Introduction to Logics of Knowledge and Belief

Eric Pacuit

University of Maryland

pacuit.org

epacuit@umd.edu

May 13, 2019

Knowing what

i knows what the value of c

$\exists x K_i(c = x)$

Knowing what

$$\varphi ::= \top \mid p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_i\varphi \mid Kv_i c$$

where $p \in \mathsf{At}$ and $c \in \mathbf{C}$ (a set of constant symbols)

$\mathcal{M} = \langle W, D, (R_i)_{i\in\mathcal{A}}, V, V_C\rangle$

where $W \neq \emptyset$, each R_i is a relation on W, $V : \mathsf{At} \to \wp(W)$, D is the constant domain and $V_C : \mathbf{C} \times W \to D$ assigns to each $c \in \mathbf{C}$ and world w a value $d \in D$.

$$\mathcal{M}, w \models Kv_i c \text{ iff for any } v_1, v_2, \text{ if } wR_iv_1 \text{ and } wR_iv_2,$$

$$\text{then } V_C(v_1, c) = v_C(c, v_2)$$

$$K_i Kv_j c \wedge \neg Kv_j c \quad \text{vs.} \quad K_i K_j p \wedge \neg K_i p$$

$$K_i Kv_j c \wedge \neg Kv_j c \quad \text{vs.} \quad K_i K_j p \wedge \neg K_i p$$

$$\varphi ::= \top \mid p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_i\varphi \mid Kv_i c \mid [\varphi]\varphi$$

$(\langle p \rangle Kv_i c \wedge \langle q \rangle Kv_i c) \to \langle p \vee c \rangle Kv_i c$ is not derivable is S5 with recursion axioms.

Y. Wang and J. Fan. *Knowing that, knowing what, and public communication: Public announcement logic with Kv operators*. In: Proceedings of IJCAI?13, pp 1139 - 1146, 2013.

A. Baltag. *To Know is to Know the Value of a Variable*. AiML, 2016.

Y. Wang. *A New Modal Framework for Epistemic Logic*. TARK 2017.

Know how

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. Woodridge 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.

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.

Knowledge, action, abilities

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

A. Herzig. *Logics of knowledge and action: critical analysis and challenges*. Autonomous Agent and Multi-Agent Systems, 2014.

J. Broeresen, A. Herzig and N. Troquard. *What groups do, can do and know they can do: An analysis in normal modal logics*. Journal of Applied and Non-Classical Logics, 19:3, pgs. 261 - 289, 2009.

W. van der Hoek and M. Wooldridge. *Cooperation, knowledge and time: Alternating-time temporal epistemic logic and its applications*. Studia Logica, 75, pgs. 125 - 157, 2003.

Example

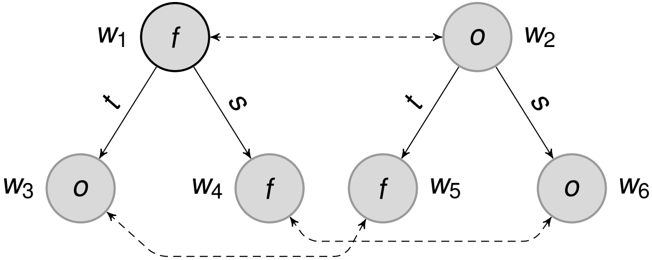
A. Herzig and N. Troquard. *Knowing how to play: uniform choices in logics of agency*. In Proceedings of AAMAS 2006.

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):

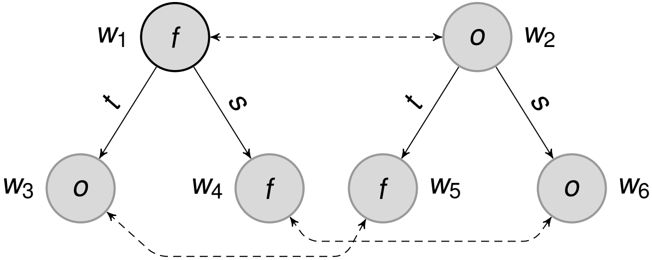
Example

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):



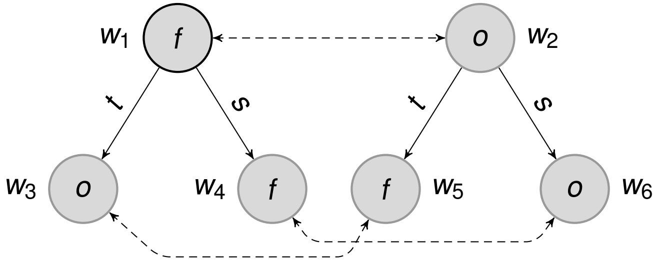
Example

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):



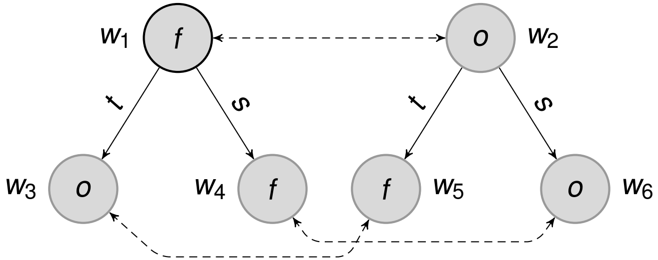
Does she have the *ability* to turn the light on? Is she *capable* of turning the light on? Does she *know how* to turn the light on?

Example



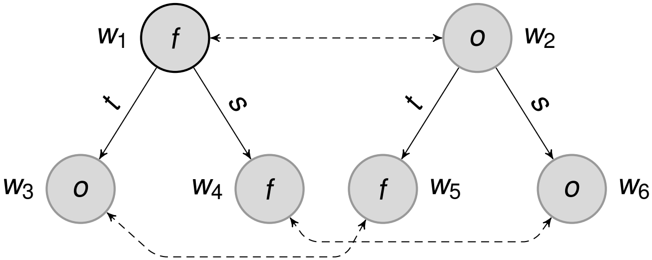
$w_1 \models \neg\Box f$: “Ann does not know the light is on”

Example



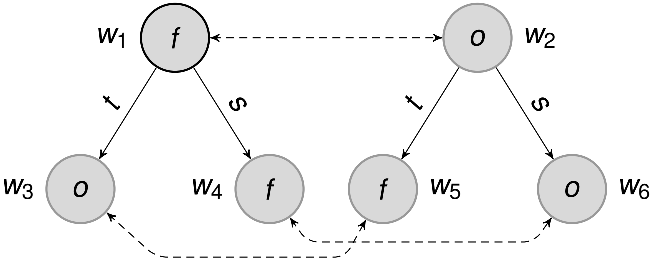
$w_1 \models \langle t \rangle o$ "after toggling the light switch, the light will be on"

Example



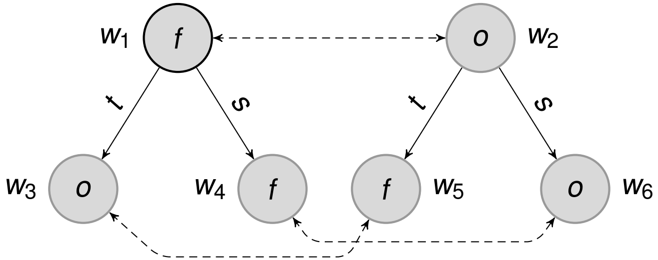
$w_1 \models \neg\Box\langle t\rangle o$: “Ann does not know that after toggling the light switch, the light will be on”

Example



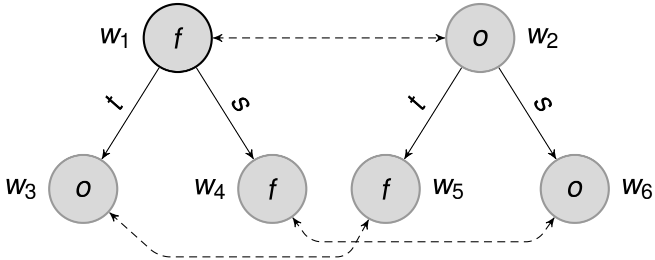
$w_1 \models \Box(\langle t\rangle\top \wedge \langle s\rangle\top)$: "Ann knows that she can toggle the switch and she can do nothing"

Example



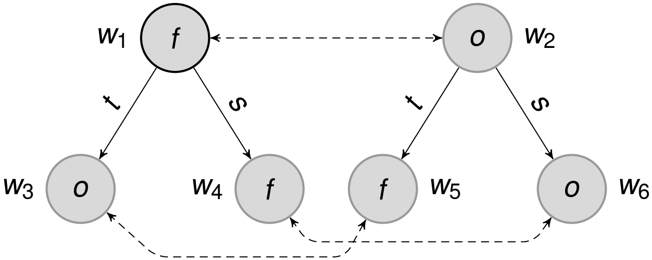
$w_1 \models \langle t \rangle \neg \Box o$: "after toggling the switch Ann does not know that the light is on"

Example



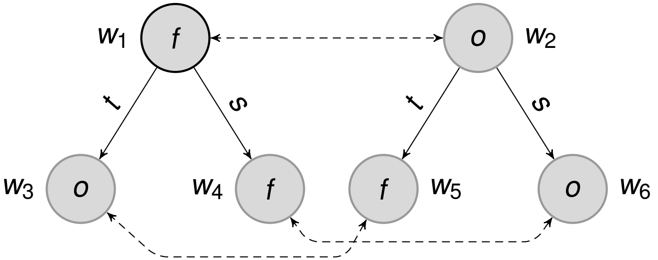
Let l be "turn the light on": a choice between t and s

Example



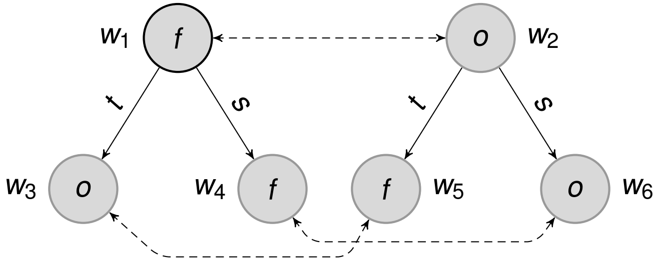
$w_1 \models \langle l \rangle^{\exists} o \wedge \neg \langle l \rangle^{\forall} o$: executing l can lead to a situation where the light is on, but this is not *guaranteed* (i.e., the plan may fail)

Example



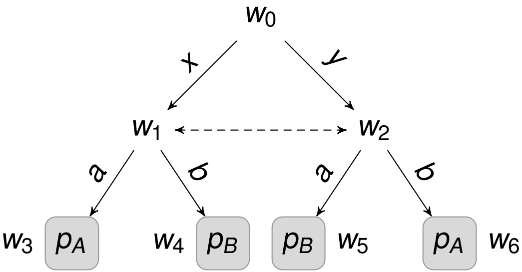
$w_1 \models \Box\langle l\rangle^{\exists} o$: Ann knows that she is capable of turning the light on. She has *de re* knowledge that she can turn the light on.

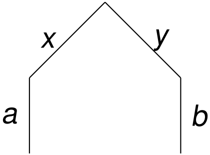
Example



$w_1 \models \neg\langle l\rangle^{\diamond} o$: 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).

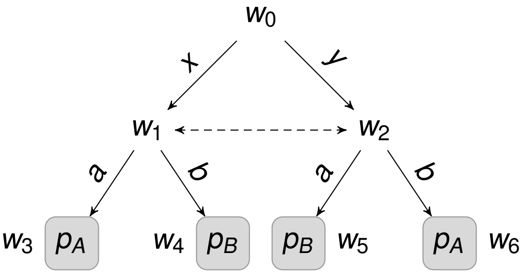
Knowing How to Win

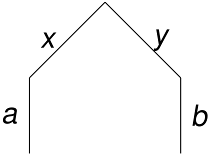




"the plan is a winning strategy for Ann."

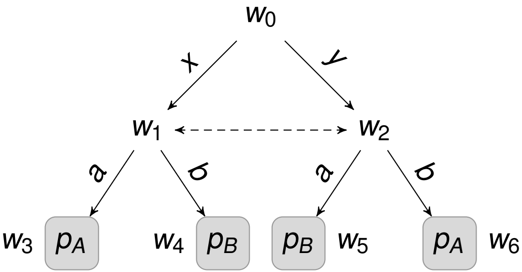
Knowing How to Win

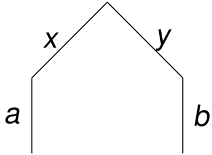




"Ann knows that the plan is a winning strategy."

Knowing How to Win





" the plan can be executed, but Ann does not know how to use it to win."

Y. Wang. *A New Modal Framework for Epistemic Logic*. TARK 2017.

The *know-wh* modalities all share a general *de re* schema: $\exists x \Box \varphi(x)$ (*mention-some*)

"knowing how to achieve φ" roughly means that there exists a way such that you know that it is a way to ensure that φ

"knowing why φ" means that there exists an explanation such that you know that it is an explanation to the fact φ.
mention-all interpretation: "knowing who came to the party" means, under an exhaustive reading, that for each relevant person, you know whether (s)he came to the party or not: $\forall x(\Box\varphi(x) \vee \Box\neg\varphi(x))$.

“knowing [what] the value of c [is]” means, under the interpretation of mention-some, that there exists a value such that you know that it is the value of c, which is equivalent to the mention-all interpretation: for any value, you know whether it is the value of c, given there is one and only one real value of c.

The logical core of the "mention-some" logics: $\Box^x\varphi$ is a *packaging* of $\exists x\Box$.

"I know a theorem of which I do not know any proof": $\Box^x\neg\Box^y Prove(y,x)$

"*i* knows a country which *j* knows its capital": $\Box_i^x\Box_j^y Capital(y,x)$

Let **X** be a set of variables and **P** a set of predicate symbols.

$$\varphi \ ::= \ x \approx y \mid P\overline{x} \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \Box^x \varphi$$

where $x, y \in X$ and $P \in \mathbf{P}$

$$\mathcal{M} = \langle W, D, \delta, R, \rho \rangle$$

- $W \neq \emptyset$ is a set of worlds
- $D \neq \emptyset$ is the domain
- $R \subseteq W \times W$ is an accessibility relation
- $\delta : W \rightarrow \wp(D)$ assigns to each $w \in W$ a non-empty local domain such that wRv implies that $\delta(w) \subseteq \delta(v)$ (write D_w for $\delta(w)$)
- $\rho : \mathbf{P} \times W \rightarrow \bigcup_{n \in \omega} \wp(D^n)$ assignes to each n-ary predicate and world, an n-ary relation on D.

