Logic and Artificial Intelligence

Lecture 18

Eric Pacuit

Currently Visiting the Center for Formal Epistemology, CMU

Center for Logic and Philosophy of Science
Tilburg University

ai.stanford.edu/∼epacuit
e.j.pacuit@uvt.nl

November 8, 2011

Ingredients of a Logical Analysis of Rational Agency

- $\Rightarrow$ informational attitudes (eg., knowledge, belief, certainty)
- $\Rightarrow$ time, actions and ability
- $\Rightarrow$ motivational attitudes (eg., preferences)
- $\Rightarrow$ group notions (eg., common knowledge and coalitional ability)
- $\Rightarrow$ normative attitudes (eg., obligations)

Ingredients of a Logical Analysis of Rational Agency

- ✓ informational attitudes (eg., knowledge, belief, certainty)
- ⇒ time, actions and ability
- ⇒ motivational attitudes (eg., preferences)
- ⇒ group notions (eg., common knowledge and coalitional ability)
- ⇒ normative attitudes (eg., obligations)

Time

One of the most "successful" applications of modal logic is in the "logic of time".

Time

One of the most “successful” applications of modal logic is in the “logic of time”.

Many variations

- discrete or continuous
- branching or linear
- point based or interval based

See, for example,

Antony Galton. *Temporal Logic*. Stanford Encyclopedia of Philosophy: `http://plato.stanford.edu/entries/logic-temporal/`.

I. Hodkinson and M. Reynolds. *Temporal Logic*. Handbook of Modal Logic, 2008.

Models of Time

$\mathcal{T} = \langle T, <, V \rangle$ where

- T is a set of **time points** (or **moments**),
- $< \subseteq T \times T$ is the **precedence relation**: $s < t$ means "time point s precedes time point t (or s occurs earlier than t)" and
- $V : \mathsf{At} \to \wp(T)$ is a valuation function (describing when the atomic propositions are true).

Models of Time

$\mathcal{T} = \langle T, <, V \rangle$ where

- T is a set of **time points** (or **moments**),
- $< \subseteq T \times T$ is the **precedence relation**: $s < t$ means "time point s precedes time point t (or s occurs earlier than t)" and
- $V : \mathrm{At} \to \wp(T)$ is a valuation function (describing when the atomic propositions are true).

$<$ is typically assumed to be **irreflexive** and **transitive** (a **strict partial order**).

Models of Time

$\mathcal{T} = \langle T, <, V \rangle$ where

- T is a set of **time points** (or **moments**),
- $< \subseteq T \times T$ is the **precedence relation**: $s < t$ means "time point s precedes time point t (or s occurs earlier than t)" and
- $V : \mathrm{At} \to \wp(T)$ is a valuation function (describing when the atomic propositions are true).

$<$ is typically assumed to be **irreflexive** and **transitive** (a **strict partial order**).

Examples: $\langle \mathbb{N}, < \rangle$, $\langle \mathbb{Z}, < \rangle$, $\langle \mathbb{Q}, < \rangle$, $\langle \mathbb{R}, < \rangle$

Other properties of $<$

- **Linearity**: for all $s, t \in T$, $s < t$ or $s = t$ of $t < s$
- **Past-linear**: for all $s, x, y \in T$, if $x < s$ and $y < s$, then either $x < y$ or $x = y$ or $y < x$
- **Denseness** for all $s, t \in T$, if $s < t$ then there is a $z \in T$ such that $s < z$ and $z < t$
- **Discreteness**: for all $s, t \in T$, if $s < t$ then there is a z such that ($s < z$ and there is no u such that $s < u$ and $u < z$)

Priorean Temporal Logic

$\mathcal{L}_t$ be defined by the following grammar

$$p \mid \neg\varphi \mid \varphi \wedge \psi \mid G\varphi \mid H\varphi$$

where $p \in \mathsf{At}$.

Priorean Temporal Logic

$\mathcal{L}_t$ be defined by the following grammar

$$p \mid \neg\varphi \mid \varphi \wedge \psi \mid G\varphi \mid H\varphi$$

where $p \in \text{At}$.

$G\varphi$: "φ is **going to** become true"

$H\varphi$: "φ **has been** true"

$F\varphi := \neg G \neg \varphi$: "$\varphi$ is true in the future"

$P\varphi := \neg H \neg \varphi$: "$\varphi$ was true some time in the past"

$\mathcal{M} = \langle T, <, V \rangle$

- $\mathcal{M}, t \models p$ iff $t \in V(p)$
- $\mathcal{M}, t \models \neg\varphi$ iff $\mathcal{M}, t \not\models \varphi$
- $\mathcal{M}, t \models \varphi \wedge \psi$ iff $\mathcal{M}, t \models \varphi$ and $\mathcal{M}, t \models \psi$
- $\mathcal{M}, t \models G\varphi$ iff for all $s \in T$, if $t < s$ then $\mathcal{M}, s \models \varphi$
- $\mathcal{M}, t \models H\varphi$ iff for all $s \in T$, if $s < t$ then$\mathcal{M}, s \models \varphi$

- $\mathcal{M}, t \models F\varphi$ iff there is $s \in T$ such that $t < s$ and $\mathcal{M}, s \models \varphi$
- $\mathcal{M}, t \models P\varphi$ iff there is $s \in T$ such that $s < t$ and $\mathcal{M}, s \models \varphi$

