Probabilistic Update Logic and Semantic Concept Learning

(based on joint work with Shalom Lappin)

Jan van Eijck

CWI Amsterdam

Kutaisi
September 27, 2011

Outline

Outline

Quick Intro to Dynamic Epistemic Logic (DEL)

Bayes' Law as a Learning Algorithm

Outline

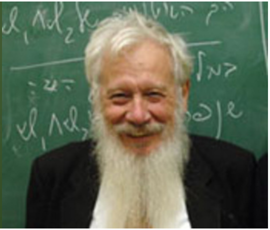
Outline

Outline

Very Brief History of (Dynamic) Epistemic Logic

Very Brief History of (Dynamic) Epistemic Logic

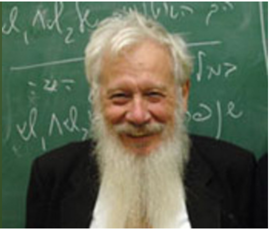




Very Brief History of (Dynamic) Epistemic Logic

David Lewis

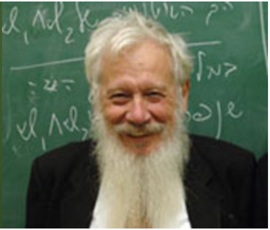
Jaakko Hintikka



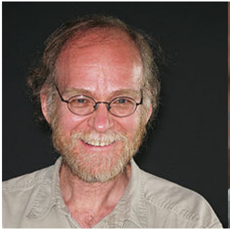
Robert Aumann

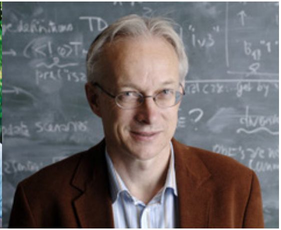
Very Brief History of (Dynamic) Epistemic Logic





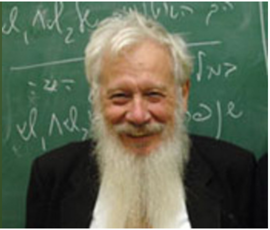
David Lewis Jaakko Hintikka Robert Aumann



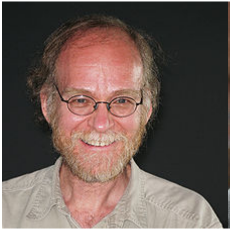



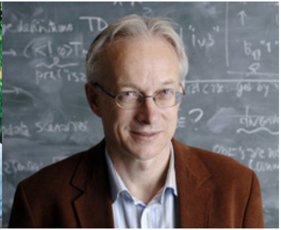
Very Brief History of (Dynamic) Epistemic Logic



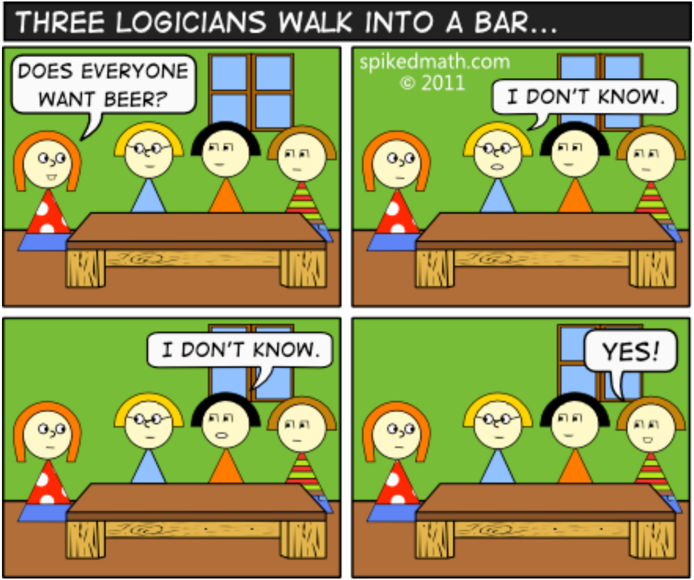


David Lewis Jaakko Hintikka Robert Aumann



Joe Halpern Jan Plaza A. Baltag Johan van Benthem


THREE LOGICIANS WALK INTO A BAR...
DOES EVERYONE WANT BEER?
spikedmath.com © 2011
I DON'T KNOW.
I DON'T KNOW.
YES!

picture by
Marco Swaen

The Muddy Children Puzzle

a clean, *b*, *c* and *d* muddy.

	a	b	c	d
at least one of you is muddy	○	●	●	●

The Muddy Children Puzzle

a clean, *b*, *c* and *d* muddy.

	a	b	c	d
at least one of you is muddy	○	●	●	●
who knows his state?				