$\sigma : \mathbf{X} \rightarrow D$ is a variable assignment.

- $\mathcal{M}, w, \sigma \models x \approx y$ iff $\sigma(x) = \sigma(y)$
- $\mathcal{M}, w, \sigma \models P(x_1, \ldots, x_n)$ iff $(\sigma(x_1), \ldots, \sigma(x_n)) \in \rho(P, w)$
- $\mathcal{M}, w, \sigma \models \neg\varphi$ iff $\mathcal{M}, w, \sigma \not\models \varphi$
- $\mathcal{M}, w, \sigma \models \varphi \wedge \psi$ iff $\mathcal{M}, w, \sigma \models \varphi$ and $\mathcal{M}, w, \sigma \models \psi$
- $\mathcal{M}, w, \sigma \models \Box^x\varphi$ iff there is an $a \in \delta(w)$ such that $\mathcal{M}, v, \sigma[x \mapsto a] \models \varphi$ for all v such that wRv

Reminder: bisimulation for modal logic

- **Language**: $p \mid \neg\varphi \mid \varphi \vee \psi \mid \Box\psi$, $p \in \mathsf{At}$ (atomic propositions), Boolean connectives defined as usual, $\Diamond\varphi \;:=\; \neg\Box\neg\varphi$
- **Frame**: $\langle W, R\rangle$, where $W \neq \emptyset$ and $R \subseteq W \times W$
- **Model**: $\langle W, R, V\rangle$, where $\langle W, R\rangle$ is a frame and $V : \mathsf{At} \to \wp(W)$ (Kripke structure)
- **Truth at a state in a model**: $\mathcal{M}, w \models \varphi$
  - $\mathcal{M}, w \models p$ iff $w \in V(p)$
  - $\mathcal{M}, w \models \neg\varphi$ iff $\mathcal{M}, w \not\models \varphi$
  - $\mathcal{M}, w \models \varphi \wedge \psi$ iff $\mathcal{M}, w \models \varphi$ and $\mathcal{M}, w \models \psi$
  - $\mathcal{M}, w \models \Box\varphi$ iff for all $v \in W$, if wRv then $\mathcal{M}, v \models \varphi$

Reminder: bisimulation for modal logic

A bisimulation between $\mathcal{M} = \langle W, R, V \rangle$and $\mathcal{M}' = \langle W', R', V' \rangle$ is a non-empty binary relation $Z \subseteq W \times W'$ such that whenever wZw':

Atomic harmony: for each $p \in \mathrm{At}$, $w \in V(p)$ iff $w' \in V'(p)$
Zig: if wRv, then $\exists v' \in W'$ such that vZv' and $w'R'v'$
Zag: if $w'R'v'$ then $\exists v \in W$ such that vZv' and wRv

Reminder: bisimulation for modal logic

- *We write $\mathcal{M}, w \underline{\leftrightarrow} \mathcal{M}', w'$ if there is a Z such that wZw'.*
- *We write $\mathcal{M}, w \equiv_{\mathcal{L}} \mathcal{M}', w'$ iff for all $\varphi \in \mathcal{L}$, $\mathcal{M}, w \models \varphi$ iff $\mathcal{M}', w' \models \varphi$.*
- **Lemma** If $\mathcal{M}, w \underline{\leftrightarrow} \mathcal{M}', w'$ then $\mathcal{M}, w \equiv_L \mathcal{M}', w'$.
- **Lemma** On finite models, if $\mathcal{M}, w \equiv_L \mathcal{M}', w'$ then $\mathcal{M}, w \underline{\leftrightarrow} \mathcal{M}', w'$.
- **Lemma** On *m*-**saturated models**, if $\mathcal{M}, w \equiv_L \mathcal{M}', w'$ then $\mathcal{M}, w \underline{\leftrightarrow} \mathcal{M}', w'$.

Reminder: monotonic neighborhood bisimulations

Let W be a non-empty set of states.

Any function $N : W \to \wp(\wp(W))$ is called a neighborhood function

A pair $\langle W, N \rangle$ is a called a neighborhood frame if W a non-empty set and N is a neighborhood function that is closed under supersets.

A neighborhood model based on $\mathfrak{F} = \langle W, N \rangle$ is a tuple $\langle W, N, V \rangle$ where $V : \mathsf{At} \to \wp(W)$ is a valuation function.

Reminder: monotonic neighborhood bisimulations

- $\mathfrak{M}, w \models p$ iff $w \in V(p)$
- $\mathfrak{M}, w \models \neg\varphi$ iff $\mathfrak{M}, w \not\models \varphi$
- $\mathfrak{M}, w \models \varphi \wedge \psi$ iff $\mathfrak{M}, w \models \varphi$ and $\mathfrak{M}, w \models \psi$
- $\mathfrak{M}, w \models \Box\varphi$ iff there is a $X \in N(w)$ such that $X \subseteq [\![\varphi]\!]_{\mathfrak{M}}$

where $[\![\varphi]\!]_{\mathfrak{M}} = \{w \mid \mathfrak{M}, w \models \varphi\}$.

Reminder: monotonic neighborhood bisimulations

M. Pauly. *Bisimulation for Non-normal Modal Logic*. Manuscript, 1999.

H. Hansen. *Monotonic Modal Logic*. ILLC, Masters Thesis, 2003.

Reminder: monotonic neighborhood bisimulations

Suppose that $\mathfrak{M} = \langle W, N, V \rangle$ and $\mathfrak{M}' = \langle W', N', V' \rangle$ are two monotonic neighborhood models. A relation $Z \subseteq W \times W'$ is a **monotonic bisimulation** provided that, whenever wZw':

Atomic harmony: for each $p \in \mathrm{At}$, $w \in V(p)$ iff $w' \in V'(p)$.

Zig: If $w\ N\ X$ then there is an $X' \subseteq W'$ such that $w'\ N'\ X'$ and $\forall x' \in X'$, $\exists x \in X$ such that $x\ Z\ x'$.

Zag: If $w'\ N'\ X'$ then there is an $X \subseteq W$ such that $w\ N\ X$ and $\forall x \in X$, $\exists x' \in X'$ such that $x\ Z\ x'$.

Write $\mathfrak{M}, w \underline{\leftrightarrow} \mathfrak{M}', w'$ when there is a monotonic bisimulation $Z \subseteq dom(\mathfrak{M}) \times dom(\mathfrak{M}')$ such that $w\ Z\ w'$.

Reminder: monotonic neighborhood bisimulations

- **Lemma**. If $\mathcal{M}$ is a monotonic model, $\mathfrak{M}, w \,\underline{\leftrightarrow}\, \mathfrak{M}', w'$ implies $\mathfrak{M}, w \equiv_{\mathcal{L}} \mathfrak{M}', w'$.

Reminder: monotonic neighborhood bisimulations

- Suppose that $\mathcal{F}$ is a monotonic collection of subsets of W. The **non-monotonic core**, denoted $\mathcal{F}^{nc}$, is a subset of $\mathcal{F}$ defined as follows: $\mathcal{F}^{nc} = \{X \mid X \in \mathcal{F}$ and for all $X' \subseteq W$, if $X' \subsetneq X$, then $X' \notin \mathcal{F}\}$.
  A monotonic collection of sets $\mathcal{F}$ is **core-complete** provided for all $X \in \mathcal{F}$, there exists a $Y \in \mathcal{F}^{nc}$ such that $Y \subseteq X$.
- A neighborhood model $\mathfrak{M} = \langle W, N, V \rangle$ is **locally core-finite** provided that $\mathfrak{M}$ is core-complete and for each $w \in W$, $N^{nc}(w)$ is finite, and for all $X \in N^{nc}(w)$, X is finite.

  Lemma. Suppose that $\mathfrak{M} = \langle W, N, V \rangle$ and $\mathfrak{M}' = \langle W', N', V' \rangle$ are monotonic, locally core-finite models. Then, for all $w \in W$, $w' \in W'$, $\mathfrak{M}, w \equiv_{\mathcal{L}} \mathfrak{M}', w'$ iff $\mathfrak{M}, w \underline{\leftrightarrow} \mathfrak{M}', w'$.

- $\mathcal{M}, w, \sigma \models x \approx y$ iff $\sigma(x) = \sigma(y)$
- $\mathcal{M}, w, \sigma \models P(x_1, \ldots, x_n)$ iff $(\sigma(x_1), \ldots, \sigma(x_n)) \in \rho(P, w)$
- $\mathcal{M}, w, \sigma \models \neg\varphi$ iff $\mathcal{M}, w, \sigma \not\models \varphi$
- $\mathcal{M}, w, \sigma \models \varphi \wedge \psi$ iff $\mathcal{M}, w, \sigma \models \varphi$ and $\mathcal{M}, w, \sigma \models \psi$
- $\mathcal{M}, w, \sigma \models \Box^x\varphi$ iff there is an $a \in \delta(w)$ such that $\mathcal{M}, v, \sigma[x \mapsto a] \models \varphi$ for all v such that wRv

Let $\mathcal{M}$, $\mathcal{N}$ be two models, a relation $Z \subseteq (W^{\mathcal{M}} \times D^*_{\mathcal{M}}) \times (W^{\mathcal{N}} \times D^*_{\mathcal{N}})$ is an $\exists\Box$-bisimulation if for every $((w, \overline{(}a)), (v, \overline{b})) \in Z$ such that $|a| = |b|$, the following holds:

PISO: $\overline{a}$ and $\overline{b}$ form a partial isomporphism wr.t. identity and the interpretations of the predicates at w and v

$\exists\Box$Zig: for any $c \in D^{\mathcal{M}}_w$, there is a $d \in D^{\mathcal{N}}_v$ such that for any $v' \in W^{\mathcal{N}}$, if $vR^{\mathcal{N}}v'$ then there is a $w' \in W^{\mathcal{M}}$ such that $wR^{\mathcal{M}}w'$ and $w'\overline{a}cZv'\overline{b}d$

$\exists\Box$Zag: for any $d \in D^{\mathcal{N}}_v$, there is a $c \in D^{\mathcal{M}}_w$ such that for any $w' \in W^{\mathcal{M}}$, if $wR^{\mathcal{M}}w'$ then there is a $v' \in W^{\mathcal{N}}$ such that $vR^{\mathcal{N}}v'$ and $w'\overline{a}cZv'\overline{b}d$

We say that $\mathcal{M}, w\overline{a}$ and $\mathcal{N}, v\overline{b}$ are $\exists\Box$-bisimilar (denoted $\mathcal{M}, w\overline{a} \underline{\leftrightarrow}_{\exists\Box} \mathcal{N}, v\overline{b}$) if $|a| = |b|$ and there is a $\exists\Box$-bisimulation connecting $w\overline{a}$ and $v\overline{b}$

- **Lemma.** If $\mathcal{M}, w\overline{a} \underline{\leftrightarrow}_{\exists\Box} \mathcal{N}, v\overline{b}$, then $\mathcal{M}, w\overline{a} \equiv_{MLMS^{\approx}} \mathcal{N}, v\overline{b}$
- **Lemma.** If $\mathcal{M}, w \underline{\leftrightarrow}_{\exists\Box} \mathcal{N}, v$, then for all closed formula φ, $\mathcal{M}, w \models \varphi$ iff $\mathcal{N}, v \models \varphi$.
- $\Box\exists xPx$, $\exists x\Diamond Px$ and $\Diamond\exists xPx$ are not expressible in $MLMS^{\approx}$.
- If $\mathcal{M}$ and $\mathcal{N}$ are finite ($\exists\Box$-saturated) and $|\overline{a}| = |\overline{b}|$, then $\mathcal{M}, w\overline{a} \underline{\leftrightarrow}_{\exists\Box} \mathcal{N}, v\overline{b}$ iff $\mathcal{M}, w\overline{a} \equiv_{MLMS^{\approx}} \mathcal{N}, v\overline{b}$

A New Epistemic Logic

Let **X** be a set of variables and **P** a set of predicate symbols.

$$\varphi \ ::= \ x \approx y \mid P\overline{x} \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \Box^x \varphi \mid \Box\varphi$$

where $x, y \in X$ and $P \in \mathbf{P}$

The models are the same except:

- Each R is an equivalence relation
- For all $w \in W$, $D_w = D$

$\mathcal{M}, w, \sigma \models \blacksquare^x \varphi$ iff for each $d \in D$, either $\mathcal{M}, w\sigma[x \mapsto d] \models \Box\varphi$ or $\mathcal{M}, w\sigma[x \mapsto d] \models \Box\neg\varphi$

$\mathcal{M}, w, \sigma \models \blacksquare^x \varphi$ iff for each $d \in D$, either $\mathcal{M}, w\sigma[x \mapsto d] \models \Box\varphi$ or $\mathcal{M}, w\sigma[x \mapsto d] \models \Box\neg\varphi$

$\mathcal{M}, w, \sigma \models \Box^{\forall x}\varphi$ iff for each $d \in D$, $\mathcal{M}, w\sigma[x \mapsto d] \models \Box\varphi$

$\mathcal{M}, w, \sigma \models \blacksquare^x \varphi$ iff for each $d \in D$, either $\mathcal{M}, w\sigma[x \mapsto d] \models \Box\varphi$ or $\mathcal{M}, w\sigma[x \mapsto d] \models \Box\neg\varphi$

$\mathcal{M}, w, \sigma \models \Box^{\forall x} \varphi$ iff for each $d \in D$, $\mathcal{M}, w\sigma[x \mapsto d] \models \Box\varphi$

$\mathcal{M}, w, \sigma \models \Box^{x_1 \cdots x_n} \varphi$ iff there is $d_1, \ldots, d_n \in D$ such that $\mathcal{M}, w\sigma[\overline{x} \mapsto \overline{d}] \models \Box\varphi$

$$\blacksquare^x \varphi \quad \leftrightarrow \quad \Diamond^x(\Box\varphi \vee \Box\neg\varphi)$$

$$\blacksquare^{x}\varphi \quad \leftrightarrow \quad \Diamond^{x}(\Box\varphi \vee \Box\neg\varphi)$$

$$\Box^{\forall x}\varphi \quad \leftrightarrow \quad \Diamond^{x}\Box\varphi$$

$$\blacksquare^x \varphi \quad \leftrightarrow \quad \Diamond^x(\Box\varphi \vee \Box\neg\varphi)$$

$$\Box^{\forall x}\varphi \quad \leftrightarrow \quad \Diamond^x \Box\varphi$$

$$\Box^{\overline{x}}\varphi \quad \leftrightarrow \quad \Box^{x_1}\cdots\Box^{x_n}\varphi$$

$\Box^x(\varphi \rightarrow \psi) \rightarrow (\Box^x\varphi \rightarrow \Box^x\psi)$ is not valid.
So, $\Box^x$ is a non-normal modality.

