Probabilistic Unawareness

Mikaël Cozic

IHPST (Paris I-ENS Ulm-CNRS), GREGHEC (HEC-CNRS) & DEC (ENS Ulm)

IHPST-Tilburg-LSE Workshop
07/XII/2007

doxastic models

- two main families of doxastic models:

(i) **epistemic logic** for *full beliefs*:
- Pierre believes that ϕ
- Pierre believes that $\neg\phi$
- Pierre neither believes that ϕ nor believes that $\neg\phi$

(ii) **probability** for *partial beliefs*:
- Pierre believes that ϕ to degree *r*

doxastic models

- two main families of doxastic models:

(i) **epistemic logic** for *full beliefs*:
- Pierre believes that ϕ
- Pierre believes that $\neg\phi$
- Pierre neither believes that ϕ nor believes that $\neg\phi$

(ii) **probability** for *partial beliefs*:
- Pierre believes that ϕ to degree *r*

- it is well known that epistemic logic suffers from two main cognitive idealizations:

(i) **logical omniscience** (LO): closure under logical consequence, substituability of logically equivalent formulas, etc.

(ii) **full awareness**(FA): full understanding of the state space

a plea for impossible states

- it is rarely recognized that *probabilistic models of beliefs suffer from (LO) and (FA) as well*

a plea for impossible states

- it is rarely recognized that *probabilistic models of beliefs suffer from (LO) and (FA) as well*
- **broad question**: how to weaken (LO) and (FA) in models of partial beliefs ?
- **broad aim:** defend the impossible states (or worlds) approach as a unifying way to weaken cognitive idealizations = **a plea for impossible states**

Outline of the talk

what is (un)awareness ?

- *Modica & Rustichini 1999:*
  • "ignorance about the state space"
  • "some of the facts that determine which state of nature occurs are not present in the subject's mind"
  • "the agent does not know, does not know that she does not know, does not know that she does not know that she does not know, and so on..."
- *Heifetz, Meier & Schipper 2007b:*
  • "Unawareness refers to lack of conception rather than to lack of information."

example

- Pierre plans to rent a house for the holiday; three main factors from the modeler point of view:
- p: the house is no more than 1 km far from the sea
- q: the house is no more than 1 km far from a bar
- r: the house is no more than 1 km far from an airport

example

- Pierre plans to rent a house for the holiday; three main factors from the modeler point of view:
  - p: the house is no more than 1 km far from the sea
  - q: the house is no more than 1 km far from a bar
  - r: the house is no more than 1 km far from an airport
- "simple", factual, ignorance of r: Pierre doesn't know whether there is an airport no more than 1 km far from the house - there are both r-states and $\neg r$-states which are epistemically accessible

example

- Pierre plans to rent a house for the holiday; three main factors from the modeler point of view:
  - p: the house is no more than 1 km far from the sea
  - q: the house is no more than 1 km far from a bar
  - r: the house is no more than 1 km far from an airport
- "simple", factual, ignorance of r: Pierre doesn't know whether there is an airport no more than 1 km far from the house - there are both r-states and $\neg r$-states which are epistemically accessible
- unawareness: Pierre doesn't ask to himself: "is there an airport no more than 1 km far from the house?"
  [See states in *small worlds* in Savage (1954/72)]

example, *cont.*

- the possibility that r is not simply excluded: it is out of Pierre's state space
- modeler's point of view:

pqr	$p\neg qr$	$pq\neg r$	$p\neg q\neg r$
$\neg pqr$	$\neg p\neg qr$	$\neg pq\neg r$	$\neg p\neg q\neg r$

- Pierre's point of view:

pq	$p\neg q$
$\neg pq$	$\neg p\neg q$

[See Savage: "...a smaller world is derived from a larger by neglecting some distinctions between states"]

properties of (un)awareness

- some intuitive properties of (un)awareness:

$A\phi \leftrightarrow A\neg\phi$	(Symmetry)
$A(\phi \wedge \psi) \leftrightarrow A\phi \wedge A\psi$	
$A\phi \leftrightarrow AA\phi$	(Self-Reflection)
$U\phi \rightarrow UU\phi$	(U-introspection)
$U\phi \rightarrow \neg B\phi \wedge \neg B\neg B\phi$	(Plausibility)
$U\phi \rightarrow (\neg B)\phi^n \; \forall n \in N$	(Strong Plausibility)
$\neg BU\phi$	(BU-introspection)

the modeling of (un)awareness

- it is impossible to devise a non-trivial (un)awareness operator that satisfies (most of) the intuitively appealing properties above mentioned
- for instance, in *"Standard State-Space Models Preclude Unawareness"* (1998) , Dekel, Lipman & Rustichini show that it is impossible to have