The Muddy Children Puzzle

a clean, *b*, *c* and *d* muddy.

	a	b	c	d
at least one of you is muddy	○	●	●	●
who knows his state?	N	N	N	N

The Muddy Children Puzzle

a clean, *b*, *c* and *d* muddy.

	a	b	c	d
at least one of you is muddy	○	●	●	●
who knows his state?	N	N	N	N
who knows his state now?				

The Muddy Children Puzzle

a clean, *b*, *c* and *d* muddy.

	a	b	c	d
at least one of you is muddy	○	●	●	●
who knows his state?	N	N	N	N
who knows his state now?	N	N	N	N

The Muddy Children Puzzle

a clean, *b*, *c* and *d* muddy.

	a	b	c	d
at least one of you is muddy	○	●	●	●
who knows his state?	N	N	N	N
who knows his state now?	N	N	N	N
who knows his state now?				

The Muddy Children Puzzle

a clean, *b*, *c* and *d* muddy.

	a	b	c	d
at least one of you is muddy	○	●	●	●
who knows his state?	N	N	N	N
who knows his state now?	N	N	N	N
who knows his state now?	N	Y	Y	Y

The Muddy Children Puzzle

a clean, *b*, *c* and *d* muddy.

	a	b	c	d
at least one of you is muddy	○	●	●	●
who knows his state?	N	N	N	N
who knows his state now?	N	N	N	N
who knows his state now?	N	Y	Y	Y
who knows his state now?				

The Muddy Children Puzzle

a clean, *b*, *c* and *d* muddy.

	a	b	c	d
at least one of you is muddy	○	●	●	●
who knows his state?	N	N	N	N
who knows his state now?	N	N	N	N
who knows his state now?	N	Y	Y	Y
who knows his state now?	Y			

The Muddy Children (2)

a, *b*, *c* clean, *d* muddy.

	a	b	c	d
at least one of you is muddy	○	○	○	●

The Muddy Children (2)

a, *b*, *c* clean, *d* muddy.

	a	b	c	d
at least one of you is muddy	○	○	○	●
who knows his state?				

The Muddy Children (2)

a, *b*, *c* clean, *d* muddy.

	a	b	c	d
at least one of you is muddy	○	○	○	●
who knows his state?	N	N	N	Y

The Muddy Children (2)

a, *b*, *c* clean, *d* muddy.

	a	b	c	d
at least one of you is muddy	○	○	○	●
who knows his state?	N	N	N	Y
who knows his state now?				

The Muddy Children (2)

a, *b*, *c* clean, *d* muddy.

	a	b	c	d
at least one of you is muddy	○	○	○	●
who knows his state?	N	N	N	Y
who knows his state now?	Y	Y	Y	

The Muddy Children (3)

a, *b* clean, *c*, *d* muddy.

	a	b	c	d
at least one of you is muddy	○	○	●	●

The Muddy Children (3)

a, *b* clean, *c*, *d* muddy.

	a	b	c	d
at least one of you is muddy	○	○	●	●
who knows his state?				

The Muddy Children (3)

a, *b* clean, *c*, *d* muddy.

	a	b	c	d
at least one of you is muddy	○	○	●	●
who knows his state?	N	N	N	N

The Muddy Children (3)

a, *b* clean, *c*, *d* muddy.

	a	b	c	d
at least one of you is muddy	○	○	●	●
who knows his state?	N	N	N	N
who knows his state now?				

The Muddy Children (3)

a, *b* clean, *c*, *d* muddy.

	a	b	c	d
at least one of you is muddy	○	○	●	●
who knows his state?	N	N	N	N
who knows his state now?	N	N	Y	Y

The Muddy Children (3)

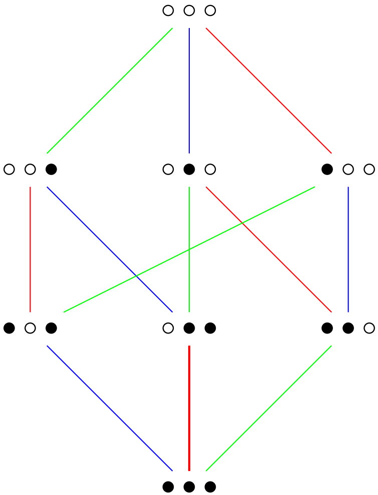
a, *b* clean, *c*, *d* muddy.