- Taut: all axioms of propositional logic
- DISTK: $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$
- T: $\Box\varphi \to \varphi$
- 4MS: $\Box^x\varphi \to \Box\Box^x\varphi$
- 5MS: $\neg\Box^x\varphi \to \Box\neg\Box^x\varphi$
- KtoMS: $\Box(\varphi[y/x]) \to \Box^x\varphi$ (if $\varphi[y/x]$ is admissible)
- MStoK: $\Box^x\varphi \to \Box\varphi$ (if $x \notin FV(\varphi)$)
- MStoMSK: $\Box^x\varphi \to \Box^x\Box\varphi$
- KT: $\Box\top$
- MP: $\dfrac{\varphi, \varphi \to \psi}{\psi}$
- MONOMS: $\dfrac{\varphi \to \psi}{\Box^x\varphi \to \Box^x\psi}$

Theorem. (Wang) MLMSK is strongly complete over S5 models.

Theorem. (Wang) $\text{MLMSK}^{\approx}$ is strongly complete over S5 models.

Theorem. (Wang) MLMSK is undecidable over S5 models.

A. Baltag. *To Know is to Know the Value of a Variable*. AiML, 2016.

Knowing the value of a variable:
$w \models K_i x$ iff for all v, if $w \sim_i v$, then $w(x) = v(x)$

When D is finite, this is equivalent to $\bigvee_{d \in D} K_i(x = d)$

$[!\varphi]\psi$: after publicly announcing φ, ψ is true.

Completeness needs $K_i^\varphi x$ ("conditionally knowing what"), with the intuitive meaning that agent i could find the value of x if given the additional information that φ was the case.

Axiomatization for a logic that combines the operators for "knowledge that" ($K\varphi$) K, "knowledge of a value" (Kx), propositional public announcements $[!\varphi]$ and public announcements of values $[!x]$.

$$w \models K_i^{x_1,\ldots,x_n} y \text{ iff for all } v \sim_i w (\text{if } w(\overline{x}) = v(\overline{x}), \text{ then } w(y) = v(y))$$

Constants: $(x = c) \rightarrow (K_i x \leftrightarrow K_i(x = c))$

When the value of x is c, then knowing the value of x is the same as knowing that this value is c

Constants: $(x = c) \rightarrow (K_i x \leftrightarrow K_i(x = c))$

When the value of x is c, then knowing the value of x is the same as knowing that this value is c

Fluctuating variables: $?_\varphi$ stores the truth value of formula φ. Terms of the form $?_\varphi$ are even "more non-rigid" than the generic variables x, in that they can change their value while this value is being learnt: while x keeps its value when that value is publicly announced, terms $?_\varphi$ corresponding to Moore sentences (such as "$x = 0$ but you don't know it") may change their values after being learnt.

$K_i\varphi \leftrightarrow (\varphi \wedge K_i ?_\varphi)$

$$\varphi \ ::= \ p \mid R(\bar{t}) \mid \varphi \rightarrow \varphi \mid K_i^{\bar{t}} t$$

$$t \ ::= \ x \mid c \mid ?_p hi \mid f(\bar{t})$$

where $x \in Var$, $c \in Const$, $i \in \mathcal{A}$, $R \in \mathcal{R}$ (the set of predicate symbols including $=$), and $f \in \mathcal{F}$ (the set of function symbols).

$$\mathcal{M} = (W, D, \mathbf{0}, \mathbf{1}, \sim_i, [\![\bullet]\!], \bullet(\bullet), \mathbf{f}, \mathbf{R})_{i \in A, f \in \mathcal{F}, R \in \mathcal{R}}$$

- W is a nonempty set of worlds
- D is a nonempty domain with $\mathbf{0}, \mathbf{1} \in D$ and $\mathbf{0} \neq \mathbf{1}$
- $\sim_i$ are equivalence relations
- $\sim_i$ are equivalence relations
- $[\![\bullet]\!]$ maps atomic propositions to sets of worlds
- $\bullet(\bullet) : W \times (Var \cup Const) \to D$
- For each $f \in \mathcal{F}$, $\mathbf{f} : D^n \to D$
- For each $R \in \mathcal{R}$, $\mathbf{R} \subseteq D^n$

$$\mathcal{M} = (W, D, \mathbf{0}, \mathbf{1}, \sim_i, [\![\bullet]\!], \bullet(\bullet), \mathbf{f}, \mathbf{R})_{i \in A, f \in \mathcal{F}, R \in \mathcal{R}}$$

- $[\![R(\bar{t})]\!]_{\mathcal{M}} = \{w \mid w(\bar{t}) \in \mathbf{R}\}$
- $[\![\varphi \to \psi]\!]_{\mathcal{M}} = W \setminus [\![\varphi]\!]_{\mathcal{M}} \cup [\![\psi]\!]_{\mathcal{M}}$
- $[\![K^{\bar{t}} t]\!]_{\mathcal{M}} = \{w \mid \forall v \in W($if $w \sim_i v$ and $w(\bar{t}) = v(\bar{t})$, then $w(t) = v(t)\}$

- $w(?_{\varphi}) = \mathbf{1}$ iff $w \in [\![\varphi]\!]_{\mathcal{M}}$
- $w(?_{\varphi}) = \mathbf{0}$ iff $w \notin [\![\varphi]\!]_{\mathcal{M}}$
- $w(f(\bar{t})) = \mathbf{f}(w(\bar{t}))$

$$K_i^{\bar{t}}\varphi \;:=\; \varphi \wedge K_i^{\bar{t}}?_\varphi$$

$$\langle K_i^{\bar{t}}\rangle\varphi \;:=\; \neg K_i^{\bar{t}}\neg\varphi$$

$K_i\varphi \;:=\; K_i^{\lambda}\varphi$, where λ is the empty sequence

$$K_i^{\varphi}\psi \;:=\; K_i(\varphi \to \psi)$$

Alice and Bob have each a natural number written on their foreheads. It is common knowledge that Alice?s number x_a is the immediate successor of Bob?s number x_b. Both are blindfolded, so nobody can see the numbers.

The model has: $Var = \{x_a, x_b\}$, $D = C = \mathbb{N}$ is the set of natural numbers; $\mathcal{F} = \{+, \times\}$ and $\mathcal{R} = \{=, >\}$ contain the usual operations and relations on $\mathbb{N}$; the set W of worlds consists of all functions $w : Var \to \mathbb{N}$ satisfying $w(x_a) \to w(x_b) + 1$; the epistemic relations are given by the universal relations: $\sim_a = \sim_b = W \times W$.

$$\neg K_a x_a \wedge \neg K_b x_b \wedge K_a(x_a > x_b) \wedge K_b(x_a > x_b) \wedge K_a^{x_b} x_a \wedge K_b^{x_a} x_b$$

is true in all worlds.

So nobody knows his/her number, but both know that Alice?s number is larger, and both could come to know the numbers if given only the other?s number.

- Propositional substitution: From φ infer $\varphi[p/\theta]$
- Variable substitution: From φ infer $\varphi[x/t]$
- Modus Ponens: From φ and $\varphi \rightarrow \psi$ infer ψ
- Necessitation: From φ infer $K_i\varphi$
- Existence-of-Value Rule (EVR): From $(x = c) \rightarrow \varphi$, infer φ, provided that c does not occur in φ

- All classical propositional tautologies
- All S5 axioms for K_i
- Knowedge De Re:

$$(\overline{x} = \overline{c} \wedge y = d) \rightarrow (K_i^{\overline{x}} y \leftrightarrow K_i^{\overline{x}=\overline{c}} y = d)$$

- Equality Axioms $x = x$ $(x = y) \rightarrow (y = x)$
  $((x = y) \wedge (y = z)) \rightarrow (x = z)$ $(\overline{x} = \overline{y}) \rightarrow f(\overline{x}) = f(\overline{x})$
  $(x = y \wedge R(\overline{z}, x, \overline{y})) \rightarrow R(\overline{z}, y, \overline{y})$
- Characteristic Functions:
  $?_\varphi = 1 \leftrightarrow \varphi$
  $?_\varphi = 0 \leftrightarrow \neg\varphi$
- Knowledge of Functions: $K_i^{\overline{x}} f(\overline{x})$

Theorem (Baltag) The logic is sound and strongly complete

Theorem (Baltag) The logic has the finite model property (and is decidable)

$$
\begin{aligned}
\varphi &::= p \mid R(\bar{t}) \mid \varphi \rightarrow \varphi \mid K_i^{\bar{t}} t \mid \langle ! \bar{t} \rangle \varphi \\
t &::= x \mid c \mid ?_p hi \mid f(\bar{t}) \mid \langle ! \bar{t} \rangle t
\end{aligned}
$$

$$\varphi \ ::= \ p \mid R(\bar{t}) \mid \varphi \to \varphi \mid K_i^{\bar{t}} t \mid \langle ! \bar{t} \rangle \varphi$$
$$t \ ::= \ x \mid c \mid ?_p hi \mid f(\bar{t}) \mid \langle ! \bar{t} \rangle t$$

$$[\![\langle ! \bar{t} \rangle \varphi]\!]_{\mathcal{M}} = [\![\varphi]\!]_{\mathcal{M}^{\bar{t}}}$$

$$w(\langle ! \bar{t} \rangle t')_{\mathcal{M}} = w(t')_{\mathcal{M}^{\bar{t}}}$$

$$\varphi ::= p \mid R(\bar{t}) \mid \varphi \rightarrow \varphi \mid K_i^{\bar{t}} t \mid \langle ! \bar{t} \rangle \varphi$$
$$t ::= x \mid c \mid ?_p hi \mid f(\bar{t}) \mid \langle ! \bar{t} \rangle t$$

$$[\![\langle ! \bar{t} \rangle \varphi]\!]_{\mathcal{M}} = [\![\varphi]\!]_{\mathcal{M}^{\bar{t}}}$$

$$w(\langle ! \bar{t} \rangle t')_{\mathcal{M}} = w(t')_{\mathcal{M}^{\bar{t}}}$$

$$\mathcal{M}^{\bar{t}} = (W, D, \mathbf{0}, \mathbf{1}, \sim_i^{\bar{t}}, [\![\bullet]\!], \bullet(\bullet), \mathbf{f}, \mathbf{R})_{i \in A, f \in \mathcal{F}, R \in \mathcal{R}}$$

$$\sim_i^{\mathcal{M}^{\bar{t}}} = \{(w, s) \in W \times W \mid w \sim_i s, w(\bar{t})_{\mathcal{M}} = s(\bar{t})_{\mathcal{M}}\}$$

In the previous example, $\langle !x_a \rangle (K_a x_b \wedge K_b x_b)$ is true at all worlds.

$$\langle !\bar{t}\rangle p \leftrightarrow p$$

$$\langle !\bar{t}\rangle R(t_1,\ldots,t_n) \leftrightarrow R(\langle !\bar{t}\rangle t_1,\ldots,\langle !\bar{t}\rangle t_n)$$

$$\langle !\bar{t}\rangle(\varphi \rightarrow \psi) \leftrightarrow (\langle !\bar{t}\rangle\varphi \rightarrow \langle !\bar{t}\rangle\psi)$$

$$\langle !\bar{t}\rangle K_i^{t_1,\ldots,t_n} t' \leftrightarrow K_i^{\langle !\bar{t}\rangle t_1,\ldots,\langle !\bar{t}\rangle t_n}\langle !\bar{t}\rangle t'$$

$$\langle !\bar{t}\rangle p \leftrightarrow p$$

$$\langle !\bar{t}\rangle R(t_1, \ldots, t_n) \leftrightarrow R(\langle !\bar{t}\rangle t_1, \ldots, \langle !\bar{t}\rangle t_n)$$

$$\langle !\bar{t}\rangle(\varphi \to \psi) \leftrightarrow (\langle !\bar{t}\rangle\varphi \to \langle !\bar{t}\rangle\psi)$$

$$\langle !\bar{t}\rangle K_i^{t_1,\ldots,t_n} t' \leftrightarrow K_i^{\langle !\bar{t}\rangle t_1,\ldots,\langle !\bar{t}\rangle t_n}\langle !\bar{t}\rangle t'$$

$$\langle !\bar{t}\rangle c = c$$

$$\langle !\bar{t}\rangle x = x$$

$$\langle !\bar{t}\rangle ?_\varphi = ?_{\langle !\bar{t}\rangle\varphi}$$

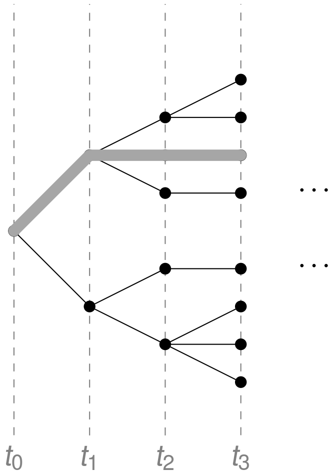
$$\langle !\bar{t}\rangle f(t_1, \ldots, t_n) = f(\langle !\bar{t}\rangle t_1, \ldots, \langle !\bar{t}\rangle t_n)$$

Theorem (Baltag) The above proof system is sound and weakly complete and has the same expressivity as LED.

