} \langle sv \rangle \Phi(\mathcal{N}, v, s)$$
$$\Phi(\mathcal{N}, t, s) = \bigwedge \Lambda(t) \wedge \bigwedge_{s \neq v \in E(t)} \langle tv \rangle \Phi(\mathcal{N}, v, t)$$

Note the role of converses for all the diamonds here!

We can show the following:

Proposition

$\Delta(\mathcal{N}, s)$ is consistent iff $\Delta(\mathcal{N}, t)$ is consistent, for any two nodes in any network $\mathcal{N}$

by using the $Converse_1$ and $Converse_2$ axioms (when an edge is marked with anything else than a D) and the D_2 axiom (when an edge is marked with D).

We can show the following:

Proposition

$\Delta(\mathcal{N}, s)$ is consistent iff $\Delta(\mathcal{N}, t)$ is consistent, for any two nodes in any network $\mathcal{N}$

by using the $Converse_1$ and $Converse_2$ axioms (when an edge is marked with anything else than a D) and the D_2 axiom (when an edge is marked with D).

Definition 9.11 (Coherence)

A network is coherent if $\Delta(\mathcal{N}, s)$ is consistent for any s.

Possible defects in a network:

$D1(s, \phi)$ where s is a node and ϕ a formula, and $\phi \notin \Lambda(s)$ and $\neg\phi \notin \Lambda(s)$

$D2(s)$ there is no formula ϕ such that $\phi \wedge \neg\langle D\rangle\phi \in \Lambda(s)$

$D3(s, \langle\xi\rangle\phi)$ ($\xi \in \{i, C, i^s, C^s, D\}$) where s is a node and $\langle\xi\rangle\phi \in \Lambda(s)$ and for all $(s,t) \in E$ such that $d(\{s,t\}) = \mathrm{Rel}(\xi)$ and $r(\{s,t\}) = s$ it is the case that $\phi \notin \Lambda(t)$

$D4(s, \langle\xi^c\rangle\phi)$ ($\xi \in \{i, C, i^s, C^s\}$) where s is a node and $\langle\xi^c\rangle\phi \in \Lambda(s)$ and for all $(s,t) \in E$ such that $d(\{s,t\}) = \mathrm{Rel}(\xi)$ and $r(\{s,t\}) = t$ it is the case that $\phi \notin \Lambda(t)$

Proposition

For any defect in a coherent network $\mathcal{N}$, there is a coherent network $\mathcal{N}'$ extending $\mathcal{N}$ lacking that effect.

Repairing defects: standard approach

Repairing D2-defects with the D-rule

$D2(s)$ there is no formula ϕ such that $\phi \wedge \neg\langle D\rangle\phi \in \Lambda(s)$

- Let p be an atom not occurring in $\Delta(\mathcal{N}, s)$ (recall that we assumed there are infinitely many)
- Alternative statement of the D-rule:

$$\text{If } \Phi \text{ is consistent and does not contain } p$$
$$\Downarrow$$
$$(p \wedge \neg\langle D\rangle p) \wedge \Phi \text{ is consistent}$$

- $\Delta(\mathcal{N}, s)$ consistent so $\Delta(\mathcal{N}, s) \wedge p \wedge \neg\langle D\rangle p$ is consistent
- Define the new network by adding $p \wedge \neg\langle D\rangle p$ to $\Lambda(s)$
- Clearly, it is coherent

Repairing D3- and D4-defects

$D3(s, \langle\xi\rangle\phi)$ ($\xi \in \{i, C, i^s, C^s, D\}$) where s is a node and $\langle\xi\rangle\phi \in \Lambda(s)$ and for all $(s,t) \in E$ such that $d(\{s,t\}) = \mathrm{Rel}(\xi)$ and $r(\{s,t\}) = s$ it is the case that $\phi \notin \Lambda(t)$

Repairing D3- and D4-defects

$D3(s, \langle\xi\rangle\phi)$ ($\xi \in \{i, C, i^s, C^s, D\}$) where s is a node and $\langle\xi\rangle\phi \in \Lambda(s)$ and for all $(s,t) \in E$ such that $d(\{s,t\}) =$ Rel(ξ) and $r(\{s,t\}) = s$ it is the case that $\phi \notin \Lambda(t)$

We define $\mathcal{N}'$ as follows:

- $N' = N \cup \{t\}$ for some $t \in Y \setminus N$
- $E' = E \cup \{\{s,t\}\}$
- $d' = d \cup \{\{s,t\} \mapsto \text{Rel}(\xi)\}$
- $r' = r \cup \{\{s,t\} \mapsto s\}$
- $\Lambda' = \Lambda \cup \{t \mapsto \{\phi\}\}$

where Y is a countably infinite set of “fresh” states.