Frame Correspondence

- $H\bot \lor PH\bot$ is valid

Frame Correspondence

- $H\bot \vee PH\bot$ is valid iff there is a starting point

Frame Correspondence

- $H\bot \vee PH\bot$ is valid iff there is a starting point
- $P\top$ is valid

Frame Correspondence

- $H\bot \vee PH\bot$ is valid iff there is a starting point
- $P\top$ is valid iff there is no starting point

Frame Correspondence

- $H\bot \vee PH\bot$ is valid iff there is a starting point
- $P\top$ is valid iff there is no starting point
- $G\bot \vee FG\bot$ is valid

Frame Correspondence

- $H\bot \vee PH\bot$ is valid iff there is a starting point
- $P\top$ is valid iff there is no starting point
- $G\bot \vee FG\bot$ is valid iff there is an end point

Frame Correspondence

- $H\bot \vee PH\bot$ is valid iff there is a starting point
- $P\top$ is valid iff there is no starting point
- $G\bot \vee FG\bot$ is valid iff there is an end point
- $F\top$ is valid

Frame Correspondence

- $H\bot \vee PH\bot$ is valid iff there is a starting point
- $P\top$ is valid iff there is no starting point
- $G\bot \vee FG\bot$ is valid iff there is an end point
- $F\top$ is valid iff there is no endpoint

Frame Correspondence

- $H\bot \vee PH\bot$ is valid iff there is a starting point
- $P\top$ is valid iff there is no starting point
- $G\bot \vee FG\bot$ is valid iff there is an end point
- $F\top$ is valid iff there is no endpoint
- $PF\varphi \rightarrow (P\varphi \vee \varphi \vee F\varphi)$ is valid

Frame Correspondence

- $H\bot \vee PH\bot$ is valid iff there is a starting point
- $P\top$ is valid iff there is no starting point
- $G\bot \vee FG\bot$ is valid iff there is an end point
- $F\top$ is valid iff there is no endpoint
- $PF\varphi \to (P\varphi \vee \varphi \vee F\varphi)$ is valid iff the future is not branching

Frame Correspondence

- $H\bot \vee PH\bot$ is valid iff there is a starting point
- $P\top$ is valid iff there is no starting point
- $G\bot \vee FG\bot$ is valid iff there is an end point
- $F\top$ is valid iff there is no endpoint
- $PF\varphi \to (P\varphi \vee \varphi \vee F\varphi)$ is valid iff the future is not branching
- $FP\varphi \to (F\varphi \vee \varphi \vee P\varphi)$ is valid

Frame Correspondence

- $H\bot \vee PH\bot$ is valid iff there is a starting point
- $P\top$ is valid iff there is no starting point
- $G\bot \vee FG\bot$ is valid iff there is an end point
- $F\top$ is valid iff there is no endpoint
- $PF\varphi \rightarrow (P\varphi \vee \varphi \vee F\varphi)$ is valid iff the future is not branching
- $FP\varphi \rightarrow (F\varphi \vee \varphi \vee P\varphi)$ is valid iff the past is non-branching

Frame Correspondence

- $H\bot \vee PH\bot$ is valid iff there is a starting point
- $P\top$ is valid iff there is no starting point
- $G\bot \vee FG\bot$ is valid iff there is an end point
- $F\top$ is valid iff there is no endpoint
- $PF\varphi \rightarrow (P\varphi \vee \varphi \vee F\varphi)$ is valid iff the future is not branching
- $FP\varphi \rightarrow (F\varphi \vee \varphi \vee P\varphi)$ is valid iff the past is non-branching
- $F\varphi \rightarrow FF\varphi$ is valid

Frame Correspondence

- $H\bot \vee PH\bot$ is valid iff there is a starting point
- $P\top$ is valid iff there is no starting point
- $G\bot \vee FG\bot$ is valid iff there is an end point
- $F\top$ is valid iff there is no endpoint
- $PF\varphi \to (P\varphi \vee \varphi \vee F\varphi)$ is valid iff the future is not branching
- $FP\varphi \to (F\varphi \vee \varphi \vee P\varphi)$ is valid iff the past is non-branching
- $F\varphi \to FF\varphi$ is valid iff the flow of time is dense

Frame Correspondence

- $H\bot \vee PH\bot$ is valid iff there is a starting point
- $P\top$ is valid iff there is no starting point
- $G\bot \vee FG\bot$ is valid iff there is an end point
- $F\top$ is valid iff there is no endpoint
- $PF\varphi \to (P\varphi \vee \varphi \vee F\varphi)$ is valid iff the future is not branching
- $FP\varphi \to (F\varphi \vee \varphi \vee P\varphi)$ is valid iff the past is non-branching
- $F\varphi \to FF\varphi$ is valid iff the flow of time is dense
- $(F\top \wedge \varphi \wedge H\varphi) \to FH\varphi$ is valid

Frame Correspondence

- $H\bot \vee PH\bot$ is valid iff there is a starting point
- $P\top$ is valid iff there is no starting point
- $G\bot \vee FG\bot$ is valid iff there is an end point
- $F\top$ is valid iff there is no endpoint
- $PF\varphi \to (P\varphi \vee \varphi \vee F\varphi)$ is valid iff the future is not branching
- $FP\varphi \to (F\varphi \vee \varphi \vee P\varphi)$ is valid iff the past is non-branching
- $F\varphi \to FF\varphi$ is valid iff the flow of time is dense
- $(F\top \wedge \varphi \wedge H\varphi) \to FH\varphi$ is valid iff the flow of time is discrete