(i) a non-trivial awareness operator which satisfies Plausibility, U-introspection and BU-introspection

(ii) a belief operator which satisfies either Necessitation or Monotonicity

main models of unawareness

- two main ways to circumvent the issue:

(i) endogenous characterization: awareness defined in terms of beliefs : Modica & Rustichini (1999), Heifetz, Meier & Schipper (2006), (2007a)

$$\mathcal{M}, s \models A\phi \Leftrightarrow \mathcal{M}, s \models B\phi \lor B\neg B\phi$$

main models of unawareness

- two main ways to circumvent the issue:

(i) endogenous characterization: awareness defined in terms of beliefs : Modica & Rustichini (1999), Heifetz, Meier & Schipper (2006), (2007a)

$$\mathcal{M}, s \models A\phi \Leftrightarrow \mathcal{M}, s \models B\phi \vee B\neg B\phi$$

(ii) exogenous characterization: Fagin & Halpern (1988), Halpern (2001) awareness© structures

$$\mathcal{M}, s \models A\phi \Leftrightarrow \phi \in \mathcal{A}(s)$$

where $\mathcal{A} : S \rightarrow \wp(\mathcal{L}(At))$

GSM structures, example

- state space S based on $At = \{p, q, r\}$:

pqr	$p\neg qr$	$pq\neg r$	$p\neg q\neg r$
$\neg pqr$	$\neg p\neg qr$	$\neg pq\neg r$	$\neg p\neg q\neg r$

- in actual state $s = pqr$, Pierre believes that p, does not know whether q and is unaware of r ; his non-standard state space $S_{\{p,q\}}$ and accessibility correspondance in s are

$[pq]$	$p\neg q$
$\neg pq$	$\neg p\neg q$

GSM structures, example

pqr	$p\neg qr$	$pq\neg r$	$p\neg q\neg r$
$\neg pqr$	$\neg p\neg qr$	$\neg pq\neg r$	$\neg p\neg q\neg r$

$[pq]$	$p\neg q$
$\neg pq$	$\neg p\neg q$

- suppose that pqr in projected in pq: $\rho(pqr) = pq$
  • if pqr and $pq\neg r$ are projected in pq, pqr and $pq\neg r$ agree on p and q
  • if pqr and $pq\neg r$ are projected in pq, $R(pqr) = R(pq\neg r)$
  • $R(pqr) \subseteq S_{\{p,q\}}$

GSM structures

A GSM structure is a t-uple $\mathcal{M} = (S, S', \pi, R, \rho)$

(i) S is a state space

(ii) $S' = \bigcup_{X \subseteq At} S'_X$ (where S'_X are disjoint) is a (non-standard) state space

(iii) $\pi : At \times S \rightarrow \{0, 1\}$ is a valuation for S

(iv) $R : S \rightarrow \wp(S')$ is an accessibility correspondence

GSM structures

A GSM structure is a t-uple $\mathcal{M} = (S, S', \pi, R, \rho)$

(i) S is a state space

(ii) $S' = \bigcup_{X \subseteq At} S'_X$ (where S'_X are disjoint) is a (non-standard) state space

(iii) $\pi : At \times S \rightarrow \{0, 1\}$ is a valuation for S

(iv) $R : S \rightarrow \wp(S')$ is an accessibility correspondence

(v) $\rho : S \rightarrow S'$ is a projection s.t. (1) if $\rho(s) = \rho(t) \in S'_X$, then (a) for each atomic formula $p \in X$, $\pi(s, p) = \pi(t, p)$ and (b) $R(s) = R(t)$ and (2) if $\rho(s) \in S'_X$, then $R(s) \subseteq S'_X$

GSM structures

A GSM structure is a t-uple $\mathcal{M} = (S, S', \pi, R, \rho)$

(i) S is a state space

(ii) $S' = \bigcup_{X \subseteq At} S'_X$ (where S'_X are disjoint) is a (non-standard) state space