Repairing D3- and D4-defects

$D3(s, \langle\xi\rangle\phi)$ ($\xi \in \{i, C, i^s, C^s, D\}$) where s is a node and $\langle\xi\rangle\phi \in \Lambda(s)$ and for all $(s,t) \in E$ such that $d(\{s,t\}) = \mathrm{Rel}(\xi)$ and $r(\{s,t\}) = s$ it is the case that $\phi \notin \Lambda(t)$

We define $\mathcal{N}'$ as follows:

- $N' = N \cup \{t\}$ for some $t \in Y \setminus N$
- $E' = E \cup \{\{s,t\}\}$
- $d' = d \cup \{\{s,t\} \mapsto \mathrm{Rel}(\xi)\}$
- $r' = r \cup \{\{s,t\} \mapsto s\}$
- $\Lambda' = \Lambda \cup \{t \mapsto \{\phi\}\}$

where Y is a countably infinite set of “fresh” states.

We have that $\Delta(\mathcal{N}', s) = \Delta(\mathcal{N}, s) \wedge \langle\xi\rangle\phi$. Since $\langle\xi\rangle\phi \in \Lambda(s)$, it already is a conjunct of $\Delta(\mathcal{N}, s)$. Thus, $\mathcal{N}'$ is coherent.

Fix a consistent formula

$$\hat{\phi}$$

We now will construct a model for it.

$\mathcal{N}_i$

For every number i define a network
$\mathcal{N}_i = (N_i, E_i, d_i, r_i, \Lambda_i)$:

- $\mathcal{N}_0$ has a single node y labelled with $\{\hat{\phi}\}$. Clearly, $\mathcal{N}_0$ is coherent.
- When $n > 0$, $\mathcal{N}_{n+1}$ is the (coherent) network obtained by repairing the next (according to some enumeration) defect, by the rules given above.

$\mathcal{N}_i$

For every number i define a network
$\mathcal{N}_i = (N_i, E_i, d_i, r_i, \Lambda_i)$:

- $\mathcal{N}_0$ has a single node y labelled with $\{\hat{\phi}\}$. Clearly, $\mathcal{N}_0$ is coherent.
- When $n > 0$, $\mathcal{N}_{n+1}$ is the (coherent) network obtained by repairing the next (according to some enumeration) defect, by the rules given above.

Note that:

- $\mathcal{N}_j$ extends $\mathcal{N}_i$ when $i < j$
- a repaired defect will never can never be reintroduced
- for any defect of $\mathcal{N}_i$ there is a $j > i$ such that $\mathcal{N}_j$ lacks that defect

Collect repairs: $\mathcal{N}$

Collect all repairs by defining $\mathcal{N} = (N, E, d, r, \Lambda)$:

- $N = \bigcup_{i \in \mathbb{N}} Ni$
- $E = \bigcup_{i \in \mathbb{N}} E_i$
- $d = \bigcup_{i \in \mathbb{N}} d_i$
- $r = \bigcup_{i \in \mathbb{N}} r_i$
- $\Lambda(s) = \cup\{\Lambda_i(s) : i \in \mathbb{N}, s \in N_i\}$

Collect repairs: $\mathcal{N}$

Collect all repairs by defining $\mathcal{N} = (N, E, d, r, \Lambda)$:

- $N = \bigcup_{i \in \mathbb{N}} Ni$
- $E = \bigcup_{i \in \mathbb{N}} E_i$
- $d = \bigcup_{i \in \mathbb{N}} d_i$
- $r = \bigcup_{i \in \mathbb{N}} r_i$
- $\Lambda(s) = \cup \{\Lambda_i(s) : i \in \mathbb{N}, s \in N_i\}$

Proposition

For every s, $\Lambda(s)$ is a maximal consistent set of formulae.

- Maximality: repair of $D1$ effects
- Consistency: from consistency of each $\mathcal{N}_i$

The Model

Define the model M (for $\hat{\phi}$) by restricting the canonical model for MCGL to the MCSs that appear in $\mathcal{N}$, i.e. to

$$W = \{\Lambda(s) : s \in N\}$$

The Model

Define the model M (for $\hat{\phi}$) by restricting the canonical model for MCGL to the MCSs that appear in $\mathcal{N}$, i.e. to

$$W = \{\Lambda(s) : s \in N\}$$

(remove the other states, restrict the relations and valuation function accordingly)

Truth Lemma

Proposition

$$M, \Gamma \models \psi \Leftrightarrow \psi \in \Gamma$$

for any $\Gamma \in W$ and any ψ

Truth Lemma

Proposition

$$M, \Gamma \models \psi \Leftrightarrow \psi \in \Gamma$$

for any $\Gamma \in W$ and any ψ

- First: show that we don't throw away too much, that for any diamond $\diamond$ we have that $\Gamma \in W$ whenever $\diamond\psi \in \Gamma$, there is a $\Delta \in W$ such that $\psi \in Delta$ and Γ, Δ are related by the canonical relation. Easily shown by construction.
- Then: induction on ϕ

It remains to be shown that M has all the properties we required.

It remains to be shown that M has all the properties we required.