Epistemizing logics of action and ability

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 stit*]φ which is intended to mean that the agent *i* can "see to it that φ is true".
  - [*i stit*]φ 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.

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

STIT Model

A STIT models is $\mathcal{M} = \langle T, <, Choice, V \rangle$ 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.
- $Choice : \mathcal{A} \times T \rightarrow \wp(\wp(H))$ 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 $s_t : \mathcal{A} \rightarrow \wp(H_t)$ 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.
- $Choice : \mathcal{A} \times T \rightarrow \wp(\wp(H))$ 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 $s_t : \mathcal{A} \rightarrow \wp(H_t)$ such that $s_t(i) \in Choice_i^t$, we have $\bigcap_{i \in \mathcal{A}} s_t(i) \neq \emptyset$
- $V : At \rightarrow \wp(T \times Hist)$ is a valuation function assigning to each atomic proposition

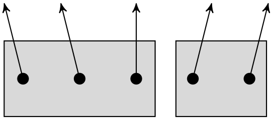
STIT Model

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

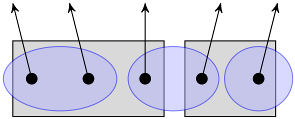
The previous model assumes there is *one* agent that “controls” the transition system.



Many Agents

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

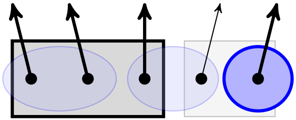
What if there is more than one agent?



Many Agents

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

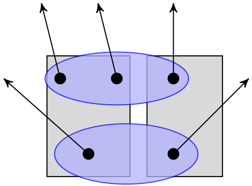
What if there is more than one agent? *Independence of agents*



Many Agents

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

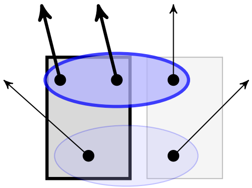
What if there is more than one agent? *Independence of agents*



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

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

- $\mathcal{M}, t/h \models p$ 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 $h' \in Choice_i^t(h)$

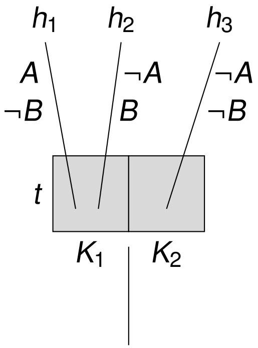
STIT Language

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

- $\mathcal{M}, t/h \models p$ 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 $\Box$: $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$, $\Box\varphi \rightarrow \varphi$, $\Box\varphi \rightarrow \Box\Box\varphi$, $\neg\Box\varphi \rightarrow \Box\neg\Box\varphi$

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

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

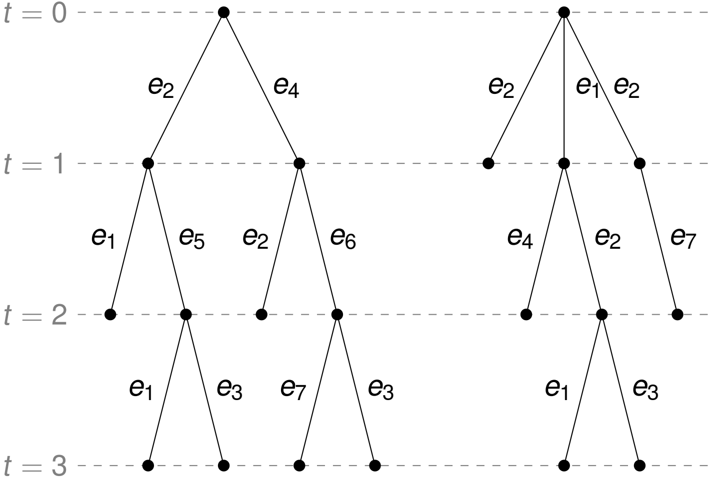
- $F\varphi$: φ is true at some moment in *the* future
- $\exists F\varphi$: there is a history where φ is true some moment in the future
- $[\alpha]\varphi$: after doing action α, φ is true
- $[\delta\varphi]\psi$: after bringing about φ, ψ is true
- $[i\ \mathit{stit}]\varphi$: the agent can "see to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

Epistemic Temporal Logic

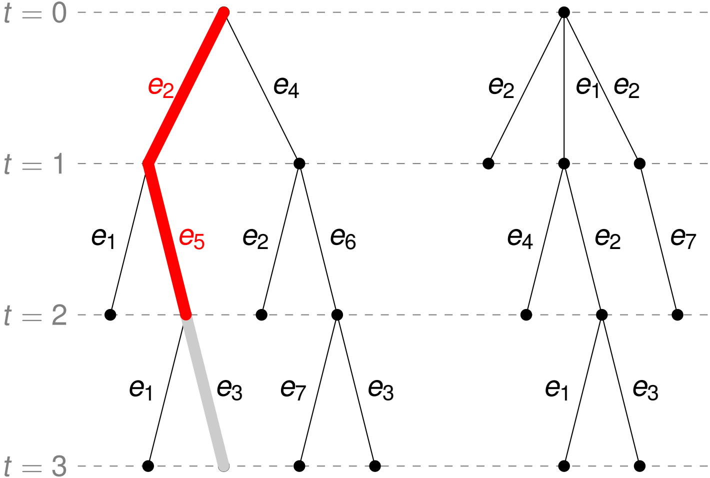
R. Parikh and R. Ramanujam. *A Knowledge Based Semantics of Messages*. *Journal of Logic, Language and Information*, 12: 453 – 467, 1985, 2003.

FHMV. Reasoning about Knowledge. MIT Press, 1995.

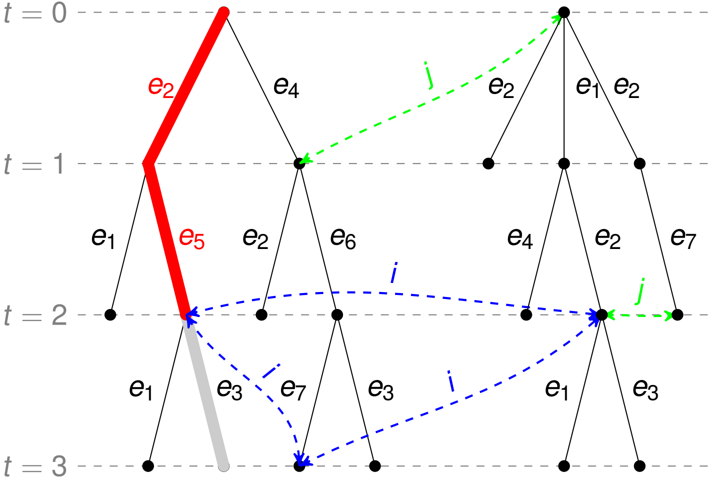
The ‘Playground’



The 'Playground'



The ‘Playground’



The ‘Playground’: Notation

The 'Playground': Notation

- Let Σ be any set. The elements of Σ are called **events.**

The ‘Playground’: Notation

- Let Σ be any set. The elements of Σ are called **events.**
- Given any set X, X^* is the set of finite strings over X and X^ω the set of infinite strings over X. Elements of $\Sigma^* \cup \Sigma^\omega$ will be called **histories.**

The 'Playground': Notation

- Let Σ be any set. The elements of Σ are called **events.**
- Given any set X, X^* is the set of finite strings over X and X^ω the set of infinite strings over X. Elements of $\Sigma^* \cup \Sigma^\omega$ will be called **histories.**
- Given $H \in \Sigma^* \cup \Sigma^\omega$, $\text{len}(H)$ is the **length** of H.

The ‘Playground’: Notation

- Let Σ be any set. The elements of Σ are called **events.**

- Given any set X, X^* is the set of finite strings over X and X^ω the set of infinite strings over X. Elements of $\Sigma^* \cup \Sigma^\omega$ will be called **histories.**

- Given $H \in \Sigma^* \cup \Sigma^\omega$, $\mathrm{len}(H)$ is the **length** of H.

- Given $H, H' \in \Sigma^* \cup \Sigma^\omega$, we write $H \preceq H'$ if H is a *finite* prefix of H'.

The 'Playground': Notation

- Let Σ be any set. The elements of Σ are called **events.**
- Given any set X, X^* is the set of finite strings over X and X^ω the set of infinite strings over X. Elements of $\Sigma^* \cup \Sigma^\omega$ will be called **histories.**
- Given $H \in \Sigma^* \cup \Sigma^\omega$, $\mathrm{len}(H)$ is the **length** of H.
- Given $H, H' \in \Sigma^* \cup \Sigma^\omega$, we write $H \preceq H'$ if H is a *finite* prefix of H'.
- $\mathrm{FinPre}(\mathcal{H}) = \{H \mid \exists H' \in \mathcal{H} \text{ such that } H \preceq H'\}$ is the set of finite prefixes of the elements of $\mathcal{H}$.

The 'Playground': Notation

- Let Σ be any set. The elements of Σ are called **events.**
- Given any set X, X^* is the set of finite strings over X and X^ω the set of infinite strings over X. Elements of $\Sigma^* \cup \Sigma^\omega$ will be called **histories.**
- Given $H \in \Sigma^* \cup \Sigma^\omega$, $\text{len}(H)$ is the **length** of H.
- Given $H, H' \in \Sigma^* \cup \Sigma^\omega$, we write $H \preceq H'$ if H is a *finite* prefix of H'.
- $\text{FinPre}(\mathcal{H}) = \{H \mid \exists H' \in \mathcal{H} \text{ such that } H \preceq H'\}$ is the set of finite prefixes of the elements of $\mathcal{H}$.
- ϵ is the empty string and $\text{FinPre}_{-\epsilon}(\mathcal{H}) = \text{FinPre}(\mathcal{H}) - \{\epsilon\}$.

History-based Frames

Definition

Let Σ be any set of events. A set $\mathcal{H} \subseteq \Sigma^* \cup \Sigma^\omega$ is called a protocol provided $\text{FinPre}_{-\epsilon}(\mathcal{H}) \subseteq \mathcal{H}$. A rooted protocol is any set $\mathcal{H} \subseteq \Sigma^* \cup \Sigma^\omega$ where $\text{FinPre}(\mathcal{H}) \subseteq \mathcal{H}$.

History-based Frames

Definition

Let Σ be any set of events. A set $\mathcal{H} \subseteq \Sigma^* \cup \Sigma^\omega$ is called a protocol provided $\text{FinPre}_{-\epsilon}(\mathcal{H}) \subseteq \mathcal{H}$. A rooted protocol is any set $\mathcal{H} \subseteq \Sigma^* \cup \Sigma^\omega$ where $\text{FinPre}(\mathcal{H}) \subseteq \mathcal{H}$.

Definition

An ETL frame is a tuple $\langle \Sigma, \mathcal{H}, \{\sim_i\}_{i \in \mathcal{A}} \rangle$ where Σ is a (finite or infinite) set of events, $\mathcal{H}$ is a protocol, and for each $i \in \mathcal{A}$, $\sim_i$ is an equivalence relation on the set of finite strings in $\mathcal{H}$.

History-based Frames

Definition

Let Σ be any set of events. A set $\mathcal{H} \subseteq \Sigma^* \cup \Sigma^\omega$ is called a protocol provided $\text{FinPre}_{-\epsilon}(\mathcal{H}) \subseteq \mathcal{H}$. A rooted protocol is any set $\mathcal{H} \subseteq \Sigma^* \cup \Sigma^\omega$ where $\text{FinPre}(\mathcal{H}) \subseteq \mathcal{H}$.

Definition

An ETL frame is a tuple $\langle \Sigma, \mathcal{H}, \{\sim_i\}_{i \in \mathcal{A}} \rangle$ where Σ is a (finite or infinite) set of events, $\mathcal{H}$ is a protocol, and for each $i \in \mathcal{A}$, $\sim_i$ is an equivalence relation on the set of finite strings in $\mathcal{H}$.

Some assumptions:

1. If Σ is assumed to be finite, then we say that $\mathcal{F}$ is **finitely branching.**
2. If $\mathcal{H}$ is a rooted protocol, $\mathcal{F}$ is a **tree frame.**

Formal Languages

- $P\varphi$ (φ is true *sometime* in the past),
- $F\varphi$ (φ is true *sometime* in the future),
- $Y\varphi$ (φ is true at *the* previous moment),
- $N\varphi$ (φ is true at *the* next moment),
- $N_e\varphi$ (φ is true after event e)
- $K_i\varphi$ (agent i knows φ) and
- $C_B\varphi$ (the group $B \subseteq \mathcal{A}$ commonly knows φ).

History-based Models

An ETL **model** is a structure $\langle \mathcal{H}, \{\sim_i\}_{i\in\mathcal{A}}, V\rangle$ where $\langle \mathcal{H}, \{\sim_i\}_{i\in\mathcal{A}}\rangle$ is an ETL frame and

$V : \mathsf{At} \to 2^{\mathsf{finite}(\mathcal{H})}$ is a valuation function.

Formulas are interpreted at pairs H, t:

$$H, t \models \varphi$$

Truth in a Model

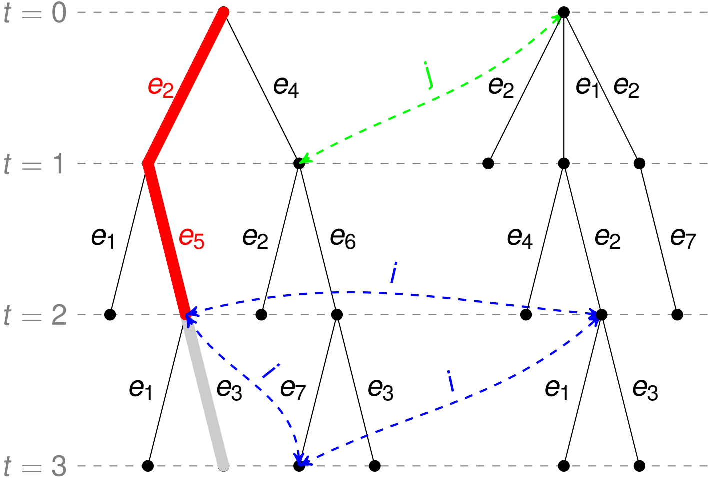
- $H, t \models P\varphi$ iff there exists $t' \leq t$ such that $H, t' \models \varphi$
- $H, t \models F\varphi$ iff there exists $t' \geq t$ such that $H, t' \models \varphi$
- $H, t \models N\varphi$ iff $H, t + 1 \models \varphi$
- $H, t \models Y\varphi$ iff $t > 1$ and $H, t - 1 \models \varphi$
- $H, t \models K_i\varphi$ iff for each $H' \in \mathcal{H}$ and $m \geq 0$ if $H_t \sim_i H'_m$ then $H', m \models \varphi$
- $H, t \models C\varphi$ iff for each $H' \in \mathcal{H}$ and $m \geq 0$ if $H_t \sim_* H'_m$ then $H', m \models \varphi$.

where $\sim_*$ is the reflexive transitive closure of the union of the $\sim_i$.