(iii) $\pi : At \times S \rightarrow \{0, 1\}$ is a valuation for S

(iv) $R : S \rightarrow \wp(S')$ is an accessibility correspondence

(v) $\rho : S \rightarrow S'$ is a projection s.t. (1) if $\rho(s) = \rho(t) \in S'_X$, then (a) for each atomic formula $p \in X$, $\pi(s, p) = \pi(t, p)$ and (b) $R(s) = R(t)$ and (2) if $\rho(s) \in S'_X$, then $R(s) \subseteq S'_X$

- one can extend R and π to the whole state space with π^*: if $s' \in S'_X$, then $\pi^*(s', p) = 1$ iff (a) $p \in X$ and (b) for all $s \in \rho^{-1}(s'), \pi(s, p) = 1$.

satisfaction relation

- one may then define as follows the satisfaction relation for each $s^* \in S^* = S \cup S'$ (Halpern's 2001 version):

(i) $\mathcal{M}, s^* \vDash p$ iff $\pi^*(s^*, p) = 1$

(ii) $\mathcal{M}, s^* \vDash \phi \wedge \psi$ iff $\mathcal{M}, s^* \vDash \phi$ and $\mathcal{M}, s^* \vDash \psi$

satisfaction relation

- one may then define as follows the satisfaction relation for each $s^* \in S^* = S \cup S'$ (Halpern's 2001 version):

(i) $\mathcal{M}, s^* \models p$ iff $\pi^*(s^*, p) = 1$

(ii) $\mathcal{M}, s^* \models \phi \wedge \psi$ iff $\mathcal{M}, s^* \models \phi$ and $\mathcal{M}, s^* \models \psi$

(iii) $\mathcal{M}, s^* \models \neg\phi$ **iff** $\mathcal{M}, s^* \not\models \phi$ **and (1) either** $s^* \in S$**, (2) or** $s^* \in S'_X$ **and** $\phi \in \mathcal{L}(X)$

satisfaction relation

- one may then define as follows the satisfaction relation for each $s^* \in S^* = S \cup S'$ (Halpern's 2001 version):

(i) $\mathcal{M}, s^* \vDash p$ iff $\pi^*(s^*, p) = 1$

(ii) $\mathcal{M}, s^* \vDash \phi \wedge \psi$ iff $\mathcal{M}, s^* \vDash \phi$ and $\mathcal{M}, s^* \vDash \psi$

(iii) $\mathcal{M}, s^* \vDash \neg\phi$ **iff** $\mathcal{M}, s^* \nvDash \phi$ **and (1) either** $s^* \in S$**, (2) or** $s^* \in S'_X$ **and** $\phi \in \mathcal{L}(X)$

(iv) $\mathcal{M}, s^* \vDash B\phi$ iff for each $t^* \in R^*(s^*)$, $\mathcal{M}, t^* \vDash \phi$

- crucial point: (iii) introduces partiality: if $p \notin X$ and $s^* \in S'_X$ then neither $\mathcal{M}, s^* \vDash p$ nor $\mathcal{M}, s^* \vDash \neg p$ (for short, $\mathcal{M}, s^* \Uparrow p$). More generally,

$$\mathcal{M}, s^* \Downarrow \phi \text{ for } s^* \in S'_X \text{ iff } \phi \in \mathcal{L}(X)$$

what (un)awareness is not

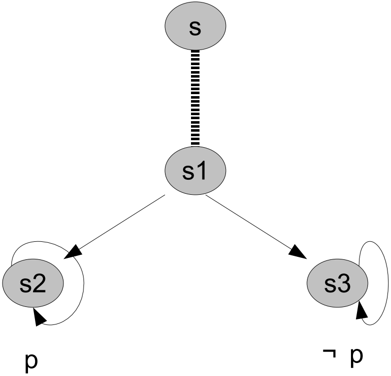
- MR 1999 and HMS 2006 define unawareness in terms of beliefs; this is OK *given their assumption that accessibility correspondences are partitional*