- **REFL** $\forall_{i \in N} R_i$ is reflexive
- **TRANS** $\forall_{i \in N} R_i$ is transitive
- **COMPL** $\forall_{i \in N} R_i$ is complete
- **STRICT** $\forall_{i \in N} R_i^s wu$ iff both $R_i wu$ and not $R_i uw$
- **DIFF** $D = \{(w,u) : w \neq u\}$
- **INTERSECTION** $\forall_{C \in \mathcal{C}} R_C = \bigcap_{i \in C} R_i$
- **INTERSECTION-STRICT** $\forall_{C \in \mathcal{C}} R_C^s = \bigcap_{i \in C} R_i^s$
- **CONVERSE** Rwv iff $R^c vw$, for $R \in \{R_i, R_i^s, R_C, R_C^s, D\}$

M is named

Because we removed $D2$-defects, for every state w there exists a formula ϕ_w such that

$$M, w \models \phi_w \wedge \neg\langle D\rangle\phi_w$$

M is named

Because we removed $D2$-defects, for every state w there exists a formula ϕ_w such that

$$M, w \models \phi_w \wedge \neg\langle D\rangle\phi_w$$

But from DIFF it follows that ϕ_w is uniquely true at w, that

$$M, w \models \phi_w$$

and for any $u \neq w$

$$M, u \not\models \phi_w$$

COMPL

- Assume that $w \neq u$, $\neg R_i uw$ and $\neg R_i wu$

COMPL

- Assume that $w \neq u$, $\neg R_i uw$ and $\neg R_i wu$
- We have that $M, u \not\models \langle i^c \rangle \phi_w$: otherwise
  - there is a v s.t. $R_i^c uv$ and $M, v \models \phi_w$
  - thus $v = w$, and $R_i^c uw$
  - by CONVERSE: $R_i wu$ – a contradiction

COMPL

- Assume that $w \neq u$, $\neg R_i uw$ and $\neg R_i wu$
- We have that $M, u \not\models \langle i^c \rangle \phi_w$: otherwise
  - there is a v s.t. $R_i^c uv$ and $M, v \models \phi_w$
  - thus $v = w$, and $R_i^c uw$
  - by CONVERSE: $R_i wu$ – a contradiction
- Thus $M, u \not\models (\phi_w \vee \langle i^c \rangle \phi_w \vee \langle i \rangle \phi_w)$

COMPL

- Assume that $w \neq u$, $\neg R_i uw$ and $\neg R_i wu$
- We have that $M, u \not\models \langle i^c \rangle \phi_w$: otherwise
  - there is a v s.t. $R_i^c uv$ and $M, v \models \phi_w$
  - thus $v = w$, and $R_i^c uw$
  - by CONVERSE: $R_i wu$ – a contradiction
- Thus $M, u \not\models (\phi_w \vee \langle i^c \rangle \phi_w \vee \langle i \rangle \phi_w)$
- And since $w \neq u$, $M, w \not\models [D]((\phi_w \vee \langle i^c \rangle \phi_w \vee \langle i \rangle \phi_w)$

COMPL

- Assume that $w \neq u$, $\neg R_i uw$ and $\neg R_i wu$
- We have that $M, u \not\models \langle i^c \rangle \phi_w$: otherwise
  - there is a v s.t. $R_i^c uv$ and $M, v \models \phi_w$
  - thus $v = w$, and $R_i^c uw$
  - by CONVERSE: $R_i wu$ – a contradiction
- Thus $M, u \not\models (\phi_w \vee \langle i^c \rangle \phi_w \vee \langle i \rangle \phi_w)$
- And since $w \neq u$, $M, w \not\models [D]((\phi_w \vee \langle i^c \rangle \phi_w \vee \langle i \rangle \phi_w)$
- Also: $M, w \models (\phi_w \wedge [i]\langle i^c \rangle \phi_w)$

COMPL

- Assume that $w \neq u$, $\neg R_i uw$ and $\neg R_i wu$
- We have that $M, u \not\models \langle i^c \rangle \phi_w$: otherwise
  - there is a v s.t. $R_i^c uv$ and $M, v \models \phi_w$
  - thus $v = w$, and $R_i^c uw$
  - by CONVERSE: $R_i wu$ – a contradiction
- Thus $M, u \not\models (\phi_w \vee \langle i^c \rangle \phi_w \vee \langle i \rangle \phi_w)$
- And since $w \neq u$, $M, w \not\models [D]((\phi_w \vee \langle i^c \rangle \phi_w \vee \langle i \rangle \phi_w)$
- Also: $M, w \models (\phi_w \wedge [i]\langle i^c \rangle \phi_w)$
- Contradicts the Trichotomy axiom

9.6 References

[1] M. J. Osborne and A. Rubinstein.
A Course in Game Theory.
The MIT Press: Cambridge, MA, 1994.

[2] Thomas Ågotnes, Wiebe van der Hoek, and Michael Wooldridge.
Reasoning about coalitional games.
To appear in *Artificial Intelligence*, 2008.

[3] M. de Rijke.
Extended Modal Logic.
PhD thesis, University of Amsterdam, 1993.

[4] P. Blackburn, M. de Rijke, and V. Venema.
Modal Logic.
Cambridge Univ. Press, 2001.

Thank you
for your attention!