Basic Temporal Logic

All classical propositional tautologies

Distribution
$G(\varphi \to \psi) \to (G\varphi \to G\psi)$
$G(\varphi \to \psi) \to (G\varphi \to G\psi)$

Converse
$\varphi \to GP\varphi$
$\varphi \to HF\varphi$

Transitivity: $G\varphi \to GG\varphi$

Modus Ponens: from φ and $\varphi \to \psi$ infer ψ
Temporal Generalization: from φ infer $F\varphi$; from φ infer $G\varphi$

Basic Temporal Logic

All classical propositional tautologies

Distribution
$G(\varphi \to \psi) \to (G\varphi \to G\psi)$
$G(\varphi \to \psi) \to (G\varphi \to G\psi)$

Converse
$\varphi \to GP\varphi$
$\varphi \to HF\varphi$

Transitivity: $G\varphi \to GG\varphi$

Modus Ponens: from φ and $\varphi \to \psi$ infer ψ
Temporal Generalization: from φ infer $F\varphi$; from φ infer $G\varphi$

Theorem. The above logic is sound and complete with respect to the class of all flows of time

Logic of Linear Time

Theorem. The above logic with the linearity axioms is sound and complete with respect to the class of all **linear** flows of time

- $PF\varphi \to (P\varphi \vee \varphi \vee F\varphi)$
- $FP\varphi \to (F\varphi \vee \varphi \vee P\varphi)$

Other Languages: Since and Until

$\mathcal{M}, t \models \varphi U \psi$ iff $\mathcal{M}, s \models \psi$ for some s such that $t < s$ and $\mathcal{M}, u \models \varphi$ for all u with $t < u < s$

$\mathcal{M}, t \models \varphi S \psi$ iff $\mathcal{M}, s \models \psi$ for some s such that $s < t$ and $\mathcal{M}, u \models \varphi$ for all u with $s < u < t$

Other Languages: Since and Until

$\mathcal{M}, t \models \varphi U \psi$ iff $\mathcal{M}, s \models \psi$ for some s such that $t < s$ and $\mathcal{M}, u \models \varphi$ for all u with $t < u < s$

$\mathcal{M}, t \models \varphi S \psi$ iff $\mathcal{M}, s \models \psi$ for some s such that $s < t$ and $\mathcal{M}, u \models \varphi$ for all u with $s < u < t$

Theorem (Kamp). Over the class of linear, continuous orderings, every temporal operator can be defined using the above modalities

Branching Time

Each moment $t \in T$ can be decided into the $Past(t) = \{s \in T \mid s < t\}$ and the $Future(t) = \{s \in T \mid t < s\}$ ("A-series")

Typically, it is assumed that the past is linear, but the future may be branching.

Branching Time

Each moment $t \in T$ can be decided into the $\mathit{Past}(t) = \{s \in T \mid s < t\}$ and the $\mathit{Future}(t) = \{s \in T \mid t < s\}$ ("A-series")

Typically, it is assumed that the past is linear, but the future may be branching.

$F\varphi$: "it will be the case that φ"

φ will be the case "in the case in the actual course of events" or "no matter what course of events"

Branching Time Logics

A **branch** b in $\langle T, < \rangle$ is a maximal linearly ordered subset of T

$s \in T$ is **on a branch** b **of** T provided $s \in b$ (we also say "b is a branch going through t").

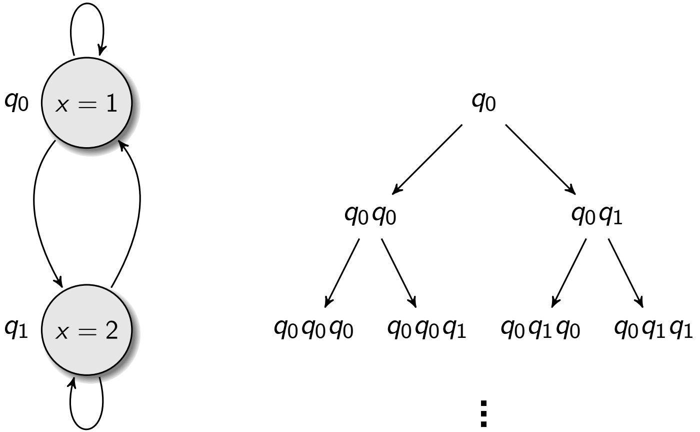
Branching Time Logics

A **branch** b in $\langle T, < \rangle$ is a maximal linearly ordered subset of T

$s \in T$ is **on a branch** b **of** T provided $s \in b$ (we also say "b is a branch going through t").

- $\mathcal{M}, t, b \models p$ iff $t \in V(p)$
- $\mathcal{M}, t, b \models \neg\varphi$ iff $\mathcal{M}, t, b \not\models \varphi$
- $\mathcal{M}, t, b \models \varphi \wedge \psi$ iff $\mathcal{M}, t, b \models \varphi$ and $\mathcal{M}, t \models \psi$
- $\mathcal{M}, t, b \models G\varphi$ iff for all $s \in T$, if s is on b and $t < s$ then $\mathcal{M}, s, b \models \varphi$
- $\mathcal{M}, t, b \models H\varphi$ iff for all $s \in T$, if s is on b and $s < t$ then $\mathcal{M}, s, b \models \varphi$
- $\mathcal{M}, t, b \models \forall\varphi$ iff $\mathcal{M}, s, c \models \varphi$ for all branches c through t

Computational vs. Behavioral Structures



Temporal Logics

Temporal Logics

- *Linear Time Temporal Logic*: Reasoning about computation paths:

$$F\varphi\text{: } \varphi \text{ is true some time in } \textit{the} \text{ future.}$$

A. Pnuelli. *A Temporal Logic of Programs*. in *Proc. 18th IEEE Symposium on Foundations of Computer Science* (1977).

Temporal Logics

- *Linear Time Temporal Logic*: Reasoning about computation paths:

  $F\varphi$: φ is true some time in *the* future.

