Reasoning about Knowledge and Beliefs

Lecture 21

Eric Pacuit

University of Maryland, College Park

pacuit.org
epacuit@umd.edu

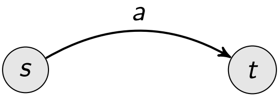
December 9, 2013

Actions

1. Actions as *transitions between states, or situations*:

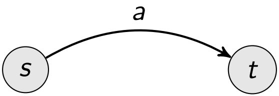
Actions

1. Actions as *transitions between states, or situations*:

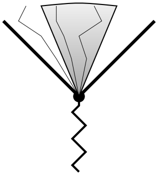


Actions

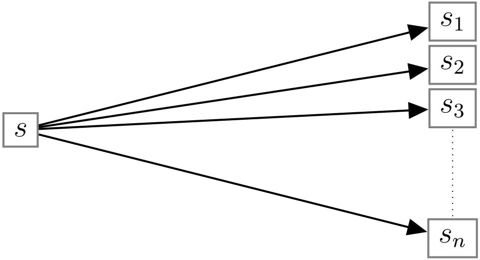
1. Actions as *transitions between states, or situations*:



2. Actions *restrict* the set of possible future histories.



J. van Benthem, H. van Ditmarsch, J. van Eijck and J. Jaspers. *Chapter 6: Propositional Dynamic Logic*. Logic in Action Online Course Project, 2011.


s
s_1
s_2
s_3
s_n

Propositional Dynamic Logic

Language: The language of propositional dynamic logic is generated by the following grammar:

$$p \mid \neg\varphi \mid \varphi \wedge \psi \mid [\alpha]\varphi$$

where $p \in \mathrm{At}$ and α is generated by the following grammar:

$$a \mid \alpha \cup \beta \mid \alpha;\beta \mid \alpha^* \mid \varphi?$$

where $a \in \mathrm{Act}$ and φ is a formula.

Propositional Dynamic Logic

Language: The language of propositional dynamic logic is generated by the following grammar:

$$p \mid \neg\varphi \mid \varphi \wedge \psi \mid [\alpha]\varphi$$

where $p \in \mathrm{At}$ and α is generated by the following grammar:

$$a \mid \alpha \cup \beta \mid \alpha;\beta \mid \alpha^* \mid \varphi?$$

where $a \in \mathrm{Act}$ and φ is a formula.

Semantics: $\mathcal{M} = \langle W, \{R_a \mid a \in \mathrm{P}\}, V\rangle$ where for each $a \in \mathrm{P}$, $R_a \subseteq W \times W$ and $V : \mathrm{At} \to \wp(W)$

Propositional Dynamic Logic

Language: The language of propositional dynamic logic is generated by the following grammar:

$$p \mid \neg\varphi \mid \varphi \wedge \psi \mid [\alpha]\varphi$$

where $p \in \mathrm{At}$ and α is generated by the following grammar:

$$a \mid \alpha \cup \beta \mid \alpha;\beta \mid \alpha^* \mid \varphi?$$

where $a \in \mathrm{Act}$ and φ is a formula.

Semantics: $\mathcal{M} = \langle W, \{R_a \mid a \in \mathrm{P}\}, V\rangle$ where for each $a \in \mathrm{P}$, $R_a \subseteq W \times W$ and $V : \mathrm{At} \to \wp(W)$

$[\alpha]\varphi$ means "after doing α, φ will be true"

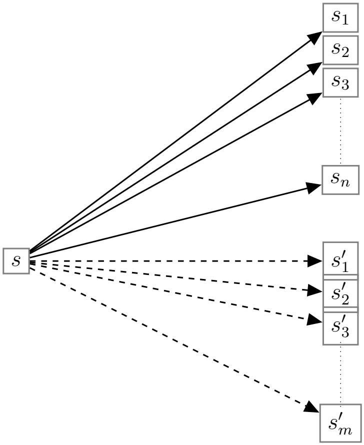
$\langle\alpha\rangle\varphi$ means "after doing α, φ may be true"

$\mathcal{M}, w \models [\alpha]\varphi$ iff for each v, if $wR_\alpha v$ then $\mathcal{M}, v \models \varphi$

$\mathcal{M}, w \models \langle\alpha\rangle\varphi$ iff there is a v such that $wR_\alpha v$ and $\mathcal{M}, v \models \varphi$

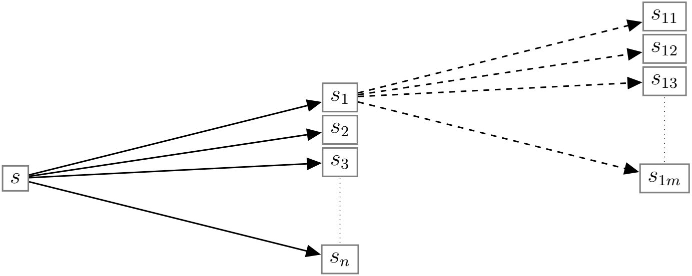
Union

$R_{\alpha\cup\beta} := R_\alpha \cup R_\beta$



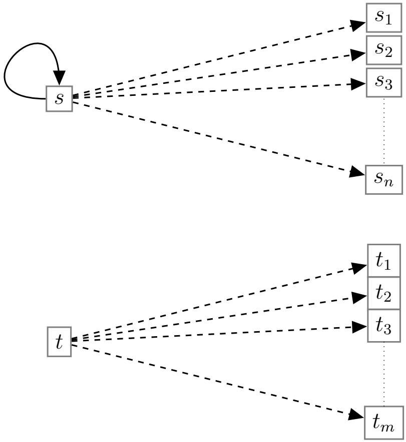
Sequence

$$R_{\alpha;\beta} := R_\alpha \circ R_\beta$$



Test

$$R_{\varphi?} = \{(w, w) \mid \mathcal{M}, w \models \varphi\}$$



Iteration

$$R_{\alpha^*} := \cup_{n \geq 0} R_\alpha^n$$

Propositional Dynamic Logic

1. Axioms of propositional logic
2. $[\alpha](\varphi \to \psi) \to ([\alpha]\varphi \to [\alpha]\psi)$
3. $[\alpha \cup \beta]\varphi \leftrightarrow [\alpha]\varphi \wedge [\beta]\varphi$
4. $[\alpha;\beta]\varphi \leftrightarrow [\alpha][\beta]\varphi$
5. $[\psi?]\varphi \leftrightarrow (\psi \to \varphi)$
6. $\varphi \wedge [\alpha][\alpha^*]\varphi \leftrightarrow [\alpha^*]\varphi$
7. $\varphi \wedge [\alpha^*](\varphi \to [\alpha]\varphi) \to [\alpha^*]\varphi$
8. Modus Ponens and Necessitation (for each program α)