what (un)awareness is not

- MR 1999 and HMS 2006 define unawareness in terms of beliefs; this is OK *given their assumption that accessibility correspondences are partitional*
- **but** this endogenous characterization is not robust under change of the accessibility correspondences' properties. In the general case,
  • it is plausible that if Pierre is unaware of ϕ, he doesn't believe that ϕ nor believe that he doesn't believe that ϕ ($U\phi \rightarrow (\neg B\phi \wedge \neg B\neg B\phi)$)
  • it is not plausible that if Pierre doesn't believe that ϕ nor believe that he doesn't believe that ϕ, he is necessarily unaware of ϕ ($(\neg B\phi \wedge \neg B\neg B\phi) \rightarrow U\phi$)

example

$\mathcal{M}, s \vDash U^{MR}p$ but $\mathcal{M}, s \vDash B(B\neg p \vee Bp)$



unawareness as partiality

- hence: keep the underlying GSM structure but change the definition of (un)awareness

unawareness as partiality

- hence: keep the underlying GSM structure but change the definition of (un)awareness
- the possible states that Pierre conceives do not “answer” to $?p$, $?q$ and $?r$: they answer only to $?p$ and $?q$
- unawareness may be seen as **partiality**: the possible states that Pierre conceives make true neither r nor $\neg r$

unawareness as partiality

- hence: keep the underlying GSM structure but change the definition of (un)awareness
- the possible states that Pierre conceives do not “answer” to $?p$, $?q$ and $?r$: they answer only to $?p$ and $?q$
- unawareness may be seen as **partiality**: the possible states that Pierre conceives make true neither r nor $\neg r$
- semantic characterization of unawareness in terms of **partiality**:

$$\mathcal{M}, s \vDash A\phi \text{ iff } \mathcal{M}, \rho(s) \Downarrow \phi$$

- Let’s call a P-GSM structure a GSM structure where the truth conditions of the unawareness operator are given in terms of partiality

BU-introspection and seriality

- BU-introspection, viz $\neg BU\phi$, is not valid in P-GSM structures
- why ? Because in the degenerate case where $R(s) = \emptyset$, then for all ψ, $\mathcal{M}, s \vDash B\psi$
- suffices to require that $R(s) \neq \emptyset$ (seriality) for $\neg BU\phi$ to be valid
- **serial P-GSM structures** (my final proposal) implies the validity of $B\phi \rightarrow A\phi$
- caution : seriality does not correspond to

  $$\text{(D) } B\phi \rightarrow \neg B\neg\phi$$

  (as in Kripke structures) but to

  $$(\text{D}_U)\ B\phi \rightarrow (\neg B\neg\phi \wedge A\phi)$$

axiom system KD_U

(PROP) Instances of propositional tautologies
(MP) From ϕ and $\phi \rightarrow \psi$ infer ψ

(K) $B\phi \wedge B(\phi \rightarrow \psi) \rightarrow B\psi$
(Gen) From ϕ infer $A\phi \rightarrow B\phi$

(D_U) $B\phi \rightarrow (\neg B\neg\phi \wedge A\phi)$
(A1) $A\phi \leftrightarrow A\neg\phi$
(A2) $A(\phi \wedge \psi) \leftrightarrow (A\phi \wedge A\psi)$
(A3) $A\phi \leftrightarrow AA\phi$
(A4) $AB\phi \leftrightarrow A\phi$
(A5) $A\phi \rightarrow BA\phi$
(Irr) If no atomic formulas in ϕ appear in ψ, from $U\phi \rightarrow \psi$ infer ψ

explicit probabilistic structures (EPS)

- **Definition** : the set of formulas of an explicit probabilistic language $\mathcal{L}L(At)$ based on a set At of propositional variables, $Form(\mathcal{L}L(At))$ is defined by :

$$\phi ::= p|\neg\phi|\phi \vee \psi|L_a\phi$$

  where $p \in At$ and $a \in [0,1] \subseteq \mathbb{Q}$.
  $L_a\phi$: P. believes that ϕ **at least to degree** a

explicit probabilistic structures (EPS)

- **Definition** : the set of formulas of an explicit probabilistic language $\mathcal{L}L(At)$ based on a set At of propositional variables, $Form(\mathcal{L}L(At))$ is defined by :

$$\phi ::= p|\neg\phi|\phi \vee \psi|L_a\phi$$

  where $p \in At$ and $a \in [0, 1] \subseteq \mathbb{Q}$.
  $L_a\phi$: P. believes that ϕ **at least to degree** a
- **Definition** : an **explicit probabilistic structure** for $\mathcal{L}L_a(At)$ is a 3-tuple $\mathcal{M} = (S, \pi, P)$ where $P : S \rightarrow \Delta(S)$ assigns to every state a probability distribution on the state space.
- Satisfaction condition for L_a :

$$\bar{\pi}(s, L_a\phi) = 1 \text{ iff } P(s)([[\phi]]) \geq a$$

EPS, *cont.*

- Some Remarks on EPS :