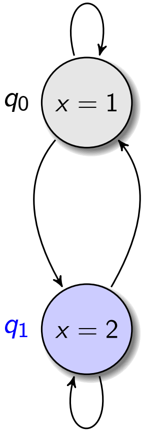
A. Pnuelli. *A Temporal Logic of Programs*. in *Proc. 18th IEEE Symposium on Foundations of Computer Science* (1977).

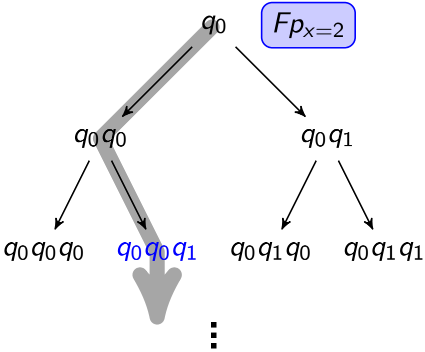
- *Branching Time Temporal Logic*: Allows quantification over paths:

  $\exists F\varphi$: there is a path in which φ is eventually true.

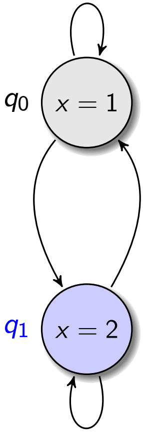
E. M. Clarke and E. A. Emerson. *Design and Synthesis of Synchronization Skeletons using Branching-time Temproal-logic Specifications*. In *Proceedings Workshop on Logic of Programs*, LNCS (1981).

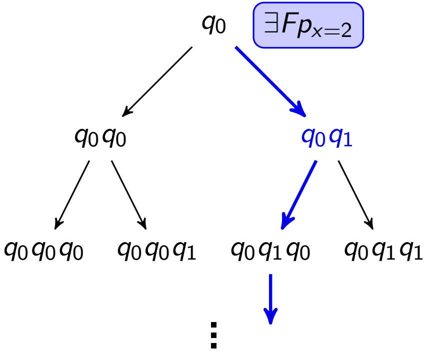
Temporal Logics



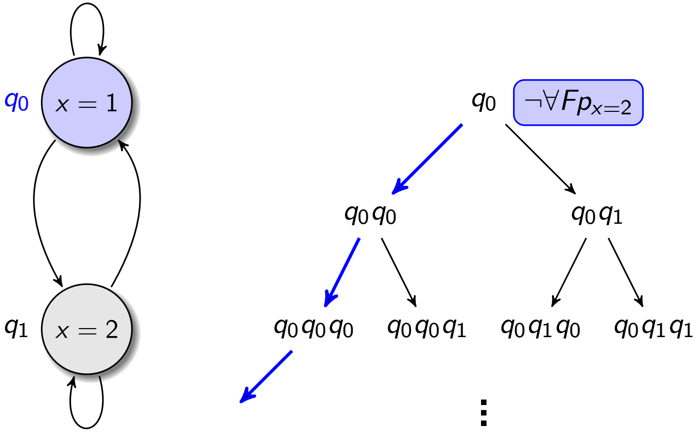


Temporal Logics





Temporal Logics



Interval Values

J. Allen and G. Ferguson. *Actions and Events in Interval Temporal Logics*. Journal of Logic and Computation, 1994.

J. Halpern and Y. Shoham. *A Propositional Modal Logic of Time Intervals*. Journal of the ACM, 38:4, pp. 935 - 962, 1991.

J. van Benthem. *Logics of Time*. Kluwer, 1991.

Interval Temporal Logics

Let $\mathcal{T} = \langle T, < \rangle$ be a frame and $I(\mathcal{T}) = \{[a, b] \mid a, b \in T$ and $a \leq b\}$ be the set of intervals over T

Models are $\mathcal{M} = \langle I(\mathcal{T}), \{R_X\}, V \rangle$ where $R_X \subseteq I(\mathcal{T}) \times I(\mathcal{T})$ and $V : \mathrm{At} \to \wp(I(\mathcal{T}))$.

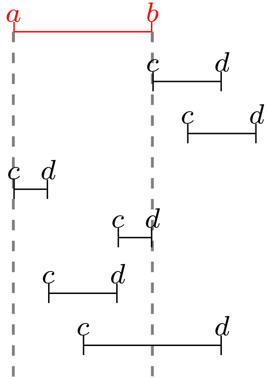
Interval Temporal Logics

Let $\mathcal{T} = \langle T, < \rangle$ be a frame and $I(\mathcal{T}) = \{[a, b] \mid a, b \in T$ and $a \leq b\}$ be the set of intervals over T

Models are $\mathcal{M} = \langle I(\mathcal{T}), \{R_X\}, V \rangle$ where $R_X \subseteq I(\mathcal{T}) \times I(\mathcal{T})$ and $V : \mathsf{At} \to \wp(I(\mathcal{T}))$.