Propositional Dynamic Logic

1. Axioms of propositional logic
2. $[\alpha](\varphi \to \psi) \to ([\alpha]\varphi \to [\alpha]\psi)$
3. $[\alpha \cup \beta]\varphi \leftrightarrow [\alpha]\varphi \wedge [\beta]\varphi$
4. $[\alpha; \beta]\varphi \leftrightarrow [\alpha][\beta]\varphi$
5. $[\psi?]\varphi \leftrightarrow (\psi \to \varphi)$
6. $\varphi \wedge [\alpha][\alpha^*]\varphi \leftrightarrow [\alpha^*]\varphi$ (Fixed-Point Axiom)
7. $\varphi \wedge [\alpha^*](\varphi \to [\alpha]\varphi) \to [\alpha^*]\varphi$ (Induction Axiom)
8. Modus Ponens and Necessitation (for each program α)

Actions and Ability

An early approach to interpret PDL as logic of actions was put forward by Krister Segerberg.

Segerberg adds an "agency" program to the PDL language δA where A is a formula.

K. Segerberg. *Bringing it about*. JPL, 1989.

Actions and Agency

The intended meaning of the program 'δA' is that the agent "brings it about that A': *formally*, δA is the set of all paths p such that

Actions and Agency

The intended meaning of the program 'δA' is that the agent "brings it about that A': *formally*, δA is the set of all paths p such that

1. p is the computation according to some program α, and
2. α only terminates at states in which it is true that A

Actions and Agency

The intended meaning of the program 'δA' is that the agent "brings it about that A': *formally*, δA is the set of all paths p such that

1. p is the computation according to some program α, and
2. α only terminates at states in which it is true that A

Interestingly, Segerberg also briefly considers a third condition:

3. p is optimal (in some sense: shortest, maximally efficient, most convenient, etc.) in the set of computations satisfying conditions (1) and (2).

Actions and Agency

The intended meaning of the program 'δA' is that the agent "brings it about that A': *formally*, δA is the set of all paths p such that

1. p is the computation according to some program α, and
2. α only terminates at states in which it is true that A

Interestingly, Segerberg also briefly considers a third condition:

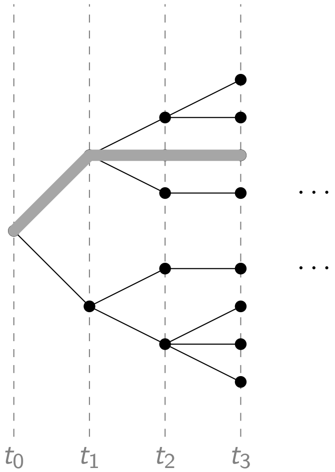
3. p is optimal (in some sense: shortest, maximally efficient, most convenient, etc.) in the set of computations satisfying conditions (1) and (2).

The axioms:

1. $[\delta A]A$
2. $[\delta A]B \rightarrow ([\delta B]C \rightarrow [\delta A]C)$

Actions and Agency in Branching Time

Alternative accounts of agency do not include explicit description of the actions:



STIT

- Each node represents a choice point for the agent.

STIT

- Each node represents a choice point for the agent.
- A **history** is a maximal branch in the above tree.

STIT

- Each node represents a choice point for the agent.
- A **history** is a maximal branch in the above tree.
- Formulas are interpreted at history moment pairs.

STIT

- Each node represents a choice point for the agent.
- A **history** is a maximal branch in the above tree.
- Formulas are interpreted at history moment pairs.
- At each moment there is a choice available to the agent (partition of the histories through that moment)

STIT

- Each node represents a choice point for the agent.
- A **history** is a maximal branch in the above tree.
- Formulas are interpreted at history moment pairs.
- At each moment there is a choice available to the agent (partition of the histories through that moment)
- The key modality is $[i\ \mathit{stit}]\varphi$ which is intended to mean that the agent i can "see to it that φ is true".
  - $[i\ \mathit{stit}]\varphi$ is true at a history moment pair provided the agent can choose a (set of) branch(es) such that every future history-moment pair satisfies φ

STIT

We use the modality '$\Diamond$' to mean historic possibility.

$\Diamond[i\ \mathit{stit}]\varphi$: "the agent has the ability to bring about φ".

STIT Model

A STIT models is $\mathcal{M} = \langle T, <, \mathit{Choice}, V \rangle$ where

STIT Model

A STIT models is $\mathcal{M} = \langle T, <, Choice, V \rangle$ where

- $\langle T, < \rangle$: T a set of moments, $<$ a tree-like ordering on T (irreflexive, transitive, linear-past)

STIT Model

A STIT models is $\mathcal{M} = \langle T, <, Choice, V \rangle$ where

- $\langle T, < \rangle$: T a set of moments, $<$ a tree-like ordering on T (irreflexive, transitive, linear-past)
- Let *Hist* be the set of all histories, and $H_t = \{h \in Hist \mid t \in h\}$ the histories through t.

STIT Model

A STIT models is $\mathcal{M} = \langle T, <, Choice, V \rangle$ where

- $\langle T, < \rangle$: T a set of moments, $<$ a tree-like ordering on T (irreflexive, transitive, linear-past)
- Let $Hist$ be the set of all histories, and $H_t = \{h \in Hist \mid t \in h\}$ the histories through t.
- $Choice : \mathcal{A} \times T \to \wp(\wp(H))$ is a function mapping each agent to a partition of H_t
  - $Choice_i^t \neq \emptyset$
  - $K \neq \emptyset$ for each $K \in Choice_i^t$
  - For all t and mappings $s_t : \mathcal{A} \to \wp(H_t)$ such that $s_t(i) \in Choice_i^t$, we have $\bigcap_{i \in \mathcal{A}} s_t(i) \neq \emptyset$