(i) this language is the one proposed by Aumann (1999) and Heifetz and Mongin (2001). Fagin, Halpern and Meggido 1990 and Halpern 2003 use a different language (linear inequalities)

EPS, *cont.*

- Some Remarks on EPS :

(i) this language is the one proposed by Aumann (1999) and Heifetz and Mongin (2001). Fagin, Halpern and Meggido 1990 and Halpern 2003 use a different language (linear inequalities)

(ii) EPS correspond to the *type spaces* used in games of incomplete information, in the same way that Kripke structures (with R as an equivalence relation) corresponds to the information partitions

EPS, *cont.*

- Some Remarks on EPS :

(i) this language is the one proposed by Aumann (1999) and Heifetz and Mongin (2001). Fagin, Halpern and Meggido 1990 and Halpern 2003 use a different language (linear inequalities)

(ii) EPS correspond to the *type spaces* used in games of incomplete information, in the same way that Kripke structures (with R as an equivalence relation) corresponds to the information partitions

(iii) from the explicit probabilistic language, one can define

- $M_a\phi = L_{1-a}\neg\phi$ (P. believes **at most** to degree a that ϕ)
- $E_a\phi = M_a\phi \wedge L_a\phi$ (P. believes **exactly** to degree a that ϕ)

EPS, *cont.*

- EPS is not compact :
  - $\Sigma = \{L_{1/2-1/n}\phi : n \geq 2, n \in \mathbb{N}\}$
  - $\psi = \neg L_{1/2}\phi$

  Clearly,
  - for each finite $\Sigma' \subset \Sigma$, $\Sigma' \cup \psi$ is coherent
  - but $\Sigma \cup \psi$ is not coherent
- consequence : no strong completeness (Meier (2001) provides a strong completeness proof with an *infinitary* language)

EPS, axiomatization

System HM (Heifetz and Mongin 2001)

(PROP) Instances of propositional tautologies
(MP) From ϕ and $\phi \rightarrow \psi$ infer ψ

(A1) $L_0\phi$
(A2) $L_a\top$
(A5)$L_a\phi \rightarrow \neg L_b\neg\phi$ $(a + b > 1)$
(Def M) $M_a\phi \leftrightarrow L_{1-a}\neg\phi$
(A8)$\neg L_a\phi \rightarrow M_a\phi$
(RE) From $\phi \leftrightarrow \psi$ infer $L_a\phi \leftrightarrow L_a\psi$

(B) From $((\phi_1, ..., \phi_m) \leftrightarrow (\psi_1, ..., \psi_n))$ infer
$((\bigwedge_{i=1}^{m} L_{ai}\phi_i) \wedge (\bigwedge_{j=2}^{n} M_{bj}\psi_j) \rightarrow L_{(a1+...+am)-(b1+...+bn)}\psi_1)$

EPS, axiomatization, *cont.*

- some theorems :
  (RN) From $\phi \rightarrow \psi$ infer $L_a\phi \rightarrow L_a\psi$
  (A2+) $\neg L_a\bot$ $(a > 0)$
  (A3) $L_a(\phi \wedge \psi) \wedge L_b(\phi \wedge \neg\psi) \rightarrow L_{a+b}\phi$ $(a + b \leq 1)$ (finite superadditivity)
  (A4)$\neg L_a(\phi \wedge \psi) \wedge \neg L_b(\phi \wedge \neg\psi) \rightarrow \neg L_{a+b}\phi$ $(a + b \leq 1)$ (finite subadditivity)
  (A7) $L_a\phi \rightarrow L_b\phi$ $(b < a)$

the rule (B)

- the sum of the probabilities of the members of two partitions of an event are equal: let
  • $E_1 \cup ... \cup E_m = A$ with E_i pairwise disjoint
  • $F_1 \cup ... \cup F_n = A$ with F_j pairwise disjoint
  Clearly,

$$\sum_{i=1}^{m} P(E_i) = \sum_{j=1}^{n} P(F_j) = P(A)$$

- Remarks:

