(59

MAIN TRACK:
LOGICS OF AGENCY

(60

LOGICS OF KNOWLEDGE AND BELIEF

- logics of knowledge
  = epistemic logics
- logics of belief
  = doxastic logics

- Knowledge (naively) understood as true belief.

- Main problem with simplistic approaches: logical omniscience (agents are assumed to have a perfect reasoning capability).

(61

Traditional approach

(Hintikka, ...)

- Use the multi-modal logic determined by all
  reflexive (knowledge) [serial (belief)],
  transitive,
  euclidean
  multi-relational frames (one relation $R(i)$ for each agent $i \in I$).
- This is (soundly & completely) axiomatized by
  $\mathbb{K}_I(T45)_{i\in I} = S5_I$ $[\mathbb{K}_I(D45)_{i\in I}]$.
- Define $K_i A$ $[B_i A]$ (i knows [believes] that A) as $[i]A$.

(62

reflexivity:	T:	„knowledge axiom":
$\begin{tikzcd} w \arrow[loop right, dashed, "i"] \end{tikzcd}$	$[i]A \supset A$	what an agent knows, is true
seriality:	D:	„belief axiom":
$w \xrightarrow{i} w'$	$[i]A \supset \langle i \rangle A$ $([i]A \supset \neg[i]\neg A,$ $\neg([i]A \wedge [i]\neg A))$	an agent does not believe the negation of what it believes

knowledge — consistent with the reality
beliefs — internally consistent

63

transitivity:

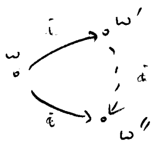
4:

$[i]A \supset [i][i]A$

<u>positive introspection</u>:

an agent knows [believes) it knows [believes] what it knows [believes]

euclideanness:



5:

$\langle i \rangle A \supset [i]\langle i \rangle A$

$(\neg[i]A \supset [i]\neg[i]A)$

<u>negative introspection</u>:

an agent knows [believes] it doesn't know [believe] what it doesn't know [believe]

(Do we always want it??)

64

<u>logical omniscience</u>

K:

$[i](A \supset B) \supset ([i]A \supset [i]B)$

an agent's knowledge [belief] is closed under implications it knows [believes]

RN:

$$\frac{A}{[i]A}$$

an agent knows [believes] all tautologies

- Example (of the power of these things!):

Suppose

$\neg K_a q$ — Alice doesn't know that q.

Then

$K_a \neg K_a q$ — Alice knows she doesn't know that q (since she can do neg. introspection).

$K_a \neg K_b K_a q$ — Alice knows that Bob doesn't know she knows that q (Bob's knowledge is consistent with the reality, to Alice's knowledge).

$K_a K_b \neg K_b K_a q$ — Alice knows that Bob in fact knows he doesn't know she knows that q (Alice knows Bob can do neg. introspection).

Common knowledge [beliefs]

- Introduce a pseudo-agent c ("any fool").
- Add this frame condition:

for any $w, w' \in W$,
$w\,R(c)\,w'$ iff, for some $\overbrace{w_0, \ldots}^{n \geq 1,}, w_n \in W$, $i_0, \ldots, i_{n-1} \in I$,
$w = w_0$, $w_0\,R(i_0)\,w_1, \ldots, w_{n-1}\,R(i_{n-1})\,w_n$, $w_n = w'$.

- Define $K_c A$ [$B_c A$] (it is commonly known [believed] that A, any fool knows [believes] that A) as $[c]A$.

- c behaves quite well as an agent:

  From the assumption that all R_i's ($i \in I$) are reflexive [serial], it follows that so is $R(c)$; besides $R(c)$ is clearly transitive.

  Check this!

- The naive logic of knowledge [belief] with "any fool" is (soundly & completely) axiomatized by $\mathbb{K}_I(T45)_{i \in I}$ [$\mathbb{K}_I(D45)_{i \in I}$] plus

$$[c]A \supset \bigwedge_{i \in I} [i](A \wedge [c]A)$$

$$\frac{B \supset \bigwedge_{i \in I} [i](A \wedge B)}{B \supset [c]A}$$

68

Distributed knowledge

- Introduce a pseudo-agent d ("a wise man").

- Add this frame condition:

  for any $w, w' \in W$,

  $w\,R(d)\,w'$ iff, for all $i \in I$, $w\,R(i)\,w'$.

- Define $K_d A$ (a wise man would know that A) as $[d]A$.

- d behaves very well as an agent wrt. knowledge: From the assumption that all R_i's ($i \in I$) are reflexive, transitive and euclidean, it follows that so is $R(d)$.

69

- The naive logic of knowledge with "a wise man" is (soundly & completely) axiomatized by $\mathbb{K}_{I \cup \{d\}}(T45)_{i \in I \cup \{d\}}$ plus

$$[i]A \supset [d]A,$$

if $|I| \geq 2$.

Other approaches

- We shall consider three "finer" approaches:
  * Fagin and Halpern's logic of (general) awareness;
  * Fagin and Halpern's logic of local reasoning;
  * Levesque's logic of only-believing (Duc's version of this).

logic of (general) awareness

(Fagin, Halpern)

- In addition to the usual notion of implicit belief, captures a notion of explicit belief based on the notion of awareness.