STIT Model

A STIT models is $\mathcal{M} = \langle T, <, Choice, V\rangle$ where

- $\langle T, <\rangle$: T a set of moments, $<$ a tree-like ordering on T (irreflexive, transitive, linear-past)
- Let $Hist$ be the set of all histories, and $H_t = \{h \in Hist \mid t \in h\}$ the histories through t.
- $Choice : \mathcal{A} \times T \to \wp(\wp(H))$ is a function mapping each agent to a partition of H_t
  - $Choice_i^t \neq \emptyset$
  - $K \neq \emptyset$ for each $K \in Choice_i^t$
  - For all t and mappings $s_t : \mathcal{A} \to \wp(H_t)$ such that $s_t(i) \in Choice_i^t$, we have $\bigcap_{i \in \mathcal{A}} s_t(i) \neq \emptyset$
- V : At $\to \wp(T \times Hist)$ is a valuation function assigning to each atomic proposition

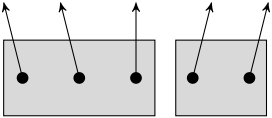
STIT Model

A STIT models is $\mathcal{M} = \langle T, <, Choice, V \rangle$ where

- $\langle T, < \rangle$: T a set of moments, $<$ a tree-like ordering on T (irreflexive, transitive, linear-past)
- Let $Hist$ be the set of all histories, and $H_t = \{h \in Hist \mid t \in h\}$ the histories through t.
- $Choice : \mathcal{A} \times T \to \wp(\wp(H))$ is a function mapping each agent to a partition of H_t
  - $Choice_i^t \neq \emptyset$
  - $K \neq \emptyset$ for each $K \in Choice_i^t$
  - For all t and mappings $s_t : \mathcal{A} \to \wp(H_t)$ such that $s_t(i) \in Choice_i^t$, we have $\bigcap_{i \in \mathcal{A}} s_t(i) \neq \emptyset$
- $V : At \to \wp(T \times Hist)$ is a valuation function assigning to each atomic proposition

Many Agents

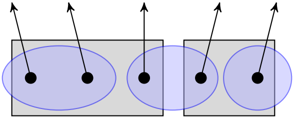
The previous model assumes there is *one* agent that "controls" the transition system.



Many Agents

The previous model assumes there is *one* agent that “controls” the transition system.

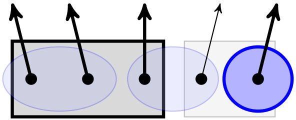
What if there is more than one agent?



Many Agents

The previous model assumes there is *one* agent that "controls" the transition system.

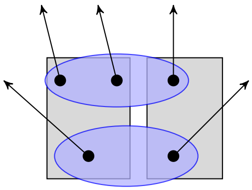
What if there is more than one agent? *Independence of agents*



Many Agents

The previous model assumes there is *one* agent that "controls" the transition system.

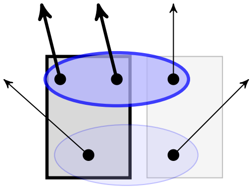
What if there is more than one agent? *Independence of agents*



Many Agents

The previous model assumes there is *one* agent that "controls" the transition system.

What if there is more than one agent? *Independence of agents*



STIT Language

$$\varphi \; = \; p \mid \neg\varphi \mid \varphi \wedge \psi \mid [i \; stit]\varphi \mid [i \; dstit : \; \varphi] \mid \Box\varphi$$

STIT Language

$$\varphi \ = \ p \mid \neg\varphi \mid \varphi \wedge \psi \mid [i \ stit]\varphi \mid [i \ dstit : \ \varphi] \mid \Box\varphi$$

- $\mathcal{M}, t/h \models p$ iff $t/h \in V(p)$

STIT Language

$$\varphi \ = \ p \mid \neg\varphi \mid \varphi \wedge \psi \mid [i \ stit]\varphi \mid [i \ dstit : \ \varphi] \mid \Box\varphi$$

- $\mathcal{M}, t/h \models p$ iff $t/h \in V(p)$
- $\mathcal{M}, t/h \models \neg\varphi$ iff $\mathcal{M}, t/h \not\models \varphi$

STIT Language

$$\varphi \ = \ p \mid \neg\varphi \mid \varphi \wedge \psi \mid [i\ stit]\varphi \mid [i\ dstit:\ \varphi] \mid \Box\varphi$$

- $\mathcal{M}, t/h \models p$ iff $t/h \in V(p)$
- $\mathcal{M}, t/h \models \neg\varphi$ iff $\mathcal{M}, t/h \not\models \varphi$
- $\mathcal{M}, t/h \models \varphi \wedge \psi$ iff $\mathcal{M}, t/h \models \varphi$ and $\mathcal{M}, t/h \models \psi$

STIT Language

$$\varphi \;=\; p \mid \neg\varphi \mid \varphi \wedge \psi \mid [i\ stit]\varphi \mid [i\ dstit:\ \varphi] \mid \Box\varphi$$

- $\mathcal{M}, t/h \models p$ iff $t/h \in V(p)$
- $\mathcal{M}, t/h \models \neg\varphi$ iff $\mathcal{M}, t/h \not\models \varphi$
- $\mathcal{M}, t/h \models \varphi \wedge \psi$ iff $\mathcal{M}, t/h \models \varphi$ and $\mathcal{M}, t/h \models \psi$
- $\mathcal{M}, t/h \models \Box\varphi$ iff $\mathcal{M}, t/h' \models \varphi$ for all $h' \in H_t$

STIT Language

$$\varphi \; = \; p \mid \neg\varphi \mid \varphi \wedge \psi \mid [i \; stit]\varphi \mid [i \; dstit: \; \varphi] \mid \Box\varphi$$

- $\mathcal{M}, t/h \models p$ iff $t/h \in V(p)$
- $\mathcal{M}, t/h \models \neg\varphi$ iff $\mathcal{M}, t/h \not\models \varphi$
- $\mathcal{M}, t/h \models \varphi \wedge \psi$ iff $\mathcal{M}, t/h \models \varphi$ and $\mathcal{M}, t/h \models \psi$
- $\mathcal{M}, t/h \models \Box\varphi$ iff $\mathcal{M}, t/h' \models \varphi$ for all $h' \in H_t$
- $\mathcal{M}, t/h \models [i \; stit]\varphi$ iff $\mathcal{M}, t/h' \models \varphi$ for all $h' \in Choice_i^t(h)$