Truth in a Model

- $H, t \models P\varphi$ iff there exists $t' \leq t$ such that $H, t' \models \varphi$
- $H, t \models F\varphi$ iff there exists $t' \geq t$ such that $H, t' \models \varphi$
- $H, t \models N\varphi$ iff $H, t + 1 \models \varphi$
- $H, t \models Y\varphi$ iff $t > 1$ and $H, t - 1 \models \varphi$
- $H, t \models K_i\varphi$ iff for each $H' \in \mathcal{H}$ and $m \geq 0$ if $H_t \sim_i H'_m$ then $H', m \models \varphi$
- $H, t \models C\varphi$ iff for each $H' \in \mathcal{H}$ and $m \geq 0$ if $H_t \sim_* H'_m$ then $H', m \models \varphi$.

where $\sim_*$ is the reflexive transitive closure of the union of the $\sim_i$.


$t = 0$
$t = 1$
$t = 2$
$t = 3$
e_2
e_4
j
e_2
e_1
e_2
e_1
e_5
e_2
e_6
e_4
e_2
e_7
i
e_1
e_3
e_7
e_3
i
e_1
e_3

An Example

Ann would like Bob to attend her talk; however, she only wants Bob to attend if he is interested in the subject of her talk, not because he is just being polite.

An Example

Ann would like Bob to attend her talk; however, she only wants Bob to attend if he is interested in the subject of her talk, not because he is just being polite.

There is a very simple procedure to solve Ann's problem: *have a (trusted) friend tell Bob the time and subject of her talk.*

An Example

Ann would like Bob to attend her talk; however, she only wants Bob to attend if he is interested in the subject of her talk, not because he is just being polite.

There is a very simple procedure to solve Ann's problem: *have a (trusted) friend tell Bob the time and subject of her talk.*

Is this procedure correct?

An Example

Yes, if

1. Ann knows about the talk.

An Example

Yes, if

1. Ann knows about the talk.
2. Bob knows about the talk.

An Example

Yes, if

1. Ann knows about the talk.
2. Bob knows about the talk.
3. Ann knows that Bob knows about the talk.

An Example

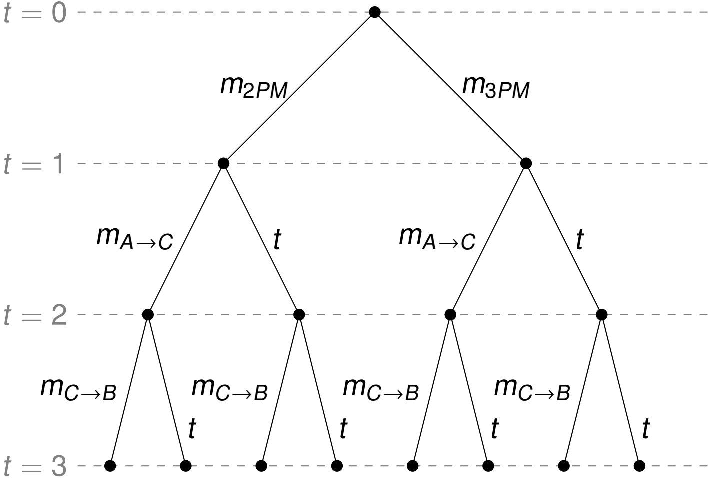
Yes, if

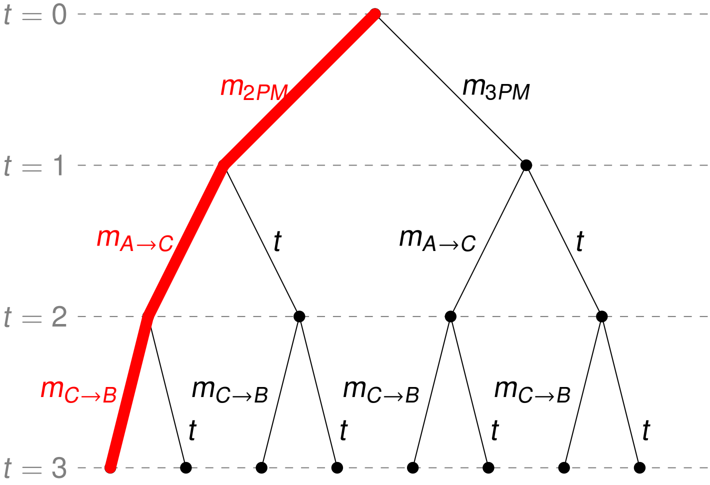
1. Ann knows about the talk.
2. Bob knows about the talk.
3. Ann knows that Bob knows about the talk.
4. Bob *does not* know that Ann knows that he knows about the talk.

An Example

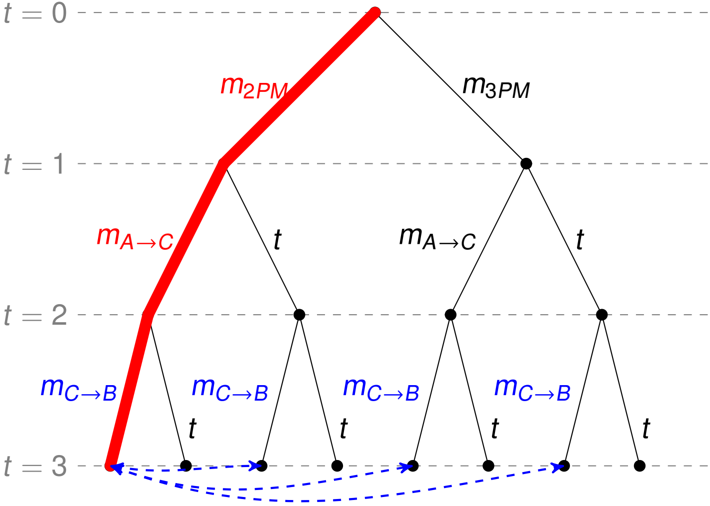
Yes, if

1. Ann knows about the talk.
2. Bob knows about the talk.
3. Ann knows that Bob knows about the talk.
4. Bob *does not* know that Ann knows that he knows about the talk.
5. *And nothing else.*

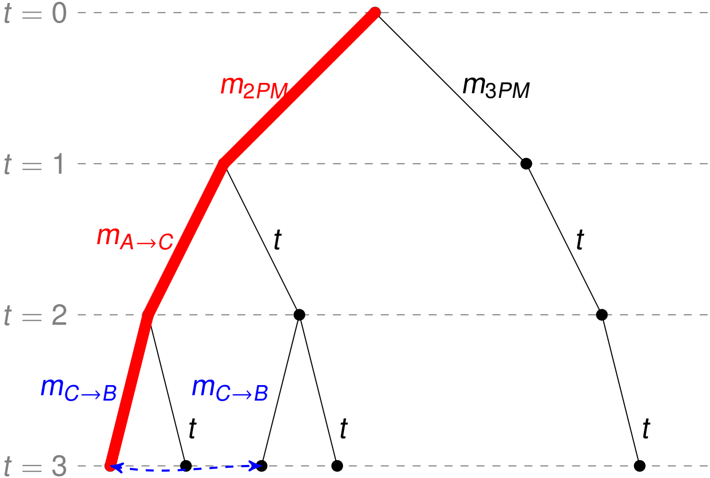

$t = 0$
$t = 1$
$t = 2$
$t = 3$
m_{2PM}
m_{3PM}
$m_{A \to C}$
t
$m_{A \to C}$
t
$m_{C \to B}$
t
$m_{C \to B}$
t
$m_{C \to B}$
t
$m_{C \to B}$
t



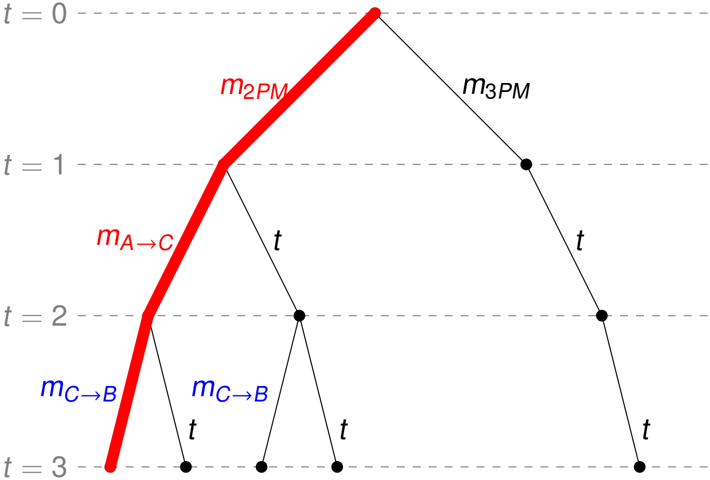
$$H, 3 \models \varphi$$



Bob's uncertainty: $H, 3 \models \neg K_B P_{2PM}$

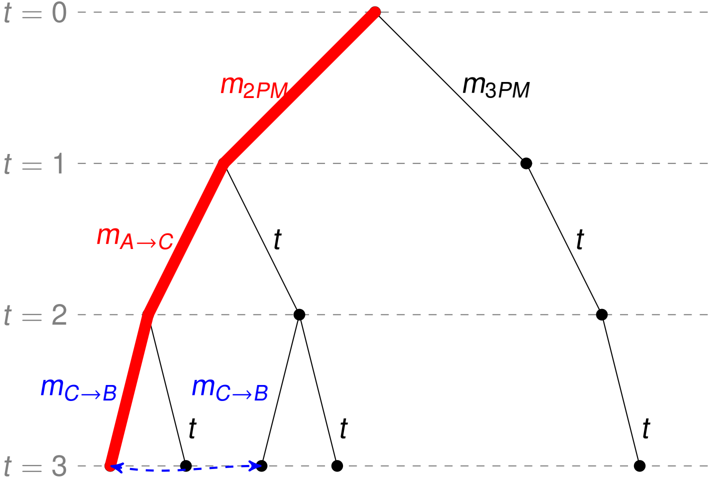


Bob's uncertainty + 'Protocol information': $H, 3 \models K_B P_{2PM}$



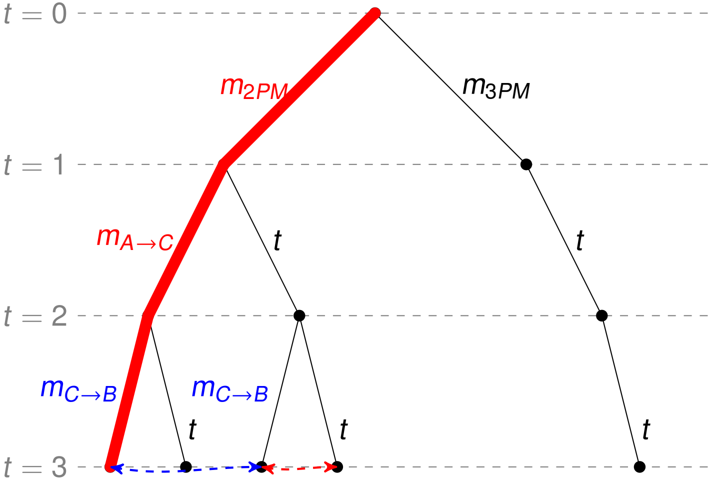
Bob's uncertainty + 'Protocol information':

$$H, 3 \models \neg K_B K_A K_B P_{2PM}$$



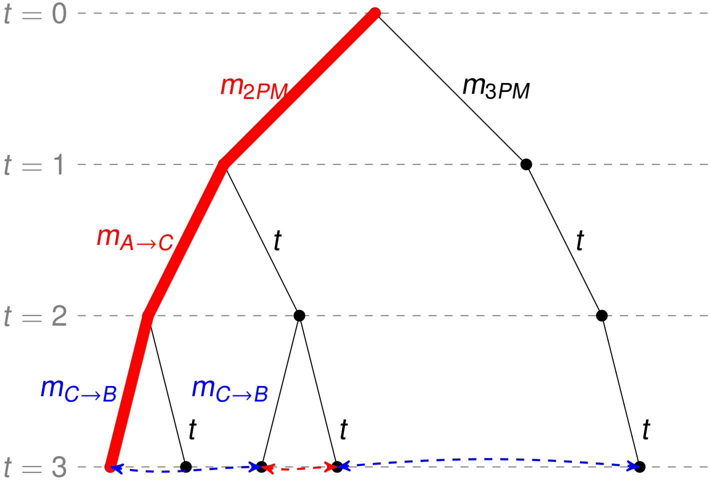
Bob's uncertainty + 'Protocol information':

$$H, 3 \models \neg K_B K_A K_B P_{2PM}$$



Bob's uncertainty + 'Protocol information':

$$H, 3 \models \neg K_B K_A K_B P_{2PM}$$



Bob's uncertainty + 'Protocol information':

$$H, 3 \models \neg K_B K_A K_B P_{2PM}$$

Living at the Edge of Decidability

Living at the Edge of Decidability

1. Expressivity of the formal language. Does the language include a common knowledge operator? A future operator? Both?

Living at the Edge of Decidability

1. Expressivity of the formal language. Does the language include a common knowledge operator? A future operator? Both?

2. Structural conditions on the underlying event structure. Do we restrict to protocol frames (finitely branching trees)? Finitely branching forests? Or, arbitrary ETL frames?