(i) $\sum_{i=1}^{m} \mathbf{I}_{E_i} = \sum_{j=1}^{n} \mathbf{I}_{F_j} = \mathbf{I}_A$

(ii) in the case $m = n$, it is clear that if $P(E_i) \geq P(F_i)$ for $1 \leq i < n$, then the two last events have to "compensate" the disequilibrium ie $P(E_n) \leq P(F_n)$

the rule (B), *cont.*

- one does not need to have partitions : the more general condition is that

  $$\sum_{i=1}^{m} \mathbf{I}_{E_i} = \sum_{j=1}^{n} \mathbf{I}_{F_j}$$

  this means that any element of S belongs to as many E_i's as F_j's
- one can think about a "compensation" closer to EPS:
  if
  $P(E_i) \geq \alpha_i$ for $i = 1, ..., n$ and
  $P(F_j) \leq \beta_j$ for $j = 2, ..., m$
  then
  $P(F_1) \geq (\alpha_1 + ... + \alpha_n) - (\beta_2 + ... + \beta_m)$

the rule (B), *cont.*

- example:
  $E_1 = A \cup B$, $F_1 = A$ and $F_2 = B$
  if
  $P(E_1 = A \cup B) \geq 1/2$ and
  $P(F_2 = B) \leq 1/6$,
  then
  $P(F_1 = A) \geq 1/3$

the rule (B), *cont.*

- example:
  $E_1 = A \cup B$, $F_1 = A$ and $F_2 = B$
  if
  $P(E_1 = A \cup B) \geq 1/2$ and
  $P(F_2 = B) \leq 1/6$,
  then
  $P(F_1 = A) \geq 1/3$
- let's come back to the inference rule:
  (B) From $((\phi_1, ..., \phi_m) \leftrightarrow (\psi_1, ..., \psi_n))$ infer
  $((\bigwedge_{i=1}^{m} L_{ai}\phi_i) \wedge (\bigwedge_{j=2}^{n} M_{bj}\psi_j) \rightarrow L_{(a1+...+am)-(b1+...+bn)}\psi_1)$

the rule (B), *cont.*

- **the premiss of (B)** : $((\phi_1, ..., \phi_m) \leftrightarrow (\psi_1, ..., \psi_n))$ is a syntactical rendering of the equality of sums of characteristic functions
- $((\phi_1, ..., \phi_m) \leftrightarrow (\psi_1, ..., \psi_n))$ is an abbreviation for

$$\bigwedge_{k=1}^{max(m,n)} \phi^{(k)} \leftrightarrow \psi^{(k)}$$

  where
  $\phi^{(k)} = \bigvee_{1 \leq l_1 \leq ... \leq l_k < m} (\phi_{l_1} \wedge ... \wedge \phi_{l_k})$ (if $k > m$, by convention $\phi^{(k)} = \bot$)
  $\phi^{(k)}$ says that at least k of the formulas ϕ_i are true
- **the conclusion of (B)** is a direct translation of the “compensation” principle in terms of inequalities

the rule (B), *cont.*

- (B) is a very powerful probabilistic inference rule. It allows e.g. to derive
  (A3) $L_a(\phi \wedge \psi) \wedge L_b(\phi \wedge \neg\psi) \rightarrow L_{a+b}\phi$ $(a + b \leq 1)$ (finite superadditivity)
- Let
  $\phi_1 = \phi \wedge \psi$, $\phi_2 = \phi \wedge \neg\psi$, $\psi_1 = \phi$
  Suffices to notice that $(\phi_1, \phi_2) \leftrightarrow \psi_1)$; hence
  $L_a\phi_1 \wedge L_b\phi_2 \rightarrow L_{a+b}\psi_1$,
  which is nothing but (A3)

probabilistic unawareness, first attempt

- **Definition** : an **P-GSM explicit probabilistic structure** for $\mathcal{L}^{LA}(At)$ is a t-tuple $\mathcal{M} = (S, S', \pi, P, \rho)$ where

(i) S is a state space

(ii) $S' = \bigcup_{\Phi \subseteq At} S'_\Phi$ (where S'_Φ are disjoint) is a state space