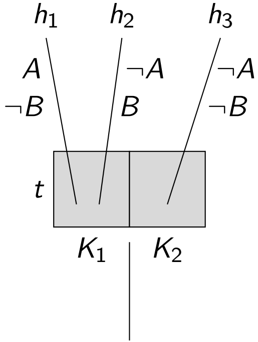
STIT Language

$$\varphi \;=\; p \mid \neg\varphi \mid \varphi \wedge \psi \mid [i\ stit]\varphi \mid [i\ dstit:\ \varphi] \mid \Box\varphi$$

- $\mathcal{M}, t/h \models p$ iff $t/h \in V(p)$
- $\mathcal{M}, t/h \models \neg\varphi$ iff $\mathcal{M}, t/h \not\models \varphi$
- $\mathcal{M}, t/h \models \varphi \wedge \psi$ iff $\mathcal{M}, t/h \models \varphi$ and $\mathcal{M}, t/h \models \psi$
- $\mathcal{M}, t/h \models \Box\varphi$ iff $\mathcal{M}, t/h' \models \varphi$ for all $h' \in H_t$
- $\mathcal{M}, t/h \models [i\ stit]\varphi$ iff $\mathcal{M}, t/h' \models \varphi$ for all $h' \in Choice_i^t(h)$
- $\mathcal{M}, t/h \models [i\ dstit]\varphi$ iff $\mathcal{M}, t/h' \models \varphi$ for all $h' \in Choice_i^t(h)$ and there is a $h'' \in H_t$ such that $\mathcal{M}, t/h \models \neg\varphi$

STIT: Example

The following are false: $A \to \Diamond[stit]A$ and $\Diamond[stit](A \vee B) \to \Diamond[stit]A \vee \Diamond[stit]B$.



J. Horty. *Agency and Deontic Logic*. 2001.

STIT: Axiomatics

- **S5** for $\Box$: $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$, $\Box\varphi \to \varphi$, $\Box\varphi \to \Box\Box\varphi$, $\neg\Box\varphi \to \Box\neg\Box\varphi$

STIT: Axiomatics

- **S5** for $\Box$: $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$, $\Box\varphi \to \varphi$, $\Box\varphi \to \Box\Box\varphi$, $\neg\Box\varphi \to \Box\neg\Box\varphi$
- **S5** for $[i\ stit]$: $[i\ stit](\varphi \to \psi) \to ([i\ stit]\varphi \to [i\ stit]\psi)$, $[i\ stit]\varphi \to \varphi$, $[i\ stit]\varphi \to [i\ stit][i\ stit]\varphi$, $\neg[i\ stit]\varphi \to [i\ stit]\neg[i\ stit]\varphi$

STIT: Axiomatics

- **S5** for $\Box$: $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$, $\Box\varphi \to \varphi$, $\Box\varphi \to \Box\Box\varphi$, $\neg\Box\varphi \to \Box\neg\Box\varphi$
- **S5** for $[i\ \mathit{stit}]$: $[i\ \mathit{stit}](\varphi \to \psi) \to ([i\ \mathit{stit}]\varphi \to [i\ \mathit{stit}]\psi)$, $[i\ \mathit{stit}]\varphi \to \varphi$, $[i\ \mathit{stit}]\varphi \to [i\ \mathit{stit}][i\ \mathit{stit}]\varphi$, $\neg[i\ \mathit{stit}]\varphi \to [i\ \mathit{stit}]\neg[i\ \mathit{stit}]\varphi$
- $\Box\varphi \to [i\ \mathit{stit}]\varphi$

STIT: Axiomatics

- **S5** for $\Box$: $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$, $\Box\varphi \to \varphi$, $\Box\varphi \to \Box\Box\varphi$, $\neg\Box\varphi \to \Box\neg\Box\varphi$
- **S5** for $[i\ \mathit{stit}]$: $[i\ \mathit{stit}](\varphi \to \psi) \to ([i\ \mathit{stit}]\varphi \to [i\ \mathit{stit}]\psi)$, $[i\ \mathit{stit}]\varphi \to \varphi$, $[i\ \mathit{stit}]\varphi \to [i\ \mathit{stit}][i\ \mathit{stit}]\varphi$, $\neg[i\ \mathit{stit}]\varphi \to [i\ \mathit{stit}]\neg[i\ \mathit{stit}]\varphi$
- $\Box\varphi \to [i\ \mathit{stit}]\varphi$
- $(\bigwedge_{i \in \mathcal{A}} \Diamond[i\ \mathit{stit}]\varphi_i) \to \Diamond(\bigwedge_{i \in \mathcal{A}}[i\ \mathit{stit}]\varphi_i)$

STIT: Axiomatics

- **S5** for $\Box$: $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$, $\Box\varphi \to \varphi$, $\Box\varphi \to \Box\Box\varphi$, $\neg\Box\varphi \to \Box\neg\Box\varphi$
- **S5** for $[i\ \mathit{stit}]$: $[i\ \mathit{stit}](\varphi \to \psi) \to ([i\ \mathit{stit}]\varphi \to [i\ \mathit{stit}]\psi)$, $[i\ \mathit{stit}]\varphi \to \varphi$, $[i\ \mathit{stit}]\varphi \to [i\ \mathit{stit}][i\ \mathit{stit}]\varphi$, $\neg[i\ \mathit{stit}]\varphi \to [i\ \mathit{stit}]\neg[i\ \mathit{stit}]\varphi$
- $\Box\varphi \to [i\ \mathit{stit}]\varphi$
- $(\bigwedge_{i \in \mathcal{A}} \Diamond[i\ \mathit{stit}]\varphi_i) \to \Diamond(\bigwedge_{i \in \mathcal{A}}[i\ \mathit{stit}]\varphi_i)$
- Modus Ponens and Necessitation for $\Box$

M. Xu. *Axioms for deliberative STIT*. Journal of Philosophical Logic, Volume 27, pp. 505 - 552, 1998.

P. Balbiani, A. Herzig and N. Troquard. *Alternative axiomatics and complexity of deliberative STIT theories*. Journal of Philosophical Logic, 37:4, pp. 387 - 406, 2008.