Living at the Edge of Decidability

1. Expressivity of the formal language. Does the language include a common knowledge operator? A future operator? Both?

2. Structural conditions on the underlying event structure. Do we restrict to protocol frames (finitely branching trees)? Finitely branching forests? Or, arbitrary ETL frames?

3. Conditions on the reasoning abilities of the agents. Do the agents satisfy perfect recall? No miracles? Do they agents' know what time it is?

Agent Oriented Properties:

- **No Miracles**: For all finite histories $H, H' \in \mathcal{H}$ and events $e \in \Sigma$ such that $He \in \mathcal{H}$ and $H'e \in \mathcal{H}$, if $H \sim_i H'$ then $He \sim_i H'e$.

- **Perfect Recall**: For all finite histories $H, H' \in \mathcal{H}$ and events $e \in \Sigma$ such that $He \in \mathcal{H}$ and $H'e \in \mathcal{H}$, if $He \sim_i H'e$ then $H \sim_i H'$.

- **Synchronous**: For all finite histories $H, H' \in \mathcal{H}$, if $H \sim_i H'$ then $\mathrm{len}(H) = \mathrm{len}(H')$.

Decidability in the Purely Temporal Setting

Theorem (Rabin)

The satisfiable problem for monadic second-order logic of the k-ary tree is decidable.

M. O. Rabin. *Decidability of Second-Order Theories and Automata on Infinite Trees*. *Transactions of the American Mathematical Society*, 141, 1969.

Theorem

The satisfiability problem for $\mathcal{L}_{TL}$ with respect to TL tree models (without epistemic structure) is decidable.

Arbitrary Agents

Theorem

The satisfiability problem (with respect to a language $\mathcal{L}_{ETL}$ with C,F, etc.) is decidable — EXPTIME-complete).

Arbitrary Agents

Theorem

The satisfiability problem (with respect to a language $\mathcal{L}_{ETL}$ with C,F, etc.) is decidable — EXPTIME-complete).

- The theorem holds if we restrict to tree models.

Ideal Agents

Assume there are two agents

Theorem

The satisfiability problem for $\mathcal{L}_{ETL}$ is highly undecidable under certain idealizations.

Ideal Agents

Assume there are two agents

Theorem

The satisfiability problem for $\mathcal{L}_{ETL}$ is highly undecidable under certain idealizations.

For example,

Theorem (Halpern & Vardi)

On interpreted systems that satisfy perfect recall or no learning, the satisfiability problem for $\mathcal{L}_{ETL}$ is Σ_1^1-complete.

(no learning: For $H, H' \in \mathcal{H}$, if $H_t \sim_i H'_{t'}$ then for all $k \geq t$ there exists $k' \geq t'$ such that $H_k \sim_i H'_{k'}$.)

J. Halpern and M. Vardi.. *The Complexity of Reasoning abut Knowledge and Time*. *J. Computer and Systems Sciences*, 38, 1989.

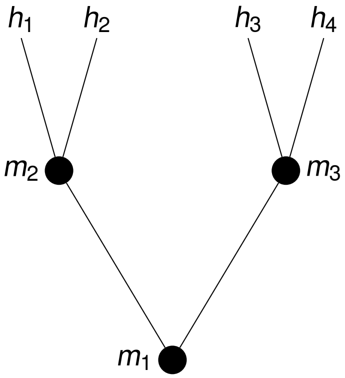
J. Horty and EP. *Action Types in Stit Semantics*. Review of Symbolic Logic, 2017.

Stit model

$$\langle Tree, <, Agent, Choice, V \rangle$$

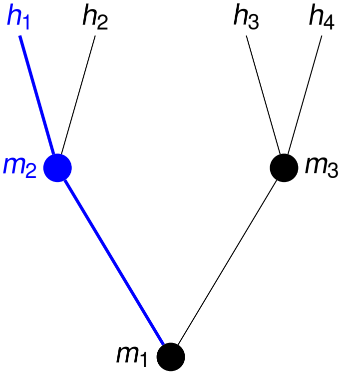
Stit model

$$\langle Tree, <, Agent, Choice, V\rangle$$



Stit model

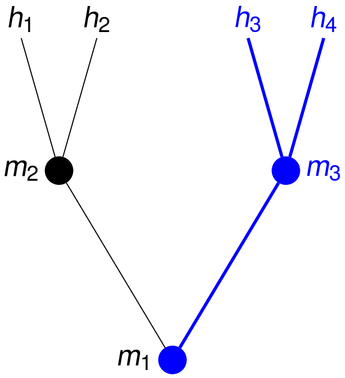
$$\langle Tree, <, Agent, Choice, V\rangle$$



m/h denotes (m, h) with $m \in h$ is called an **index**

Stit model

$$\langle Tree, <, Agent, Choice, V\rangle$$

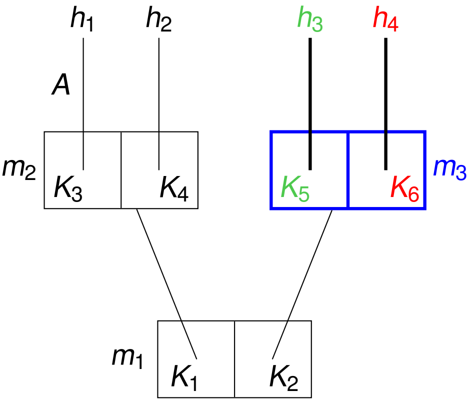


m/h denotes (m,h) with $m \in h$ is called an **index**

$$H^m = \{h \mid m \in h\}$$

Stit model

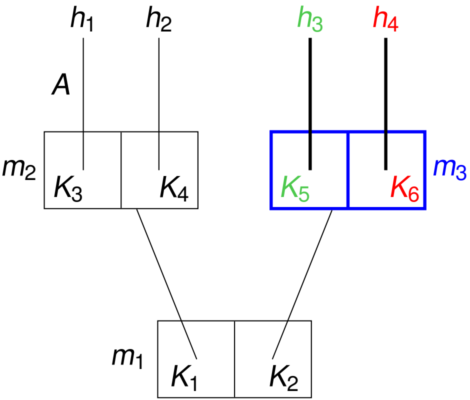
$$\langle Tree, <, Agent, Choice, V \rangle$$



For $\alpha \in Agent$, $Choice^m_\alpha$ is a partition on H^m

Stit model

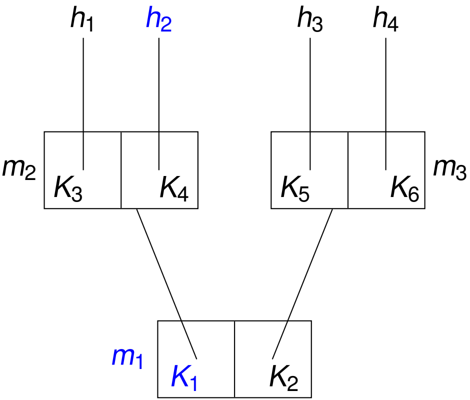
$$\langle \mathit{Tree}, <, \mathit{Agent}, \mathit{Choice}, V \rangle$$



For $\alpha \in \mathit{Agent}$, Choice^m_α is a partition on H^m

Stit model

$$\langle Tree, <, Agent, Choice, V \rangle$$

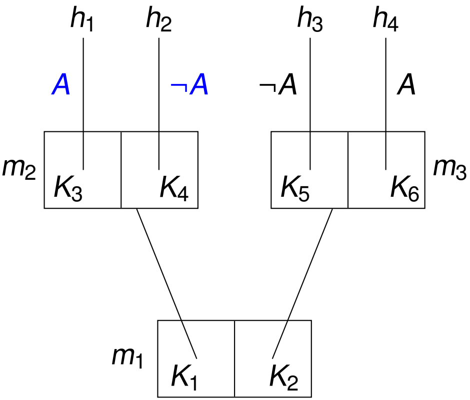


For $\alpha \in Agent$, $Choice^m_\alpha$ is a partition on H^m

$Choice^m_\alpha(h)$ is the particular action at m that contains h

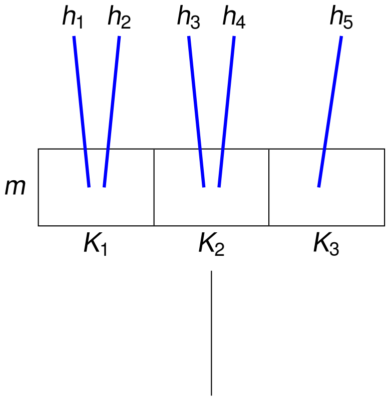
Stit model

$$\langle Tree, <, Agent, Choice, V \rangle$$

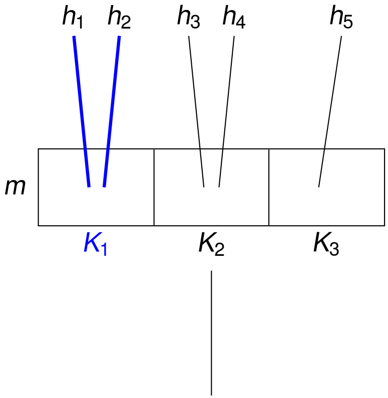


V assigns sets of indices to atomic propositions.

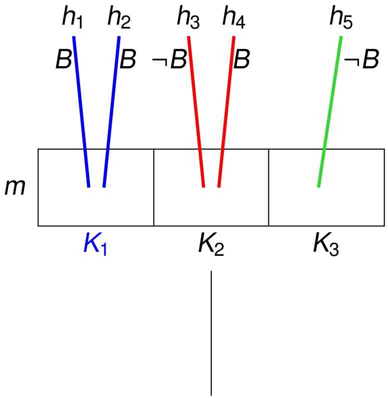
$m_2/h_1 \models A$ $\qquad$ $m_2/h_2 \not\models A$



- $\mathcal{M}, m/h \models \Box A$ if and only if $\mathcal{M}, m/h' \models A$ for all $h' \in H^m$,

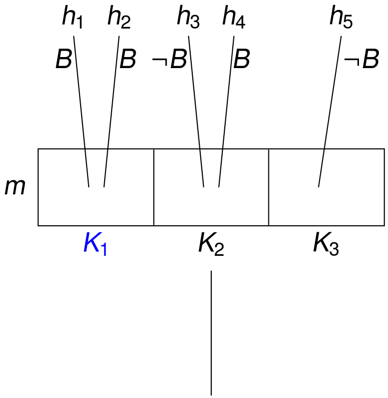


- $\mathcal{M}, m/h \models \Box A$ if and only if $\mathcal{M}, m/h' \models A$,
- $\mathcal{M}, m/h \models [\alpha\ stit\colon A]$ if and only if $Choice^m_\alpha(h) \subseteq |A|^m_\mathcal{M}$



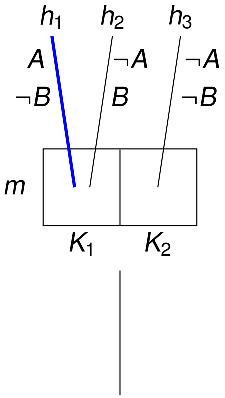
- $\mathcal{M}, m/h \models \Box A$ if and only if $\mathcal{M}, m/h' \models A$,
- $\mathcal{M}, m/h \models [\alpha\ stit\colon A]$ if and only if $Choice^m_\alpha(h) \subseteq |A|^m_\mathcal{M}$

  $m/h_1 \models [\alpha\ stit\colon B]$, $m/h_3 \not\models [\alpha\ stit\colon B]$, $m/h_5 \models [\alpha\ stit\colon \neg B]$



- $\mathcal{M}, m/h \models \Box A$ if and only if $\mathcal{M}, m/h' \models A$,
- $\mathcal{M}, m/h \models [\alpha\ stit\colon A]$ if and only if $Choice^m_\alpha(h) \subseteq |A|^m_{\mathcal{M}}$
- Temporal modalities (P, F, ...)

Ability: $\Diamond[\alpha\ stit: A]$



- $m/h_1 \not\models A \supset \Diamond[\alpha\ stit: A]$
- $m/h_1 \not\models \Diamond[\alpha\ stit: A \vee B] \supset \Diamond[\alpha\ stit: A] \vee \Diamond[\alpha\ stit: B]$

$\Diamond[\alpha$ *stit*: $A]$ is a "causal sense" of ability. But, there is also an "epistemic sense" of ability...

$\Diamond[\alpha$ *stit*: $A]$ is a "causal sense" of ability. But, there is also an "epistemic sense" of ability...

What needs to be added to stit models?

- Indistinguishability relation(s)

$\Diamond[\alpha\ \mathit{stit}\colon A]$ is a “causal sense” of ability. But, there is also an “epistemic sense” of ability...

What needs to be added to stit models?

- Indistinguishability relation(s)
- Action types

Epistemic stit models

A. Herzig. *Logics of knowledge and action: critical analysis and challenges*. Autonomous Agent and Multi-Agent Systems, 2014.

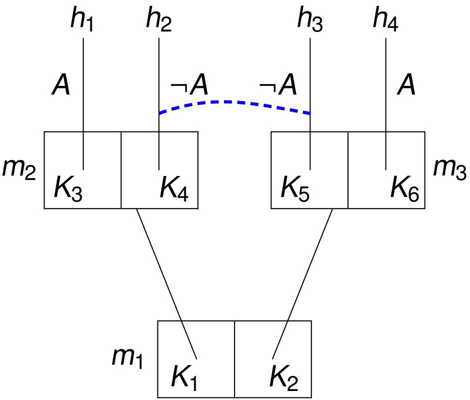
V. Goranko and EP. *Temporal aspects of the dynamics of knowledge*. in Johan van Benthem on Logic and Information Dynamics, Outstanding Contributions to Logic, (eds. Alexandru Baltag and Sonja Smets), pp. 235 - 266, 2014.

J. Broeresen, A. Herzig and N. Troquard. *What groups do, can do and know they can do: An analysis in normal modal logics*. Journal of Applied and Non-Classical Logics, 19:3, pgs. 261 - 289, 2009.