(iii) $\pi : At \times S \rightarrow \{0, 1\}$ is a valuation for S

(iv) $P : S \rightarrow \Delta(S')$

(v) $\rho : S \rightarrow S'$ is a projection s.t. (1) if $\rho(s) = \rho(t) \in S'_\Phi$, then (a) for each atomic formula $p \in \Phi$, $\pi(s, p) = \pi(t, p)$ and (b) $P(s) = P(t)$ and (2) if $\rho(s) \in S'_\Phi$, then $Supp(P(s)) \subseteq S'_\Phi$

probabilistic unawareness, first attempt

- some good news:
  for all GSM-EPS $\mathcal{M}$ and all standard state s, unawareness precludes positive probability:
  $\mathcal{M}, s \models U\phi \rightarrow \neg L_a \phi$ for $a > 0$
  $\mathcal{M}, s \models U\phi \rightarrow \neg L_a \neg\phi$ for $a > 0$
  $\mathcal{M}, s \models \neg L_a U\phi$ for $a > 0$

probabilistic unawareness, first attempt

- some good news:
  for all GSM-EPS $\mathcal{M}$ and all standard state s, unawareness precludes positive probability:
  $\mathcal{M}, s \vDash U\phi \rightarrow \neg L_a\phi$ for $a > 0$
  $\mathcal{M}, s \vDash U\phi \rightarrow \neg L_a\neg\phi$ for $a > 0$
  $\mathcal{M}, s \vDash \neg L_a U\phi$ for $a > 0$
- but some (very) bad news:
  for all GSM-EPS $\mathcal{M}$ and all standard state s,
  $\mathcal{M}, s \vDash U\phi \rightarrow L_0\phi$
  $\mathcal{M}, s \vDash U\phi \rightarrow L_0\neg\phi$
  $\mathcal{M}, s \vDash U\phi \rightarrow L_1 L_0\phi$ (!!)

probabilistic unawareness, second attempt

- satisfaction condition for $L_a\phi$:
  $\mathcal{M}, s \models L_a\phi \Leftrightarrow P(s)([[\phi]]) \geq a$ **and** $\mathcal{M}, \rho(s) \Downarrow \phi$

probabilistic unawareness, second attempt

- satisfaction condition for $L_a\phi$:
  $\mathcal{M}, s \models L_a\phi \Leftrightarrow P(s)([[\phi]]) \geq a$ **and** $\mathcal{M}, \rho(s) \Downarrow \phi$
- in this case, the following holds:

$A\phi \leftrightarrow A\neg\phi$	(Symmetry)
$A\phi \leftrightarrow AA\phi$	(Self-Reflection)
$U\phi \rightarrow UU\phi$	(U-introspection)
$U\phi \rightarrow \neg L_a\phi \wedge \neg L_a\neg L_a\phi$	(Plausibility)
$U\phi \rightarrow (\neg L_a)^n\phi\ \forall n \in N$	(Strong Plausibility)
$\neg L_a U\phi$ (a >0)	(L_aU-introspection)
$L_0\phi \leftrightarrow A\phi$	

HM_U axiom system

- an axiom system for explicit probabilistic structures with unawareness:

System HM_U, Part I
(PROP) Instances of propositional tautologies
(MP) From ϕ and $\phi \rightarrow \psi$ infer ψ
(A0) $A\phi \leftrightarrow L_0\phi$
(A1) $A\phi \leftrightarrow A\neg\phi$
(A2) $A(\phi \wedge \psi) \leftrightarrow A\phi \wedge A\psi$
(A3) $A\phi \leftrightarrow AA\phi$
(A4) $A\phi \leftrightarrow AL_a\phi$
(A5$_L$) $A\phi \rightarrow L_1A\phi$

HM_U axiom system

System HM_U, Part II

(A6) $L_a \top$
(A7) $L_a \phi \rightarrow \neg L_b \neg \phi$ $a + b > 1$
(A8$_U$) $(\neg L_a \phi \wedge A\phi) \rightarrow M_a \phi$

(RE$_U$) From $\phi \leftrightarrow \psi$ infer $A(\phi \wedge \psi) \rightarrow (L_a \phi \leftrightarrow L_a \psi)$

(B$_U$) From $((\phi_1, ..., \phi_m) \leftrightarrow (\psi_1, ..., \psi_n))$ infer $((\bigwedge_{i=1}^{m} L_{ai} \phi_i) \wedge (\bigwedge_{j=2}^{n} M_{bj} \psi_j) \rightarrow (A\psi_1 \rightarrow L_{(a1+...+am)-(b1+...+bn)} \psi_1))$