Recap: Logics of Action and Ability

- $F\varphi$: φ is true at some moment in *the* future

Recap: Logics of Action and Ability

- $F\varphi$: φ is true at some moment in *the* future
- $\exists F\varphi$: there is a history where φ is true some moment in the future

Recap: Logics of Action and Ability

- $F\varphi$: φ is true at some moment in *the* future
- $\exists F\varphi$: there is a history where φ is true some moment in the future
- $[\alpha]\varphi$: after doing action α, φ is true

Recap: Logics of Action and Ability

- $F\varphi$: φ is true at some moment in *the* future
- $\exists F\varphi$: there is a history where φ is true some moment in the future
- $[\alpha]\varphi$: after doing action α, φ is true
- $[\delta\varphi]\psi$: after bringing about φ, ψ is true

Recap: Logics of Action and Ability

- $F\varphi$: φ is true at some moment in *the* future
- $\exists F\varphi$: there is a history where φ is true some moment in the future
- $[\alpha]\varphi$: after doing action α, φ is true
- $[\delta\varphi]\psi$: after bringing about φ, ψ is true
- $[i\ stit]\varphi$: the agent can "see to it that" φ is true

Recap: Logics of Action and Ability

- $F\varphi$: φ is true at some moment in *the* future
- $\exists F\varphi$: there is a history where φ is true some moment in the future
- $[\alpha]\varphi$: after doing action α, φ is true
- $[\delta\varphi]\psi$: after bringing about φ, ψ is true
- $[i\ stit]\varphi$: the agent can "see to it that" φ is true
- $\Diamond[i\ stit]\varphi$: the agent has the ability to bring about φ

Epistemizing logics of action and ability

Knowing how to win

Consider the following game: Two cards, Ace and Joker, lie face down and the agent i must choose one. The Ace wins, the Joker loses.

Knowing how to win

Consider the following game: Two cards, Ace and Joker, lie face down and the agent i must choose one. The Ace wins, the Joker loses.

- Does the agent i have a strategy to win the game?

Knowing how to win

Consider the following game: Two cards, Ace and Joker, lie face down and the agent i must choose one. The Ace wins, the Joker loses.

- Does the agent i have a strategy to win the game?
- Does the agent i know that she has a strategy to win the game?

Knowing how to win

Consider the following game: Two cards, Ace and Joker, lie face down and the agent i must choose one. The Ace wins, the Joker loses.

- Does the agent i have a strategy to win the game?
- Does the agent i know that she has a strategy to win the game?
- Does the agent i know a strategy to win the game?

J. Fantl. *Knowing-how and knowing-that*. Philosophy Compass, 3 (2008), 451 470.

M.P. Singh. *Know-how*. In Foundations of Rational Agency (1999), M. Wooldridge and A. Rao, Eds., pp. 105 132.

Related Work: Knowing How to Execute a Plan

J. van Benthem. *Games in dynamic epistemic logic*. Bulletin of Economics Research 53, 4 (2001), 219 248..

J. Broersen. *A logical analysis of the interaction between obligation-to- do and knowingly doing*. In Proceedings of DEON 2008.

A. Herzig and N. Troquard. *Knowing how to play: uniform choices in logics of agency*. Proceedings of AAMAS 2006, pgs. 209 - 216.

Y. Lesperance, H. Levesque, F. Lin and R. Scherl. *Ability and Knowing How in the Situation Calculus*. Studia Logica 65, pgs. 165 - 186, 2000.

W. Jamroga and T. Agotnes. *Constructive Knowledge: What Agents can Achieve under Imperfect Information*. Journal of Applied Non-Classical Logics 17(4):423–425, 2007.

The Logic of Know-How

The Logic of Know-How

- $K(R \vee B) \rightarrow K(R) \vee K(B)$: "If Ann knows that she can choose a red or blue card, then either she knows that she can choose a red card or she knows that she can choose a blue card."

The Logic of Know-How

- $K(R \vee B) \rightarrow K(R) \vee K(B)$: "If Ann knows that she can choose a red or blue card, then either she knows that she can choose a red card or she knows that she can choose a blue card."
- $C(R \vee B) \rightarrow C(R) \vee C(B)$: "If Ann chooses either a red or blue card then either she chooses a red card or she chooses a black card."

The Logic of Know-How

- $K(R \vee B) \rightarrow K(R) \vee K(B)$: "If Ann knows that she can choose a red or blue card, then either she knows that she can choose a red card or she knows that she can choose a blue card."
- $C(R \vee B) \rightarrow C(R) \vee C(B)$: "If Ann chooses either a red or blue card then either she chooses a red card or she chooses a black card."
- $Abl(R \vee B) \rightarrow Abl(R) \vee Abl(B)$: "If Ann has the ability to select a red or blue card then either she has the ability to choose a red card or she has the ability to choose a black card."

The Logic of Know-How

- $K(R \vee B) \to K(R) \vee K(B)$: "If Ann knows that she can choose a red or blue card, then either she knows that she can choose a red card or she knows that she can choose a blue card."
- $C(R \vee B) \to C(R) \vee C(B)$: "If Ann chooses either a red or blue card then either she chooses a red card or she chooses a black card."
- $Abl(R \vee B) \to Abl(R) \vee Abl(B)$: "If Ann has the ability to select a red or blue card then either she has the ability to choose a red card or she has the ability to choose a black card."
- $Khow(R \vee B) \to Khow(R) \vee Khow(B)$: "If Ann knows how to select a red or blue card then either she knows how to choose a red card or she knows how to choose a black card."

Grades of Know-How

i knows how to α only if:

1. it is possible that i α
2. were i to try to α, i would α
3. were i to try to α is a suitable context, i would α
4. i is able/has the ability to α particularly well
5. i knows that w is a way to α
6. i knows that w is a way for her to α
7. i knows why w is a way for her to α

J. Fantl. *Knowing-how and knowing-that*. Philosophy Compass 3, 3 (2008), 451 470.

Example

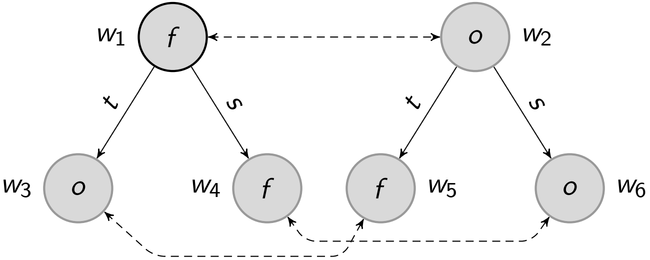
A. Herzig and N. Troquard. *Knowing how to play: uniform choices in logics of agency*. In Proceedings of AAMAS 2006.