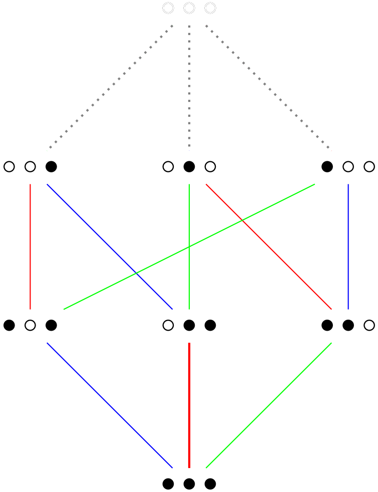
	a	b	c	d
at least one of you is muddy	○	○	●	●
who knows his state?	N	N	N	N
who knows his state now?	N	N	Y	Y
who knows his state now?				

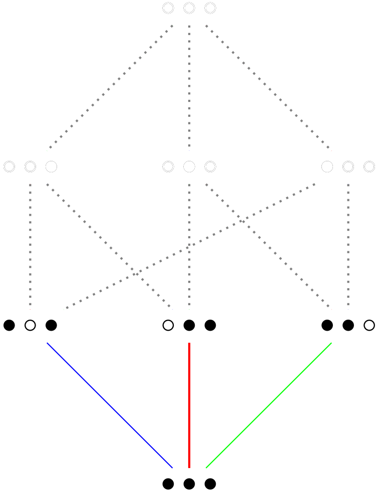
The Muddy Children (3)

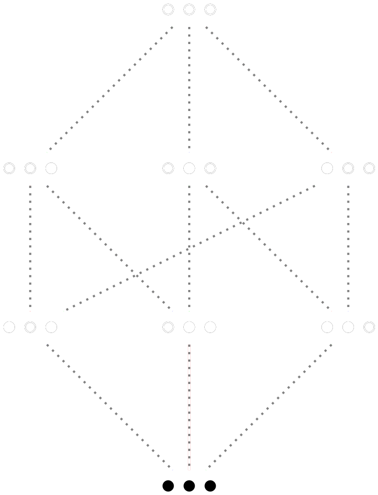
a, *b* clean, *c*, *d* muddy.

	a	b	c	d
at least one of you is muddy	○	○	●	●
who knows his state?	N	N	N	N
who knows his state now?	N	N	Y	Y
who knows his state now?	Y	Y		



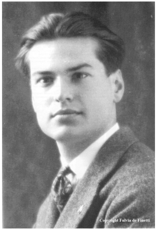






Some New Heroes

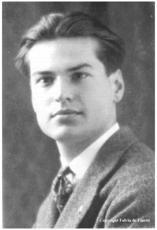
Some New Heroes



Some New Heroes

Rev Thomas Bayes

Rudolph Carnap



Bruno de Finetti

Laws of Conditional Probability

Let H be an event with positive probability. Let A be any event. Then we define:

$$P(A|H) = \frac{P(AH)}{P(A)}.$$

Laws of Conditional Probability

Let H be an event with positive probability. Let A be any event. Then we define:

$$P(A|H) = \frac{P(AH)}{P(A)}.$$

From this:

$$P(AH) = P(A|H) \cdot P(A).$$

Laws of Conditional Probability

Let H be an event with positive probability. Let A be any event. Then we define:

$$P(A|H) = \frac{P(AH)}{P(A)}.$$

From this:

$$P(AH) = P(A|H) \cdot P(A).$$

Suppose $H_1, \ldots, H_n$ are mutually exclusive events, and their union is the whole sample space Ω. That is, one of the H_i necessary occurs.

Laws of Conditional Probability

Let H be an event with positive probability. Let A be any event. Then we define:

$$P(A|H) = \frac{P(AH)}{P(A)}.$$

From this:

$$P(AH) = P(A|H) \cdot P(A).$$

Suppose $H_1, \ldots, H_n$ are mutually exclusive events, and their union is the whole sample space Ω. That is, one of the H_i necessary occurs. Then we have for any event A:

$$A = AH_1 \cup AH_2 \cup \cdots \cup AH_n.$$

Laws of Conditional Probability

Let H be an event with positive probability. Let A be any event. Then we define:

$$P(A|H) = \frac{P(AH)}{P(A)}.$$

From this:

$$P(AH) = P(A|H) \cdot P(A).$$

Suppose $H_1, \ldots, H_n$ are mutually exclusive events, and their union is the whole sample space Ω. That is, one of the H_i necessary occurs. Then we have for any event A:

$$A = AH_1 \cup AH_2 \cup \cdots \cup AH_n.$$

Since the AH_i are mutually exclusive, their probabilities add:

$$P(A) = \sum_{j=1}^{n} P(A|H_j) \cdot P(H_j).$$

Bayes' Law

For the special case of H_j we have:

$$P(H_j|A) = \frac{P(AH_j)}{P(A)}.$$

Bayes' Law

For the special case of H_j we have:

$$P(H_j|A) = \frac{P(AH_j)}{P(A)}.$$

Expanding $P(AH_j)$ and $P(A)$, we get:

$$P(H_j|A) = \frac{P(A|H_j) \cdot P(H_j)}{\sum_{i=1}^{n} P(A|H_i) \cdot P(H_i)}.$$