W. van der Hoek and M. Wooldridge. *Cooperation, knowledge and time: Alternating-time temporal epistemic logic and its applications*. Studia Logica, 75, pgs. 125 - 157, 2003.

Epistemic stit models

$$\langle Tree, <, Agent, Choice, \{\sim_\alpha\}_{\alpha \in Agent}, V \rangle$$

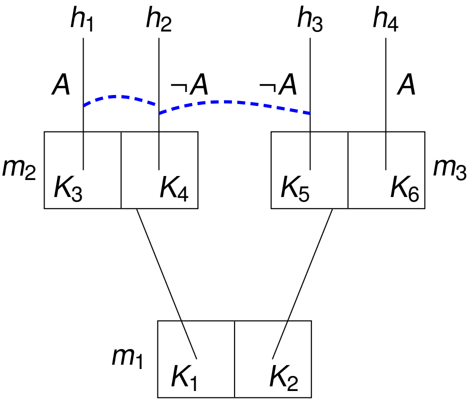


$\sim_\alpha$ is an equivalence relation on indices

$m/h \sim_\alpha m'/h'$: nothing α knows distinguishes m/h from m'/h', or m/h and m'/h' are indistinguishable

Epistemic stit models

$$\langle Tree, <, Agent, Choice, \{\sim_\alpha\}_{\alpha \in Agent}, V\rangle$$

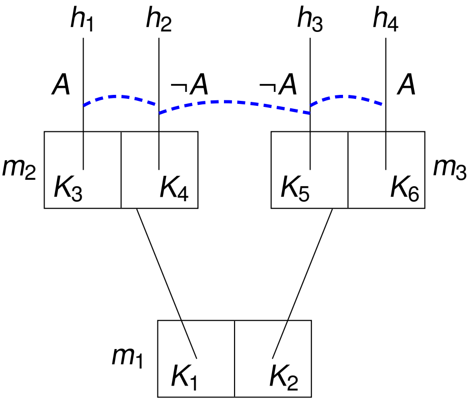


$\sim_\alpha$ is an equivalence relation on indices

$m/h \sim_\alpha m'/h'$: nothing α knows distinguishes m/h from m'/h', or m/h and m'/h' are indistinguishable

Epistemic stit models

$$\langle Tree, <, Agent, Choice, \{\sim_\alpha\}_{\alpha \in Agent}, V \rangle$$

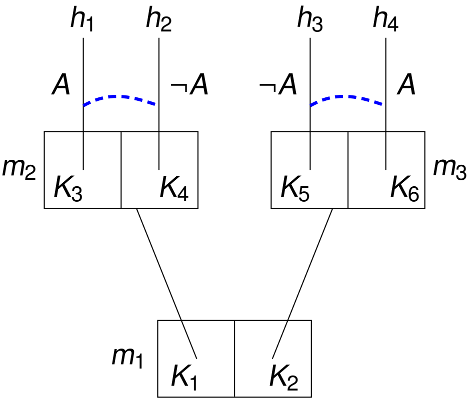


$\sim_\alpha$ is an equivalence relation on indices

$m/h \sim_\alpha m'/h'$: nothing α knows distinguishes m/h from m'/h', or m/h and m'/h' are indistinguishable

Epistemic stit models

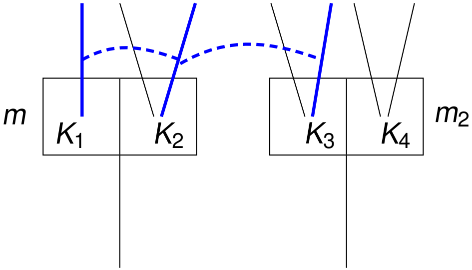
$$\langle Tree, <, Agent, Choice, \{\sim_\alpha\}_{\alpha \in Agent}, V \rangle$$



$\sim_\alpha$ is an equivalence relation on indices

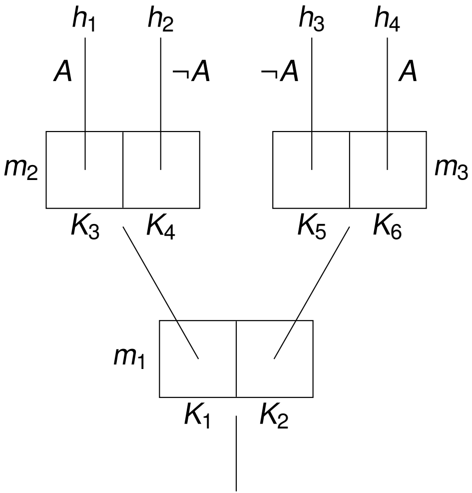
$m/h \sim_\alpha m'/h'$: nothing α knows distinguishes m/h from m'/h', or m/h and m'/h' are indistinguishable

Epistemic stit models

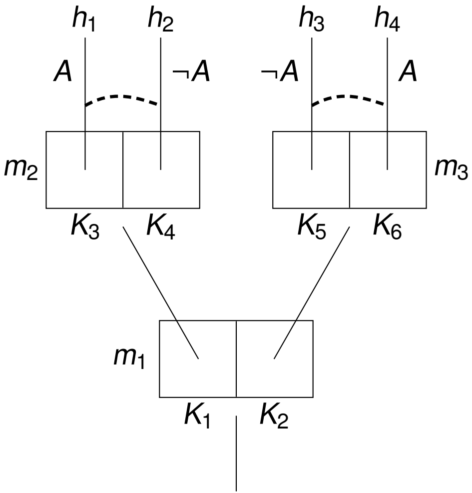


- $\mathcal{M}, m/h \models \mathsf{K}_\alpha A$ if and only if, for all m'/h', if $m/h \sim_\alpha m'/h'$, then $\mathcal{M}, m'/h' \models A$

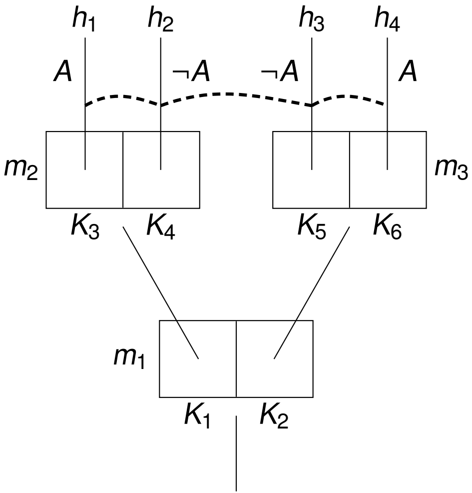
Coin game



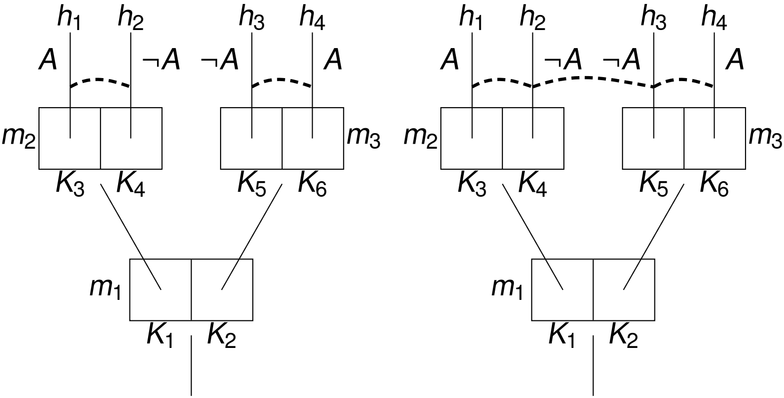
Coin game 1



Coin game 2

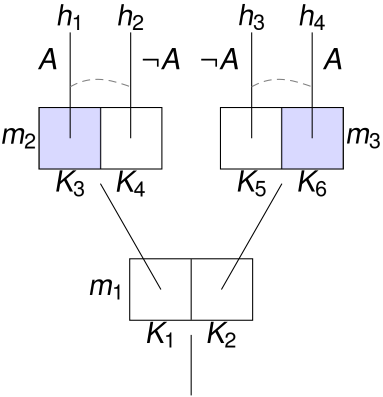


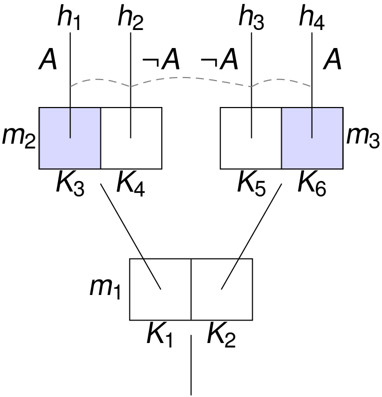
Ability



h_1
h_2
h_3
h_4
A
$\neg A$
$\neg A$
A
m_2
m_3
K_3
K_4
K_5
K_6
m_1
K_1
K_2
h_1
h_2
h_3
h_4
A
$\neg A$
$\neg A$
A
m_2
m_3
K_3
K_4
K_5
K_6
m_1
K_1
K_2

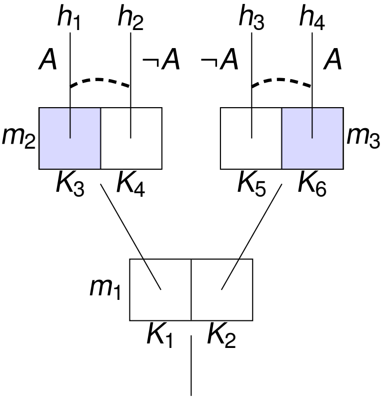
Ability

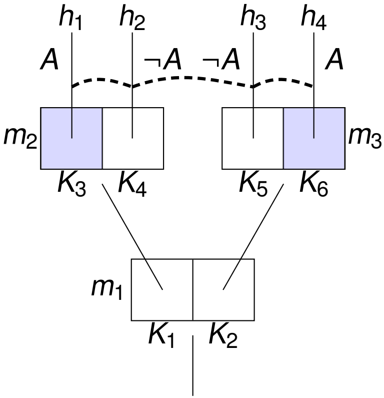




$\Diamond[\alpha$ *stit*: $A]$ is settled true in at m_2 and m_3 in both models.

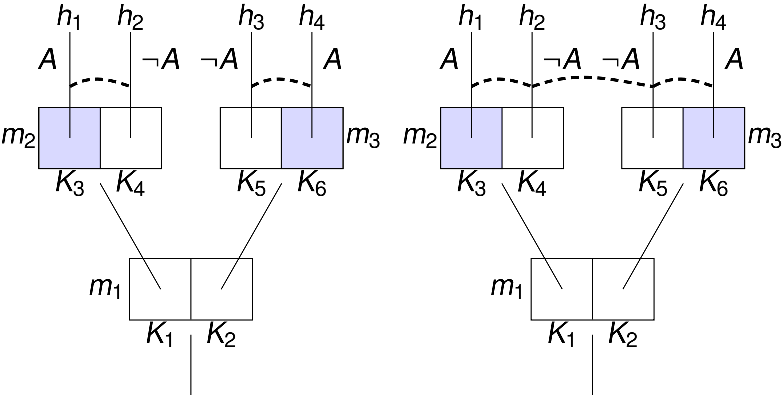
Ability





$\mathsf{K}_\alpha\Diamond[\alpha\ \mathit{stit}\colon A]$ is settled true in at m_2 and m_3 in both models.

Ability

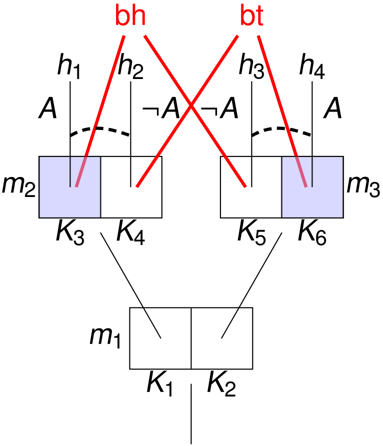


$\Diamond K_\alpha[\alpha$ *stit*: $A]$ is settled false in at m_2 and m_3 in both models.

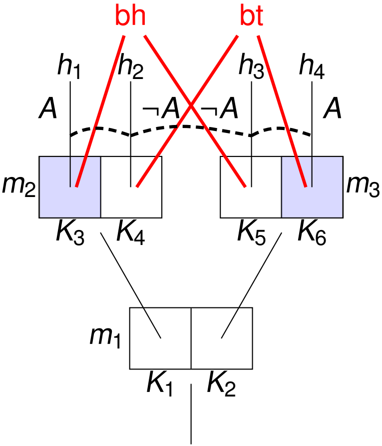
Ability

α has the ability to see to it that A in the epistemic sense just in case there is some action available to α that is known by α to guarantee the truth of A.

Ability

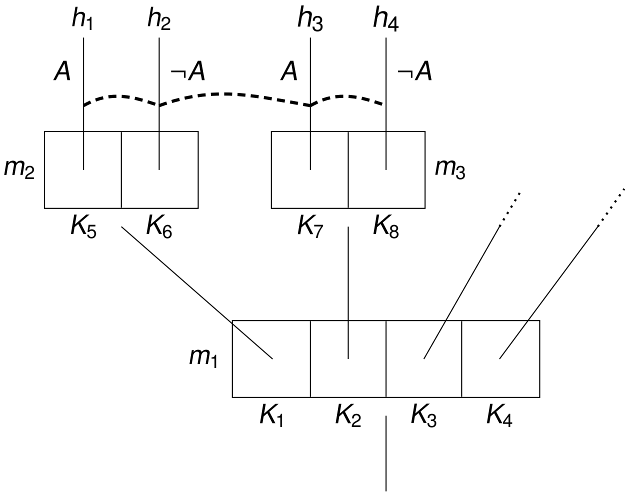


bh
bt
h_1
h_2
h_3
h_4
A
$\neg A$
$\neg A$
A
m_2
m_3
K_3
K_4
K_5
K_6
m_1
K_1
K_2



bh
bt
h_1
h_2
h_3
h_4
A
$\neg A$
$\neg A$
A
m_2
m_3
K_3
K_4
K_5
K_6
m_1
K_1
K_2

Coin game 3



Labeled stit model

$$\langle \mathit{Tree}, <, \mathit{Agent}, \mathit{Choice}, \{\sim_\alpha\}_{\alpha \in \mathit{Agent}}, \mathit{Type}, [\ ], \mathit{Label}, V \rangle$$

$\mathit{Type} = \{\tau_1, \tau_2, \ldots\}$ is a finite set of action types—general kinds of action, as opposed to the concrete action tokens already present in stit logics.