Example

Ann, who is blind, is standing with her hand on a light switch. She has two options: toggle the switch (t) or do nothing (s):

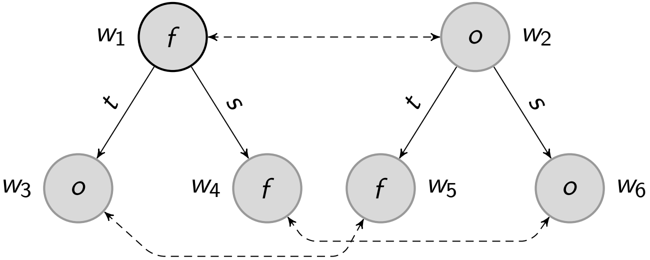
Example

Ann, who is blind, is standing with her hand on a light switch. She has two options: toggle the switch (t) or do nothing (s):



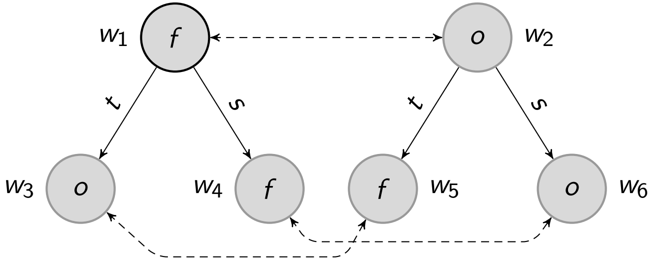
Example

Ann, who is blind, is standing with her hand on a light switch. She has two options: toggle the switch (t) or do nothing (s):



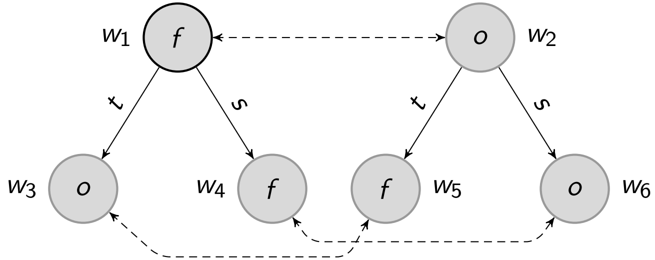
Does she have the *ability* to turn the light on? Is she *capable* of turning the light on? Does she *know how* to turn the light on?

Example



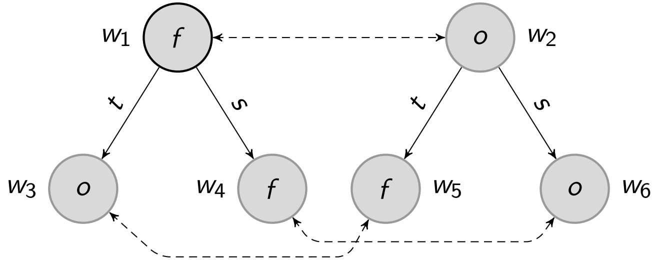
$w_1 \models \neg\Box f$: "Ann does not know the light is on"

Example



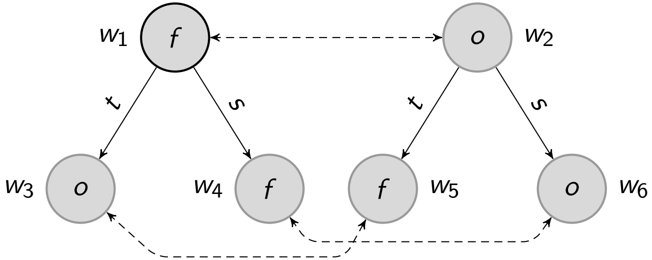
$w_1 \models \langle t \rangle o$ "after toggling the light switch, the light will be on"

Example



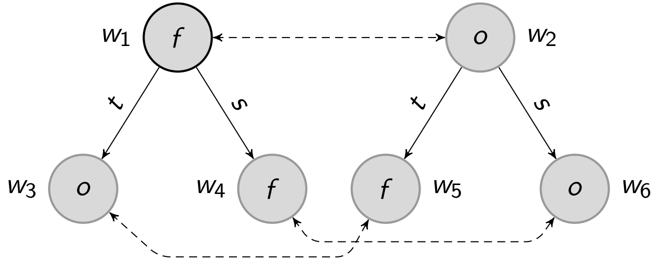
$w_1 \models \neg\Box\langle t\rangle o$: "Ann does not know that after toggling the light switch, the light will be on"

Example



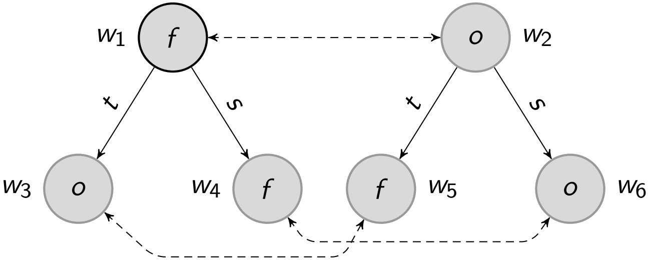
$w_1 \models \Box(\langle t \rangle \top \wedge \langle s \rangle \top)$: "Ann knows that she can toggle the switch and she can do nothing"

Example



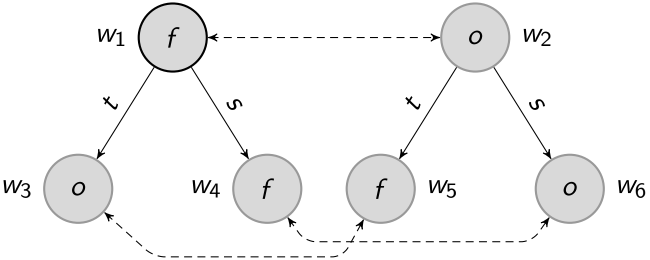
$w_1 \models \langle t \rangle \neg \Box o$: "after toggling the switch Ann does not know that the light is on"

Example



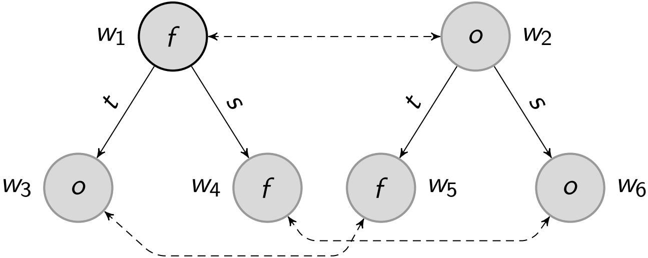
Let l be "turn the light on": a choice between t and s