Syntax

Assumed are a denumerable set $Fma_0 = \{p_0, p_1, \dots\}$ of prop. letters and a (non-empty) set I of agent identifiers. The set Fma of formulae is defined as follows:

- if $p \in Fma_0$, then $p \in Fma$;
- $\top, \bot \in Fma$,
- if $A \in Fma$, then $\neg A \in Fma$;
- ...
- if $A \in Fma$, then $A_i A$ (i is aware of 'A'), $B_i A$ (i believes A implicitly), $B_i^{exp} A$ (i believes A explicitly) $\in Fma$.

(72)

Structures for (general) awareness

A structure for (general) awareness is a quadruple $M = (W, R, A, V)$ where

* W is a non-empty set of <u>worlds</u>;
* $R \in [I \to \mathcal{P}(W \times W)]$ is a function from agent identifiers to <u>serial</u>, <u>transitive</u> and <u>euclidean</u> relations between W and W (<u>accessibility</u> relations);
* $A \in [I \to [Fma \to \mathcal{P}(W)]]$ is a function from agent identifiers to functions from formulae to subsets of W (<u>awareness</u> functions);
* $V \in [Fma_0 \to \mathcal{P}(W)]$ is a function from prop. letters to subsets of W (<u>valuation</u>)

A formula is a tautology in the logic of (general) awareness, if it is valid in all structures for (general) awareness.

(73)

Satisfaction in structures for (general) awareness

- Given a structure $M = (W, R, A, V)$ for (general) awareness, the <u>interpretation</u> function $[\![-]\!]^M \in [Fma \to \mathcal{P}(W)]$ is defined as follows ($\models^M_w A$ is short for $w \in [\![A]\!]^M$):

  * if $p \in Fma_0$, then: $\models^M_w p$ iff $w \in V(p)$;
  * $\models^M_w \top$ is true;
  * $\models^M_w \bot$ is false;
  * $\models^M_w \neg A$ iff not $\models^M_w A$;
  * ... ;
  * $\models^M_w \underline{A}_i A$ iff <u>$w \in A_i(A)$</u>;
  * $\models^M_w \underline{B}_i A$ iff, for any $w' \in W$ st $wR_i w'$, $\models^M_{w'} A$;
  * $\models^M_w \underline{B}^{exp}_i A$ iff <u>$w \in A_i(A)$</u> <u>and</u> for any $w' \in W$ st $wR_i w'$, $\models^M_{w'} A$.

74

Observations

- The following hold for any structure M for (general) awareness:

  * $\models^M \underline{B}_i^{exp} A \equiv \underline{A}_i A \wedge \underline{B}_i A$;

  * $\models^M \underline{B}_i^{exp} (A \supset B) \supset (\underline{B}_i^{exp} A \wedge \underline{A}_i B \supset \underline{B}_i^{exp} B)$;
  * If $\models^M A$, then $\models^M \underline{A}_i A \supset \underline{B}_i A$;

  * $\models^M \neg (\underline{B}_i^{exp} A \wedge \underline{B}_i^{exp} \neg A)$;
  * $\models^M \underline{B}_i^{exp} A \wedge \underline{A}_i \underline{B}_i^{exp} A \supset \underline{B}_i^{exp} \underline{B}_i^{exp} A$;
  * $\models^M \neg \underline{B}_i^{exp} A \wedge \underline{A}_i \neg \underline{B}_i A \supset \underline{B}_i^{exp} \neg \underline{B}_i^{exp} A$.

75

Axiomatization

- The logic of (general) awareness is (soundly & completely axiomatized by $\mathbb{K}_I$ plus

$$\underline{B}_i^{exp} A \equiv \underline{A}_i A \wedge \underline{B}_i A$$

Logic of local reasoning

(Fagin, Halpern)

- Agents ≈ „societies of minds", each mind with its own beliefs, possibly contradicting those of the fellow minds.

Syntax

Assumed again are a denumerable set $Fma_0 = \{p_0, p_1, \ldots\}$ of prop. letters and a (non-empty) set I of agent identifiers. The set Fma of formulae is defined as follows:

- if $p \in Fma_0$, then $p \in Fma$;
- $\top, \bot \in Fma$,
- if $A \in Fma$, then $\neg A \in Fma$;
- ...
- if $A \in Fma$, then $\underline{B}_i A$ (i believes A in some frame-of-mind), $\underline{B}_i^{imp} A$ (i believes A implicitly, "between" its different frames-of-mind)

(77)

Structures for local reasoning

A structure for local reasoning is a triple $M = (W, N, V)$ where

* W is a non-empty set of worlds;
* $N \in [I \to \mathcal{P}(W \times \mathcal{P}(W))]$ is a function from agent identifiers to relations between worlds and sets of worlds (neighbourhood relations) such that
  + for any $w \in W$, there exists an $X \subseteq W$ st wNX;
  + for no $w \in W$, $wN\emptyset$;
* $V \in [Fma_0 \in \mathcal{P}(W)]$ is a function from agent identifiers to subsets of W (valuation).