- $\mathcal{M}, [a, b] \models p$ iff $[a, b] \in V(p)$
- $\mathcal{M}, [a, b] \models \mathsf{pt}$ iff $a = b$
- $\mathcal{M}, [a, b] \models \langle X \rangle \varphi$ iff there is an interval $[c, d]$ such that $[a, b] R_X [c, d]$ and $\mathcal{M}, [c, d] \models \varphi$

$\langle A \rangle$	$[a,b]R_A[c,d] \Leftrightarrow b = c$
$\langle L \rangle$	$[a,b]R_L[c,d] \Leftrightarrow b < c$
$\langle B \rangle$	$[a,b]R_B[c,d] \Leftrightarrow a = c, d < b$
$\langle E \rangle$	$[a,b]R_E[c,d] \Leftrightarrow b = d, a < c$
$\langle D \rangle$	$[a,b]R_D[c,d] \Leftrightarrow a < c, d < b$
$\langle O \rangle$	$[a,b]R_O[c,d] \Leftrightarrow a < c < b < d$



High Undecidability!

D. Bersolin et al.. *The dark side of interval temporal logic: sharpening the undecidability border*. 2011.

Actions and Abilities: Pre-theoretic Intuitions

What does it mean for an agent to be able to do some action *a* or bring about some state of affairs φ?

Actions and Abilities: Pre-theoretic Intuitions

What does it mean for an agent to be able to do some action a or bring about some state of affairs φ?

1. **Control**: The action a should be in the control of agent i or the truth of φ should be under i's control.

Actions and Abilities: Pre-theoretic Intuitions

What does it mean for an agent to be able to do some action *a* or bring about some state of affairs φ?

1. **Control**: The action *a* should be in the control of agent *i* or the truth of φ should be under *i*'s control.
2. **Repeatability**: Agent *i* should be able to repeatedly do action *a* or repeatedly bring about formula φ.

Actions and Abilities: Pre-theoretic Intuitions

What does it mean for an agent to be able to do some action a or bring about some state of affairs φ?

1. **Control**: The action a should be in the control of agent i or the truth of φ should be under i's control.
2. **Repeatability**: Agent i should be able to repeatedly do action a or repeatedly bring about formula φ.
3. **Avoidability**: Agent i should be able to avoid doing action a.

Actions and Abilities: Pre-theoretic Intuitions

What does it mean for an agent to be able to do some action a or bring about some state of affairs φ?

1. **Control**: The action a should be in the control of agent i or the truth of φ should be under i's control.
2. **Repeatability**: Agent i should be able to repeatedly do action a or repeatedly bring about formula φ.
3. **Avoidability**: Agent i should be able to avoid doing action a.
4. **Free-will**: Agent i should be free to decide whether or not to do action a. That is, the agent should have a choice as to whether or not to do action a or bring about formula φ.

Actions and Abilities: Pre-theoretic Intuitions

What does it mean for an agent to be able to do some action *a* or bring about some state of affairs φ?

1. **Control**: The action *a* should be in the control of agent *i* or the truth of φ should be under *i*'s control.
2. **Repeatability**: Agent *i* should be able to repeatedly do action *a* or repeatedly bring about formula φ.
3. **Avoidability**: Agent *i* should be able to avoid doing action *a*.
4. **Free-will**: Agent *i* should be free to decide whether or not to do action *a*. That is, the agent should have a choice as to whether or not to do action *a* or bring about formula φ.
5. **Causality**: Agent *i* should *cause* action *a* to take place or the formula φ to become true.

Actions and Abilities: Pre-theoretic Intuitions

What does it mean for an agent to be able to do some action a or bring about some state of affairs φ?

1. **Control**: The action a should be in the control of agent i or the truth of φ should be under i's control.
2. **Repeatability**: Agent i should be able to repeatedly do action a or repeatedly bring about formula φ.
3. **Avoidability**: Agent i should be able to avoid doing action a.
4. **Free-will**: Agent i should be free to decide whether or not to do action a. That is, the agent should have a choice as to whether or not to do action a or bring about formula φ.
5. **Causality**: Agent i should *cause* action a to take place or the formula φ to become true.
6. **Intentionality**: Agent i should *intentionally* do action a or bring about φ.

Abilities

$Abl_i\varphi$: agent i has the ability to bring about (see to it that) φ is true

What are core logical principles?

Abilities

$Abl_i\varphi$: agent i has the ability to bring about (see to it that) φ is true

What are core logical principles?

1. $Abl_i\varphi \rightarrow \varphi$ (or $\varphi \rightarrow Abl_i\varphi$)

Abilities

$Abl_i\varphi$: agent i has the ability to bring about (see to it that) φ is true

What are core logical principles?

1. $Abl_i\varphi \rightarrow \varphi$ (or $\varphi \rightarrow Abl_i\varphi$)
2. $\neg Abl_i\top$

Abilities

$Abl_i\varphi$: agent i has the ability to bring about (see to it that) φ is true

What are core logical principles?

1. $Abl_i\varphi \rightarrow \varphi$ (or $\varphi \rightarrow Abl_i\varphi$)
2. $\neg Abl_i\top$
3. $(Abl_i\varphi \wedge Abl_i\psi) \rightarrow Abl_i(\varphi \wedge \psi)$

Abilities

$\mathit{Abl}_i\varphi$: agent i has the ability to bring about (see to it that) φ is true

What are core logical principles?

1. $\mathit{Abl}_i\varphi \to \varphi$ (or $\varphi \to \mathit{Abl}_i\varphi$)
2. $\neg \mathit{Abl}_i\top$
3. $(\mathit{Abl}_i\varphi \wedge \mathit{Abl}_i\psi) \to \mathit{Abl}_i(\varphi \wedge \psi)$
4. $\mathit{Abl}_i(\varphi \vee \psi) \to (\mathit{Abl}_i\varphi \vee \mathit{Abl}_i\psi)$

Abilities

$Abl_i\varphi$: agent i has the ability to bring about (see to it that) φ is true

What are core logical principles?