L_aU introspection

- the axiom of L_aU introspection

$$\neg L_a U\phi \; (a > 0)$$

  may be derived from HM_U
- intuitively,
  (1) $L_a\neg A\phi \rightarrow L_0\neg A\phi$
  (2) $L_a\neg A\phi \rightarrow A\neg A\phi$ (from (1) and (A0))
  (3) $L_a\neg A\phi \rightarrow AA\phi$ (from (2) and (A1))
  (4) $L_a\neg A\phi \rightarrow A\phi$ (from (3) and (A3))
  (5) $L_a\neg A\phi \rightarrow L_1A\phi$ (from (4) and ($A5_L$))
  (6) $L_a\neg A\phi \rightarrow \neg L_a\neg A\phi$ (from (5) and (A7))
  (7) $\bot$ (from (6) by propositional reasoning)

counter-example to (B)

- the famous inference rule (B) needs to be restricted with the help of the awareness operator.
- Let
  $\phi_1 = (p \vee \neg p)$
  $\psi_1 = (q \vee \neg q)$
  $\psi_2 = (p \vee \neg p)$
  So, clearly the premiss of (B) - $((\phi_1) \leftrightarrow (\psi_1, \psi_2))$ is satisfied
- Suppose that Pierre is aware of p but not of q ; in this case, it will be true that $L_1\phi_1$ and $M_1\psi_2$. Hence by the compensation principle, $L_0\psi_1$. But this will not be the case if Pierre is unaware of q.

completeness

- some steps remain to be checked in the proof (!), but one may confidently claim the following

Completeness Theorem : $\models_{GSM-EPS} \phi$ iff $\vdash_{HM_U} \phi$

further issues

1. becoming aware
2. multi-agent unawareness
3. applications to decision theory and game theory

becoming aware

- Pierre may be initially unaware of ϕ and become aware of ϕ :

(1) when someone gives Pierre an information that involves ϕ

(2) when someone asks Pierre what he thinks about ϕ

becoming aware

- Pierre may be initially unaware of ϕ and become aware of ϕ :

(1) when someone gives Pierre an information that involves ϕ

(2) when someone asks Pierre what he thinks about ϕ

- when one thinks about (1) for full beliefs, things may look simple:

- initially, Pierre is only aware of p, neither q nor r: $\rho(s) \in S_{\{p\}}$ and $R(s) \subseteq S_{\{p\}}$. He believes that p

p
$\neg p$

- Pierre is informed that q
  (i) first, the structure is modified such that $\rho'(s) \in S_{\{p,q\}}$
  $R'(s) \subseteq S_{\{p,q\}}$

pq	$p\neg q$
$\neg pq$	$\neg p\neg q$

  (ii) then, the $\neg q$-states are eliminated

pq	$p\neg q$
$\neg pq$	$\neg p\neg q$

becoming aware, cont.

- but *even for scenario of type (1)*, the probabilistic case is much more tricky
- one could reason like this:
- initially, Pierre is only aware of p, neither q nor r: $\rho(s) \in S_{\{p\}}$ and $Supp(P(s)) \subseteq S_{\{p\}}$

p	(α)
$\neg p$	$(1-\alpha)$

- Pierre is informed that q
  (i) first, the structure is modified: $Supp(P'(s)) \subseteq S_{\{p,q\}}$ and for each ϕ of $\mathcal{L}(\{p\})$, $P(s)([[\phi]] = P'(s)([[\phi]])$
  - state p has initially weight α
  - state p is splitted in pq and $p\neg q$, each with weight $\alpha/2$

pq	$(\alpha/2)$	$p\neg q$	$(\alpha/2)$
$\neg pq$	$((1-\alpha)/2)$	$\neg p\neg q$	$((1-\alpha)/2)$

  (ii) then, Pierre conditionalizes on the information that q

pq	(α)	$p\neg q$	
$\neg pq$	$(1-\alpha)$	$\neg p\neg q$	

- but the new probability of p could be affected by the fact that the agent learns that q (intuitively, if p and q are not independent - think about p = "the house is quiet" and q = "the house is no more than 1 km far from an airport"