(78)

Satisfaction in structures for local reasoning

Given a structure $M = (W, N, V)$ for local reasoning, the interpretation function $[\![-]\!]^M \in [Fma \to \mathcal{P}(W)]$ for this structure is defined as follows ($\models^M_w$ is short for $[\![-]\!]^M$):

* if $p \in Fma_0$, then: $\models^M_w p$ iff $w \in V(p)$;
* $\models^M_w \top$ is true;
* $\models^M_w \bot$ is false;
* $\models^M_w \neg A$ iff not $\models^M_w A$;
* $\dots$;
* $\models^M_w \underline{B}_i A$ iff, for some X st wNX, for any $w' \in X$, $\models^M_{w'} A$;
* $\models^M_w \underline{B}_i^{imp} A$ iff, for any $w' \in \bigcap\{X \mid wNX\}$, $\models^M_{w'} A$.

Observations

- For any structure M for local reasoning, the following hold:
  - $*$ if $\models^M A \supset B$, then $\models^M \underline{B}_i A \supset \underline{B}_i B$;
  - $*$ $\models^M \underline{B}_i \top$; if $\models^M A$, then $\models^M \underline{B}_i A$
  - $*$ $\models^M \neg \underline{B}_i \bot$;
  - $*$ $\models^M \underline{B}_i A \supset B_i^{imp} A$.
- There is a structure M st $\not\models^M \underline{B}_i (p \supset q) \supset (\underline{B}_i p \supset \underline{B}_i q)$.
- There is a structure M st $\not\models^M \neg (\underline{B}_i p \wedge \underline{B}_i \neg p)$ and $\not\models^M \neg B_i^{imp} \bot$.

Axiomatization

The logic of local reasoning is (soundly & completely) axiomatized by $E_I (M_i N_i P_i)_{i \in I}$, ie, E_I together with

$$\left.\begin{array}{ll} M_i: & \underline{B}_i (A \wedge B) \supset \underline{B}_i A \wedge \underline{B}_i B \\ N_i: & \underline{B}_i \top \\ P_i: & \neg \underline{B}_i \bot \end{array}\right\} \quad (i \in I)$$

plus

$$\underline{B}_i A \supset \underline{B}_i^{imp} A \qquad (i \in I)$$

(if $|I| \geqslant 2$).

Logic of only-believing

(Levesque / Duc)

- Besides the usual believing-that (believing-at-least-that) operator, study a believing-that-and-only-that (believing-just-that) operator.
- For simplicity, assume just one agent.

Syntax

From a denumerable set $Fma_0 = \{p_0, p_1, \ldots\}$ of prop. letters, define the set Fma of formulae as follows:

- if $p \in Fma_0$, then $p \in Fma$;
- $\top, \bot \in Fma$;
- if $A \in Fma$, then $\neg A \in Fma$;
- ...
- if $A \in Fma$, then $\underline{B}\,A$ (the system believes at least A), $\underline{N}\,A$ (the system believes-not at most A) $\in Fma$.

For $\underline{B}A \wedge \underline{N}\neg A$ (the system believes exactly A), write $\underline{O}\,A$.

Structures for only-believing

A structure for only-believing is a (uni)-relational structure, $M = (W, R, V)$ which is transitive and euclidean (note that seriality is not required).

Satisfaction in structures for only-believing

For a structure $M = (W, R, V)$ for only-believing, the interpretation function $[\![-]\!]^M \in [Fm_0 \to \mathcal{P}(W)]$ is defined as follows:

* if $p \in Fm_{a0}$, then $\models^M_w p$ iff $w \in V(p)$;
* $\models^M_w \top$ is true;
* $\models^M_w \bot$ is false;
* $\models^M_w \neg A$ iff $\not\models^M_w A$;
* ...;
* $\models^M_w \underline{B} A$ iff, for any w' st wRw', $\models^M_{w'} A$;
* $\models^M_w \underline{N} A$ iff, for any w' st $w \not R w'$, $\models^M_{w'} A$.

Some observations

- For any structure M for only-believing, the following holds:

  * if $\not\models^M A$, then $\models^M NA \supset \neg BA$
    (ie, $\models^M \neg(NA \wedge BA)$)

  * if $\not\models^M A$, then
    $\underline{O}(\neg \underline{B}A \supset \neg A) \supset B\neg A$

  * $\models^M \underline{O}((\neg BA \supset \neg A) \wedge A) \supset BA$

Axiomatization

The logic of only-believing is (soundly & completely) axiomatized by $\mathbb{K}45$ plus

$$\underline{N}(A \supset B) \supset (\underline{N}A \supset \underline{N}B)$$

$$\frac{A}{\underline{N}A}$$

$$\neg \underline{S} \neg (\underline{B}A \wedge \underline{N}B) \supset \underline{S}'(A \vee B)$$

where $\underline{S}$, $\underline{S}'$ are arbitrary finite (possibly empty) sequences of B's and N's

(the Humberstone axiom scheme)