1. $Abl_i\varphi \rightarrow \varphi$ (or $\varphi \rightarrow Abl_i\varphi$)
2. $\neg Abl_i\top$
3. $(Abl_i\varphi \wedge Abl_i\psi) \rightarrow Abl_i(\varphi \wedge \psi)$
4. $Abl_i(\varphi \vee \psi) \rightarrow (Abl_i\varphi \vee Abl_i\psi)$
5. $Abl_i(\varphi \wedge \psi) \rightarrow (Abl_i\varphi \wedge Abl_i\psi)$

Abilities

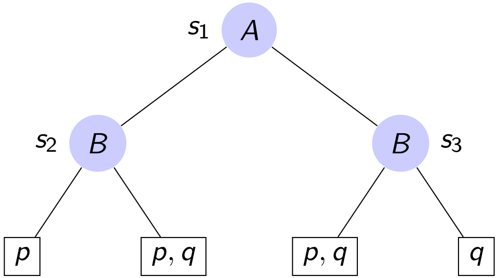
$Abl_i\varphi$: agent i has the ability to bring about (see to it that) φ is true

What are core logical principles?

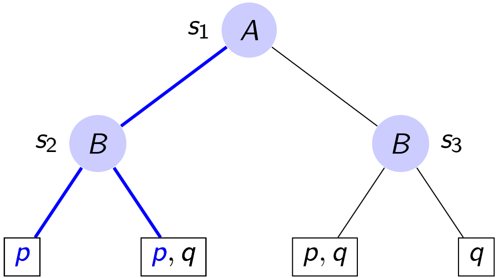
1. $Abl_i\varphi \rightarrow \varphi$ (or $\varphi \rightarrow Abl_i\varphi$)
2. $\neg Abl_i\top$
3. $(Abl_i\varphi \wedge Abl_i\psi) \rightarrow Abl_i(\varphi \wedge \psi)$
4. $Abl_i(\varphi \vee \psi) \rightarrow (Abl_i\varphi \vee Abl_i\psi)$
5. $Abl_i(\varphi \wedge \psi) \rightarrow (Abl_i\varphi \wedge Abl_i\psi)$
6. $Abl_iAbl_j\varphi \rightarrow Abl_i\varphi$, $Abl_iAbl_i\varphi \rightarrow Abl_i\varphi$

Games: $(Abl_i\varphi \wedge Abl_i\psi) \not\to Abl_i(\varphi \wedge \psi)$

Games: $(Abl_i\varphi \wedge Abl_i\psi) \not\to Abl_i(\varphi \wedge \psi)$

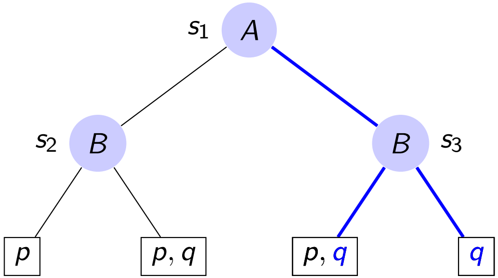


Games: $(Abl_i\varphi \wedge Abl_i\psi) \not\to Abl_i(\varphi \wedge \psi)$



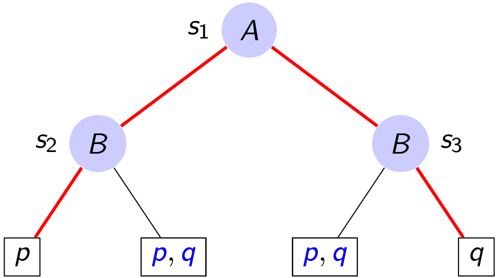
$s_1 \models Abl_A p$

Games: $(Abl_i\varphi \wedge Abl_i\psi) \not\to Abl_i(\varphi \wedge \psi)$



$$s_1 \models Abl_A p \wedge Abl_A q$$

Games: $(Abl_i\varphi \wedge Abl_i\psi) \not\to Abl_i(\varphi \wedge \psi)$



$$s_1 \models Abl_A p \wedge Abl_A q \wedge \neg Abl_A(p \wedge q)$$

Games: $(Abl_i\varphi \wedge Abl_i\psi) \not\to Abl_i(\varphi \wedge \psi)$

R. Parikh. *The Logic of Games and its Applications*. Annals of Discrete Mathematics (1985).

M. Pauly and R. Parikh. *Game Logic — An Overview*. Studia Logica (2003).

J. van Benthem. *Logic and Games*. Course notes (2007).

$\varphi \not\to Abl_i\varphi$

Suppose an agent (call her Ann) is throwing a dart and she is not a very good dart player, but she just happens to throw a bull's eye.

$\varphi \not\to Abl_i\varphi$

Suppose an agent (call her Ann) is throwing a dart and she is not a very good dart player, but she just happens to throw a bull's eye.

Intuitively, we do not want to say that Ann has the *ability* to throw a bull's eye even though it happens to be true.

$Abl_i(\varphi \vee \psi) \not\to Abl_i\varphi \vee Abl_i\psi$

Continuing with this example, suppose that Ann has the ability to hit the dart board, but has no other control over the placement of the dart.