Example



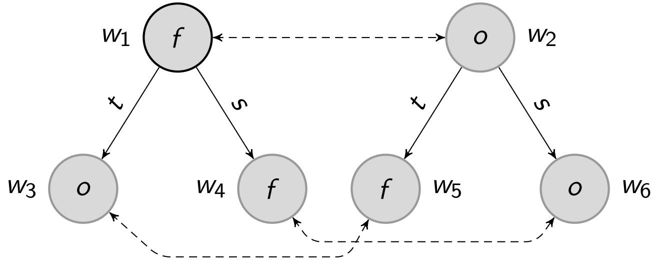
$w_1 \models \langle l \rangle^{\exists} o \wedge \neg \langle l \rangle^{\forall} o$: executing l can lead to a situation where the light is on, but this is not *guaranteed* (i.e., the plan may fail)

Example



$w_1 \models \Box\langle l\rangle^{\exists}o$: Ann knows that she is capable of turning the light on. She has *de re* knowledge that she can turn the light on.

Example



$w_1 \models \neg\langle l\rangle^{\Diamond} o$: Ann cannot knowingly turn on the light: there is no *subjective* path leading to states satisfying o (note that *all* elements of the last element of the subject path must satisfy o).

J. Broersen. *Deontic epistemic stit logic distinguishing modes of mens rea*. Journal of Applied Logic, 9, pgs. 137 - 152, 2011.

XSTIT

Language: $\varphi \ := \ p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box\varphi \mid [A\ xstit]\varphi \mid X\varphi$

XSTIT

Language: $\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box\varphi \mid [A\ xstit]\varphi \mid X\varphi$

Frame $\langle S, H, R_X, R_\Box, \{R_A \mid A \subseteq Ags\}\rangle$ such that

- S is an infinite set of static states
- $H \subseteq 2^{2^S - \emptyset}$ is a non-empty set of histories (maximal linearly ordered subsets of S). Dynamic states are tuples $\langle s, h\rangle$ with $s \in h$.
- R_X is serial and deterministic: $\langle s, h\rangle R_X \langle s', h'\rangle$ implies $h = h'$
- $R_\Box$ is defined as follows: $\langle s, h\rangle R_\Box \langle s', h'\rangle$ iff $s = s'$

XSTIT

Frame $\langle S, H, R_X, R_\Box, \{R_A \mid A \subseteq Ags\}\rangle$ such that

- R_A are effectivity relations
  - $R_\emptyset = R_\Box \circ R_X$
  - $R_{Ags} = R_X \circ R_\Box$
  - $R_A \subseteq R_B$ for $B \subseteq A$
  - For $A \cap B = \emptyset$, if $\langle s_1, h_1\rangle R_\Box \langle s_2, h_2\rangle$ and $\langle s_1, h_1\rangle R_\Box \langle s_3, h_3\rangle$, then there is $\langle s_4, h_4\rangle$ such that $\langle s_1, h_1\rangle R_\Box \langle s_4, h_4\rangle$, and if $\langle s_4, h_4\rangle R_A \langle s_5, h_5\rangle$ then $\langle s_2, h_2\rangle R_A \langle s_5, h_5\rangle$, and if $\langle s_4, h_4\rangle R_B \langle s_6, h_6\rangle$ then $\langle s_3, h_3\rangle R_B \langle s_6, h_6\rangle$

XSTIT

- $\mathcal{M}, \langle s, h \rangle \models p$ iff $\langle s, h \rangle \in V(p)$
- $\mathcal{M}, \langle s, h \rangle \models \neg\varphi$ iff $\mathcal{M}, \langle s, h \rangle \not\models \varphi$
- $\mathcal{M}, \langle s, h \rangle \models \varphi \wedge \psi$ iff $\mathcal{M}, \langle s, h \rangle \models \varphi$ and $\mathcal{M}, \langle s, h \rangle \models \psi$
- $\mathcal{M}, \langle s, h \rangle \models \Box\varphi$ iff $\langle s, h \rangle R_{\Box} \langle s', h' \rangle$ implies $\mathcal{M}, \langle s', h' \rangle \models \varphi$
- $\mathcal{M}, \langle s, h \rangle \models X\varphi$ iff $\langle s, h \rangle R_{X} \langle s', h' \rangle$ implies $\mathcal{M}, \langle s', h' \rangle \models \varphi$
- $\mathcal{M}, \langle s, h \rangle \models [A\ xstit]\varphi$ iff $\langle s, h \rangle R_{A} \langle s', h' \rangle$ implies $\mathcal{M}, \langle s', h' \rangle \models \varphi$

XSTIT Axiomatization

- $S5$ for $\Box$
- KD for each $[A\ xstit]$
- $\neg X \neg \varphi \rightarrow X\varphi$
- $[Ags\ xstit]\varphi \leftrightarrow X\Box\varphi$
- $[\emptyset\ xstit]\varphi \leftrightarrow \Box X\varphi$
- $[A\ xstit]\varphi \rightarrow [A \cup B\ stit]\varphi$
- $(\Diamond[A\ xstit]\varphi \wedge \Diamond[B\ xstit]\psi) \rightarrow \Diamond([A\ xstit]\varphi \wedge [B\ xstit]\psi)$

Epistemic XSTIT

Frame $\langle S, H, R_X, R_\Box, \{R_A \mid A \subseteq Ags\}, \{\sim_a \mid a \in Ags\}\rangle$ such that

- $\sim_a$ is an equivalence relation over dynamic states

- $\mathcal{M}, \langle s, h\rangle \models K_a\varphi$ iff $\langle s, h\rangle \sim_a \langle s', h'\rangle$ implies $\mathcal{M}, \langle s', h'\rangle \models \varphi$

Knowingly doing: $K_a[a\ xstit]\varphi$

Having the ability to do something: $\Diamond K_a[a\ xstit]\varphi$

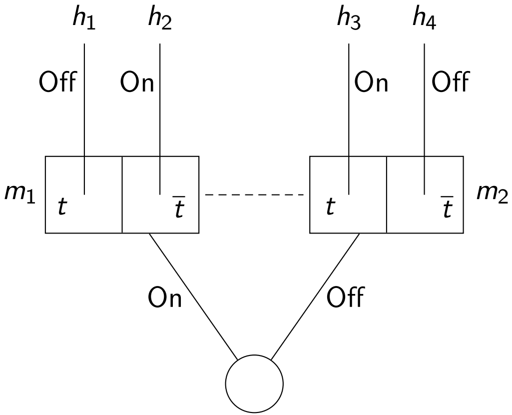
Knowing to have the capacity to cause a certain effect, without knowing what to do to cause that effect: $K_a\Diamond[a\ xstit]\varphi$

Seeing to it that a learns φ: $[a\ xstit]K_a\varphi$.