Labeled stit model

$$\langle Tree, <, Agent, Choice, \{\sim_\alpha\}_{\alpha \in Agent}, Type, [\,], Label, V\rangle$$

Type $= \{\tau_1, \tau_2, \ldots\}$ is a finite set of action types—general kinds of action, as opposed to the concrete action tokens already present in stit logics.

$[\,]$ is a partial function mapping types to the particular action token $[\tau]^m_\alpha$ that results when τ is executed by α at *m*.

Labeled stit model

$$\langle \mathit{Tree}, <, \mathit{Agent}, \mathit{Choice}, \{\sim_\alpha\}_{\alpha \in \mathit{Agent}}, \mathit{Type}, [\,], \mathit{Label}, V \rangle$$

$\mathit{Type} = \{\tau_1, \tau_2, \ldots\}$ is a finite set of action types—general kinds of action, as opposed to the concrete action tokens already present in stit logics.

$[\,]$ is a partial function mapping types to the particular action token $[\tau]^m_\alpha$ that results when τ is executed by α at m.

- $[\tau]^m_\alpha \in \mathit{Choice}^m_\alpha$

Labeled stit model

$$\langle Tree, <, Agent, Choice, \{\sim_\alpha\}_{\alpha \in Agent}, Type, [\,], Label, V\rangle$$

Type $= \{\tau_1, \tau_2, \ldots\}$ is a finite set of action types—general kinds of action, as opposed to the concrete action tokens already present in stit logics.

[ ] is a partial function mapping types to the particular action token $[\tau]^m_\alpha$ that results when τ is executed by α at m.

- $[\tau]^m_\alpha \in Choice^m_\alpha$

Label is a 1-1 function mapping $Choice^m_\alpha$ to action types.

Labeled stit model

$$\langle Tree, <, Agent, Choice, \{\sim_\alpha\}_{\alpha \in Agent}, Type, [\,], Label, V\rangle$$

Type $= \{\tau_1, \tau_2, \ldots\}$ is a finite set of action types—general kinds of action, as opposed to the concrete action tokens already present in stit logics.

[ ] is a partial function mapping types to the particular action token $[\tau]^m_\alpha$ that results when τ is executed by α at *m*.

- $[\tau]^m_\alpha \in Choice^m_\alpha$

Label is a 1-1 function mapping $Choice^m_\alpha$ to action types.

- If $K \in Choice^m_\alpha$, then $[Label(K)]^\alpha_m = K$
- If $\tau \in Type$ and $[\tau]^m_\alpha$ is defined, then $Label([\tau]^m_\alpha) = \tau$

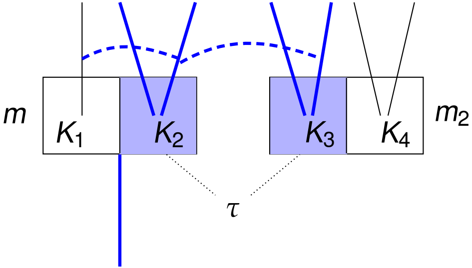
Labeled stit model, continued

$$\langle Tree, <, Agent, Choice, \{\sim_\alpha\}_{\alpha \in Agent}, Type, [\,], Label, V \rangle$$

$$Type^m_\alpha = \{Label(K) \mid K \in Choice^m_\alpha\}$$

$$Type^m_\alpha(h) = Label(Choice^m_\alpha(h))$$

kstit



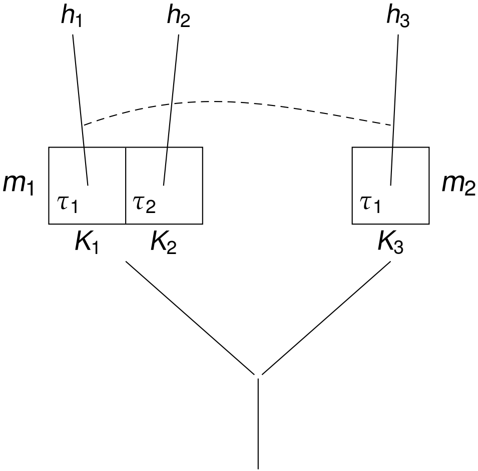
- $\mathcal{M}, m/h \models [\alpha\ kstit\colon A]$ if and only if $[Type^{m}_{\alpha}(h)]^{m'}_{\alpha} \subseteq |A|^{m'}_{\mathcal{M}}$ for all m'/h' such that $m'/h' \sim_{\alpha} m/h$.

The difference between $C1$ and $C2$

$(C1)$ If $m/h \sim_\alpha m'/h'$, then $\mathit{Type}_\alpha^m = \mathit{Type}_\alpha^{m'}$

$(C2)$ If $m/h \sim_\alpha m'/h'$, then $[\mathit{Type}_\alpha^m(h)]_\alpha^{m'}$ is defined.

Minimal Constraint



Knowledge of action types

Let A^τ_α be an atomic proposition carrying the intuitive meaning that the agent α executes the action type τ.

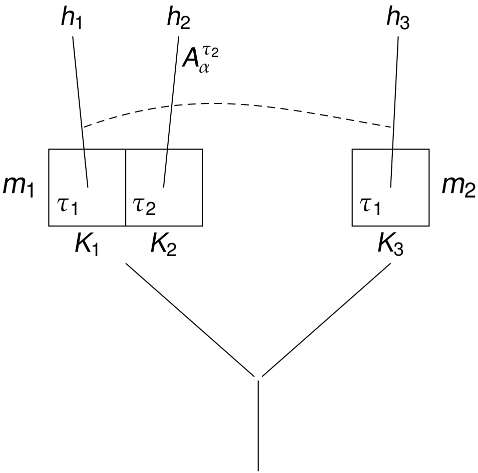
- $\mathcal{M}, m/h \models A^\tau_\alpha$ if and only if $Type^m_\alpha(h) = \tau$

Knowledge of action types

Let A_α^τ be an atomic proposition carrying the intuitive meaning that the agent α executes the action type τ.

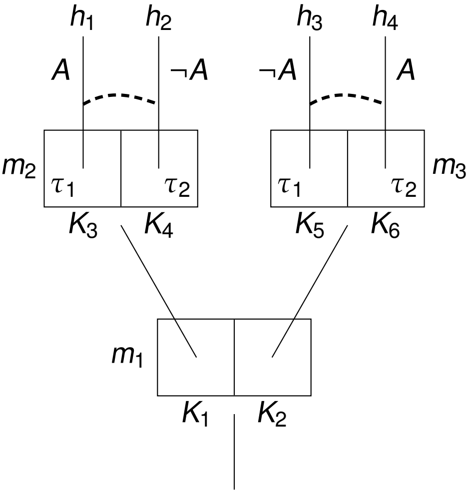
- $\mathcal{M}, m/h \models A_\alpha^\tau$ if and only if $Type_\alpha^m(h) = \tau$

C2 is satisfied iff $\Diamond A_\alpha^\tau \supset \mathsf{K}_\alpha \Diamond A_\alpha^\tau$ is valid.



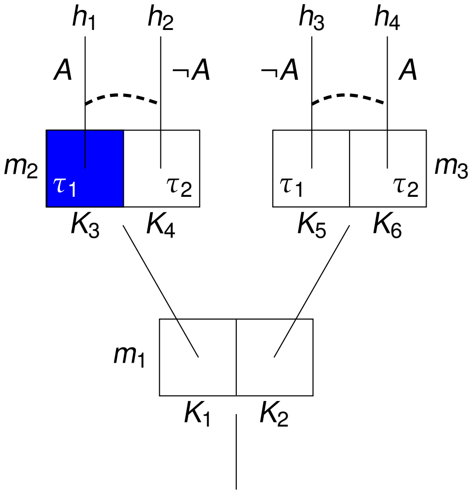
$m_1/h_1 \models \Diamond A_\alpha^{\tau_2} \qquad m_1/h_1 \not\models \mathsf{K}_\alpha \Diamond A_\alpha^{\tau_2}$

Epistemic sense of ability



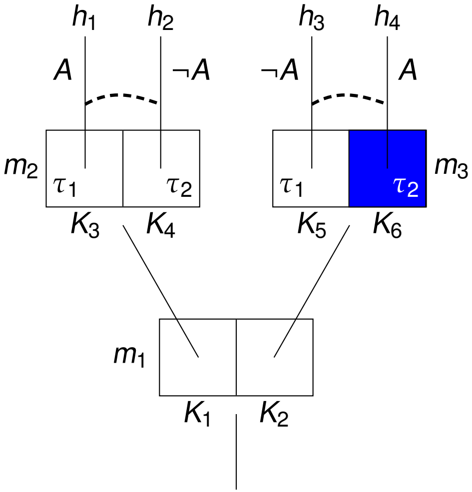
$\Diamond[\alpha$ *kstit*: $A]$ is settled true at m_2 and m_3.

Epistemic sense of ability



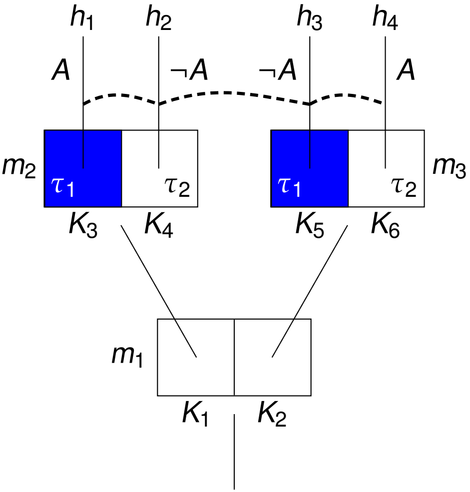
$\Diamond[\alpha\ kstit\colon A]$ is settled true at m_2 and m_3.

Epistemic sense of ability



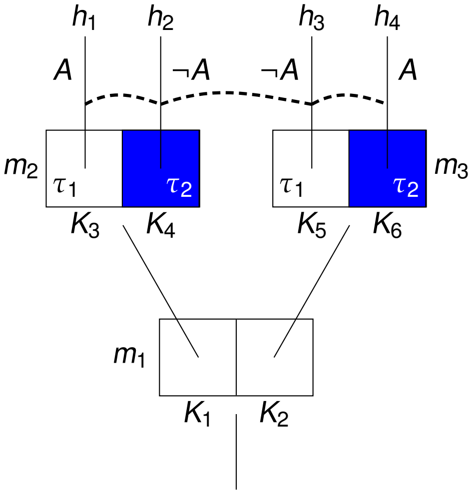
$\Diamond[\alpha$ *kstit*: $A]$ is settled true at m_2 and m_3.

Epistemic sense of ability



$\Diamond[\alpha\ kstit\colon A]$ is settled false at m_2 and m_3.

Epistemic sense of ability



$\Diamond[\alpha$ *kstit*: $A]$ is settled false at m_2 and m_3.

Discussion: Related Work

A. Herzig and N. Troquard. *Knowing how to play: uniform choices in logics of agency*. In Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multi-agent Systems (AAMAS-06), pages 209 - 216. 2006..

J. Broersen. *Deontic epistemic stit logic distinguishing modes of* mens rea. Journal of Applied Logic, 9(2):127 - 152, 2011.

A. Herzig and E. Lorini. *A Dynamic Logic of Agency I: STIT, Capabilities and Powers*. Journal of Logic, Language and Information 19(1): 89-121, 2010.

EP, R. Parikh, and E. Cogan. *The logic of knowledge based obligation*. Synthese, 149:2, pp. 311 - 341, 2006.

M. Xu. *Combinations of stit and actions*. Journal of Logic, Language, and Information, 19:485 - 503, 2010.

Discussion

Validities:

- $\mathsf{K}_\alpha[\alpha\ \textit{stit}: A] \supset [\alpha\ \textit{kstit}: A]$
- $[\alpha\ \textit{kstit}: A] \supset [\alpha\ \textit{stit}: A]$

Discussion

Validities:

- $\mathsf{K}_\alpha[\alpha\ \textit{stit}: A] \supset [\alpha\ \textit{kstit}: A]$
- $[\alpha\ \textit{kstit}: A] \supset [\alpha\ \textit{stit}: A]$

Non-Validities:

- $\Diamond[\alpha\ \textit{kstit}: A] \supset \mathsf{K}_\alpha\Diamond[\alpha\ \textit{kstit}: A]$

Constraints

$(C3)$ If $m/h \sim_\alpha m'/h'$, then $m = m'$

$(C3)$ is satisfied iff $[\alpha\ \textit{stit}\colon A] \equiv [\alpha\ \textit{kstit}\colon A]$ is valid.

$$(C4) \qquad \text{If } m/h \sim_\alpha m'/h', \text{ then } \mathit{Type}_\alpha^m(h) = \mathit{Type}_\alpha^{m'}(h')$$

$(C4)$ is satisfied iff $A_\alpha^\tau \supset \mathsf{K}_\alpha A_\alpha^\tau$ is valid.

Deliberative perspective

(*C*5) If $m/h \sim_\alpha m'/h'$, then $m/h'' \sim_\alpha m'/h'''$ for all $h'' \in H^m$ and $h''' \in H^{m'}$

Indistinguishability between moments: $m \sim_\alpha m'$ iff $m/h \sim_\alpha m'/h'$ for all $h \in H^m$ and $h' \in H^{m'}$.

Discussion

- Language/validities

  $\Box A \supset [\alpha\ stit: A]$
  $\mathsf{K}_\alpha \Box A \supset [\alpha\ kstit: A]$
  $[\alpha\ kstit: A] \equiv \mathsf{K}_\alpha^{act}[\alpha\ stit: A]$
  $\ldots$

- What do the agents know vs. What do the agents know *given what they are doing*.

- Equivalence between labeled stit models (cf. Thompson transformations specifying when two imperfect information games reduce to the same Normal form)