$Abl_i(\varphi \vee \psi) \not\rightarrow Abl_i\varphi \vee Abl_i\psi$

Continuing with this example, suppose that Ann has the ability to hit the dart board, but has no other control over the placement of the dart.

Now, when she throws the dart, as a matter of fact, it will either hit the top half of the board or the bottom half of the board.

$Abl_i(\varphi \vee \psi) \not\rightarrow Abl_i\varphi \vee Abl_i\psi$

Continuing with this example, suppose that Ann has the ability to hit the dart board, but has no other control over the placement of the dart.

Now, when she throws the dart, as a matter of fact, it will either hit the top half of the board or the bottom half of the board.

Since, Ann has the ability to hit the dart board, she has the ability to either hit the top half of the board or the bottom half of the board.

$Abl_i(\varphi \vee \psi) \not\to Abl_i\varphi \vee Abl_i\psi$

Continuing with this example, suppose that Ann has the ability to hit the dart board, but has no other control over the placement of the dart.

Now, when she throws the dart, as a matter of fact, it will either hit the top half of the board or the bottom half of the board.

Since, Ann has the ability to hit the dart board, she has the ability to either hit the top half of the board or the bottom half of the board.

However, intuitively, it seems true that Ann does not have the ability to hit the top half of the dart board, and also she does not have the ability to hit the bottom half of the dart board.

Abilities

$Abl_i\varphi$: agent i has the ability to bring about (see to it that) φ is true

What are core logical principles? Depends very much on the intended "application" and how actions are represented...

1. $Abl_i\varphi \to \varphi$ (or $\varphi \to Abl_i\varphi$)
2. $\neg Abl_i\top$
3. $(Abl_i\varphi \land Abl_i\psi) \to Abl_i(\varphi \land \psi)$
4. $Abl_i(\varphi \lor \psi) \to (Abl_i\varphi \lor Abl_i\psi)$
5. $Abl_i(\varphi \land \psi) \to (Abl_i\varphi \land Abl_i\psi)$
6. $Abl_iAbl_j\varphi \to Abl_i\varphi$, $Abl_iAbl_i\varphi \to Abl_i\varphi$

D. Elgesem. *The modal logic of agency*. Nordic Journal of Philosophical Logic 2(2), 1 - 46, 1997.

G. Governatori and A. Rotolo. *On the Axiomatisation of Elgesem's Logic of Agency and Ability*. Journal of Philosophical Logic, 34, pgs. 403 - 431 (2005).

A Minimal Logic of Abilities

$C\varphi$ means "the agent is capable of realizing φ"

$E\varphi$ means "the agent does bring about φ"

A Minimal Logic of Abilities

$C\varphi$ means "the agent is capable of realizing φ"

$E\varphi$ means "the agent does bring about φ"

1. All propositional tautologies
2. $\neg C\top$
3. $E\varphi \wedge E\psi \rightarrow E(\varphi \wedge \psi)$
4. $E\varphi \rightarrow \varphi$
5. $E\varphi \rightarrow C\varphi$
6. Modus Ponens plus from $\varphi \leftrightarrow \psi$ infer $E\varphi \leftrightarrow E\psi$ and from $\varphi \leftrightarrow \psi$ infer $C\varphi \leftrightarrow C\psi$

Brief digression: weak systems of modal logic

PC Propositional Calculus

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

M $\Box(\varphi \wedge \psi) \rightarrow (\Box\varphi \wedge \Box\psi)$

C $(\Box\varphi \wedge \Box\psi) \rightarrow \Box(\varphi \wedge \psi)$

N $\Box\top$

K $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$

RE $\dfrac{\varphi \leftrightarrow \psi}{\Box\varphi \leftrightarrow \Box\psi}$

Nec $\dfrac{\varphi}{\Box\varphi}$

MP $\dfrac{\varphi \quad \varphi \rightarrow \psi}{\psi}$

PC Propositional Calculus

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

M $\Box(\varphi \wedge \psi) \rightarrow (\Box\varphi \wedge \Box\psi)$

C $(\Box\varphi \wedge \Box\psi) \rightarrow \Box(\varphi \wedge \psi)$

N $\Box\top$

K $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$

RE $\dfrac{\varphi \leftrightarrow \psi}{\Box\varphi \leftrightarrow \Box\psi}$

Nec $\dfrac{\varphi}{\Box\varphi}$

MP $\dfrac{\varphi \quad \varphi \rightarrow \psi}{\psi}$

A modal logic **L** is classical if it contains all instances of *E* and is closed under *RE*.

PC Propositional Calculus

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

M $\Box(\varphi \wedge \psi) \to (\Box\varphi \wedge \Box\psi)$

C $(\Box\varphi \wedge \Box\psi) \to \Box(\varphi \wedge \psi)$

N $\Box\top$

K $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

RE $\dfrac{\varphi \leftrightarrow \psi}{\Box\varphi \leftrightarrow \Box\psi}$

Nec $\dfrac{\varphi}{\Box\varphi}$

MP $\dfrac{\varphi \quad \varphi \to \psi}{\psi}$

A modal logic **L** is classical if it contains all instances of *E* and is closed under *RE*.

E is the smallest classical modal logic.

PC Propositional Calculus

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

M $\Box(\varphi \wedge \psi) \to (\Box\varphi \wedge \Box\psi)$

C $(\Box\varphi \wedge \Box\psi) \to \Box(\varphi \wedge \psi)$

N $\Box\top$

K $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

RE $\dfrac{\varphi \leftrightarrow \psi}{\Box\varphi \leftrightarrow \Box\psi}$

Nec $\dfrac{\varphi}{\Box\varphi}$

MP $\dfrac{\varphi \quad \varphi \to \psi}{\psi}$

E is the smallest classical modal logic.