Epistemic XSTIT

- *knowledge about next state*: $K_a X\varphi \to K_a[a\ xstit]\varphi$ is valid iff $\sim_a \circ R_a \subseteq \sim_a \circ R_X$
- effect recollection: $K_a[a\ xstit]\varphi \to XK_a\varphi$ iff $R_X \circ \sim_a \subseteq \sim_a \circ R_a$.
- *uniformity of historical possibility*: $\Diamond K_a\varphi \to K_a\Diamond\varphi$ is valid iff if $\langle s_1, h_1\rangle R_\Box \langle s_2, h_2\rangle$ and $\langle s_1, h_1\rangle \sim_a \langle s_3, h_3\rangle$ then there is $\langle s_4, h_4\rangle$ such that $\langle s_3, h_3\rangle R_\Box \langle s_4, h_4\rangle$ and $\langle s_2, h_2\rangle \sim_a \langle s_4, h_4\rangle$

Returning to the Example

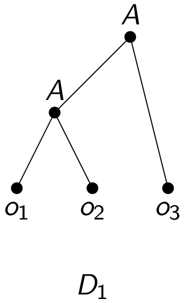


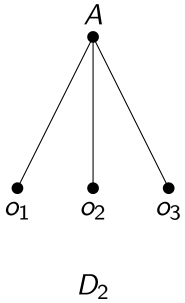
When are two games the *same*?

When are two games the *same*?

- Whose point-of-view? (players, modelers)
- Game-theoretic analysis should not depend on "irrelevant" mathematical details
- Different perspectives: transformations, structural, agent

The same decision problem



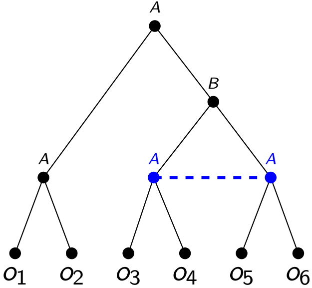


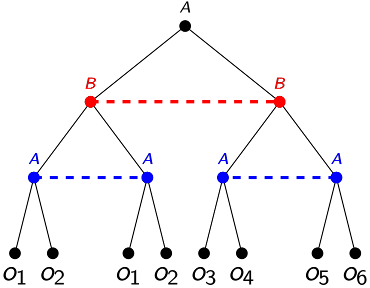
Thompson Transformations

Game-theoretic analysis should not depend on “irrelevant” features of the (mathematical) description of the game.

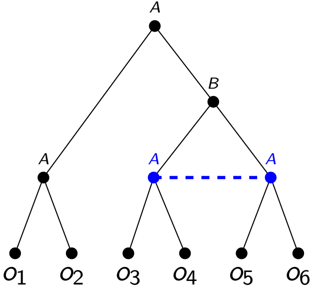
F. B. Thompson. *Equivalence of Games in Extensive Form*. Classics in Game Theory, pgs 36 - 45, 1952.

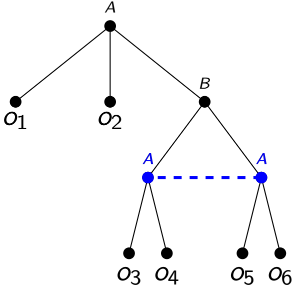
(Osborne and Rubinstein, pgs. 203 - 212)



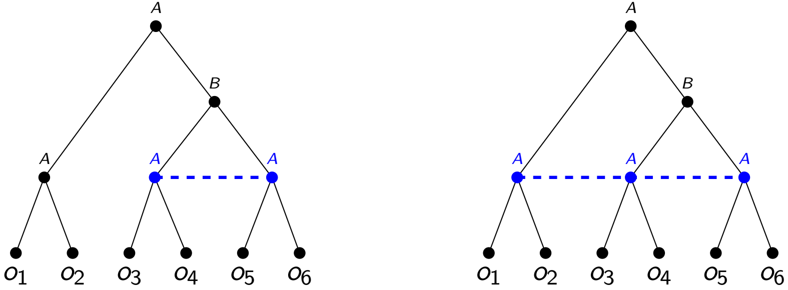


Addition of Superfluous Move

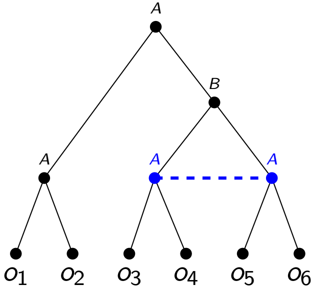




Coalescing of moves



Inflation/deflation



Interchange of moves

Theorem (Thompson) Each of the previous transformations preserves the reduced strategic form of the game. In finite extensive games (without uncertainty between subhistories), if any two games have the same reduced normal form then one can be obtained from the other by a sequence of the four transformations.

Other transformations/game forms

Kohlberg and Mertens. *On Strategic Stability of Equilibria*. Econometrica (1986).

Elmes and Reny. *On The Strategic Equivalence of Extensive Form Games*. Journal of Economic Theory (1994).

G. Bonanno. *Set-Theoretic Equivalence of Extensive-Form Games*. IJGT (1992).

Games as Processes

- When are two processes the same?

Games as Processes

- When are two processes the same?
- Extensive games are natural process models which support many familiar modal logics such as *bisimulation*.

Games as Processes

- When are two processes the same?
- Extensive games are natural process models which support many familiar modal logics such as *bisimulation*.
- From this point-of-view, "When are two games the same?" goes tandem with asking "what are appropriate languages for games"

Games as Processes

- When are two processes the same?
- Extensive games are natural process models which support many familiar modal logics such as *bisimulation*.
- From this point-of-view, "When are two games the same?" goes tandem with asking "what are appropriate languages for games"

J. van Benthem. *Extensive Games as Process Models*. IJGT, 2001.