In **E**, *M* is equivalent to

$$(Mon)\ \frac{\varphi \to \psi}{\Box\varphi \to \Box\psi}$$

PC Propositional Calculus

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

Mon $\dfrac{\varphi \rightarrow \psi}{\Box\varphi \rightarrow \Box\psi}$

C $(\Box\varphi \wedge \Box\psi) \rightarrow \Box(\varphi \wedge \psi)$

N $\Box\top$

K $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$

RE $\dfrac{\varphi \leftrightarrow \psi}{\Box\varphi \leftrightarrow \Box\psi}$

Nec $\dfrac{\varphi}{\Box\varphi}$

MP $\dfrac{\varphi \quad \varphi \rightarrow \psi}{\psi}$

E is the smallest classical modal logic.

EM is the logic **E** + *Mon*

PC 6. Propositional Calculus

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

Mon $\dfrac{\varphi \rightarrow \psi}{\Box\varphi \rightarrow \Box\psi}$

C $(\Box\varphi \wedge \Box\psi) \rightarrow \Box(\varphi \wedge \psi)$

N $\Box\top$

K $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$

RE $\dfrac{\varphi \leftrightarrow \psi}{\Box\varphi \leftrightarrow \Box\psi}$

Nec $\dfrac{\varphi}{\Box\varphi}$

MP $\dfrac{\varphi \quad \varphi \rightarrow \psi}{\psi}$

E is the smallest classical modal logic.

EM is the logic **E** + *Mon*

EC is the logic **E** + *C*

PC Propositional Calculus

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

Mon $\dfrac{\varphi \rightarrow \psi}{\Box\varphi \rightarrow \Box\psi}$

C $(\Box\varphi \wedge \Box\psi) \rightarrow \Box(\varphi \wedge \psi)$

N $\Box\top$

K $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$

RE $\dfrac{\varphi \leftrightarrow \psi}{\Box\varphi \leftrightarrow \Box\psi}$

Nec $\dfrac{\varphi}{\Box\varphi}$

MP $\dfrac{\varphi \quad \varphi \rightarrow \psi}{\psi}$

E is the smallest classical modal logic.

EM is the logic **E** $+$ *Mon*

EC is the logic **E** $+$ *C*

EMC is the smallest regular modal logic

PC Propositional Calculus

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

Mon $\dfrac{\varphi \rightarrow \psi}{\Box\varphi \rightarrow \Box\psi}$

C $(\Box\varphi \wedge \Box\psi) \rightarrow \Box(\varphi \wedge \psi)$

N $\Box\top$

K $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$

RE $\dfrac{\varphi \leftrightarrow \psi}{\Box\varphi \leftrightarrow \Box\psi}$

Nec $\dfrac{\varphi}{\Box\varphi}$

MP $\dfrac{\varphi \quad \varphi \rightarrow \psi}{\psi}$

E is the smallest classical modal logic.

EM is the logic **E** + *Mon*

EC is the logic **E** + *C*

EMC is the smallest regular modal logic

A logic is normal if it contains all instances of *E*, *C* and is closed under *Mon* and *Nec*

PC Propositional Calculus

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

Mon $\dfrac{\varphi \to \psi}{\Box\varphi \to \Box\psi}$

C $(\Box\varphi \wedge \Box\psi) \to \Box(\varphi \wedge \psi)$

N $\Box\top$

K $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

RE $\dfrac{\varphi \leftrightarrow \psi}{\Box\varphi \leftrightarrow \Box\psi}$

Nec $\dfrac{\varphi}{\Box\varphi}$

MP $\dfrac{\varphi \quad \varphi \to \psi}{\psi}$

E is the smallest classical modal logic.

EM is the logic **E** + *Mon*

EC is the logic **E** + *C*

EMC is the smallest regular modal logic

K is the smallest normal modal logic

PC Propositional Calculus

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

Mon $\dfrac{\varphi \rightarrow \psi}{\Box\varphi \rightarrow \Box\psi}$

C $(\Box\varphi \wedge \Box\psi) \rightarrow \Box(\varphi \wedge \psi)$

N $\Box\top$

K $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$

RE $\dfrac{\varphi \leftrightarrow \psi}{\Box\varphi \leftrightarrow \Box\psi}$

Nec $\dfrac{\varphi}{\Box\varphi}$

MP $\dfrac{\varphi \qquad \varphi \rightarrow \psi}{\psi}$

E is the smallest classical modal logic.

EM is the logic **E** + *Mon*

EC is the logic **E** + *C*

EMC is the smallest regular modal logic

K = **EMCN**

PC Propositional Calculus

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

Mon $\dfrac{\varphi \to \psi}{\Box\varphi \to \Box\psi}$

C $(\Box\varphi \land \Box\psi) \to \Box(\varphi \land \psi)$

N $\Box\top$

K $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

RE $\dfrac{\varphi \leftrightarrow \psi}{\Box\varphi \leftrightarrow \Box\psi}$

Nec $\dfrac{\varphi}{\Box\varphi}$

MP $\dfrac{\varphi \quad \varphi \to \psi}{\psi}$

E is the smallest classical modal logic.

EM is the logic **E** + *Mon*

EC is the logic **E** + *C*

EMC is the smallest regular modal logic

K $= PC(+E) + K + Nec + MP$