Bayes' Law

For the special case of H_j we have:

$$P(H_j|A) = \frac{P(AH_j)}{P(A)}.$$

Expanding $P(AH_j)$ and $P(A)$, we get:

$$P(H_j|A) = \frac{P(A|H_j) \cdot P(H_j)}{\sum_{i=1}^{n} P(A|H_i) \cdot P(H_i)}.$$

This is called *Bayes' Law*.

Bayes' Law

For the special case of H_j we have:

$$P(H_j|A) = \frac{P(AH_j)}{P(A)}.$$

Expanding $P(AH_j)$ and $P(A)$, we get:

$$P(H_j|A) = \frac{P(A|H_j) \cdot P(H_j)}{\sum_{i=1}^{n} P(A|H_i) \cdot P(H_i)}.$$

This is called *Bayes' Law*.
Bayes' law can be viewed as a learning algorithm: how probable is hypothesis H_j, given the data A that were observed using H_j?

Connection with Update Logic

Connection with Update Logic

- $P(H_j)$ is what we know of H_j initially,

Connection with Update Logic

- $P(H_j)$ is what we know of H_j initially,
- $P(A|H_j)$ is what we know of the update effect of A, given our initial knowledge,

Connection with Update Logic

- $P(H_j)$ is what we know of H_j initially,
- $P(A|H_j)$ is what we know of the update effect of A, given our initial knowledge,
- $P(H_j|A)$ is our new state of knowledge.

Connection with Update Logic

- $P(H_j)$ is what we know of H_j initially,
- $P(A|H_j)$ is what we know of the update effect of A, given our initial knowledge,
- $P(H_j|A)$ is our new state of knowledge.
- Persistent question in Bayesian analysis: how to get a reasonable prior $P(H_j)$?

Connection with Update Logic

- $P(H_j)$ is what we know of H_j initially,
- $P(A|H_j)$ is what we know of the update effect of A, given our initial knowledge,
- $P(H_j|A)$ is our new state of knowledge.
- Persistent question in Bayesian analysis: how to get a reasonable prior $P(H_j)$?
- Persistent question in epistemic model checking: how to get a reasonable initial epistemic model?

Probabilistic DEL

- Based on Van Benthem, Gerbrandy, Kooi [2].
- Compare also: Kooi's PhD Thesis [7], Baltag and Smets [1], Gierasimszuk [4], Halpern [5], . . .
- For simplicity, take the single agent case.

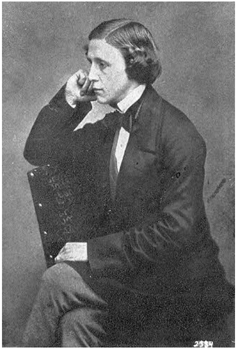
Probabilistic DEL

- Based on Van Benthem, Gerbrandy, Kooi [2].
- Compare also: Kooi's PhD Thesis [7], Baltag and Smets [1], Gierasimszuk [4], Halpern [5], ...
- For simplicity, take the single agent case.
- Probabilistic Epistemic Model M is pair (W, P), where
  - W is a (finite) set of worlds,
  - $P : W \to [0, 1]$ is a probability distribution.

Probabilistic DEL

- Based on Van Benthem, Gerbrandy, Kooi [2].
- Compare also: Kooi's PhD Thesis [7], Baltag and Smets [1], Gierasimszuk [4], Halpern [5], . . .
- For simplicity, take the single agent case.
- Probabilistic Epistemic Model M is pair (W, P), where
  - W is a (finite) set of worlds,
  - $P : W \to [0, 1]$ is a probability distribution.
  - This means: $P(w) \in \{p \mid 0 \leq p \leq 1\}$, $\sum_{w \in W} P(w) = 1$.

A Puzzle of Lewis Carroll



An urn contains a single marble, either white or black. Mr A puts another marble in the urn, a white one. The urn now contains two marbles. Next, Mrs B draws one of the two marbles from the urn. It turns out to be white. What is the probability that the other marble is also white? (Gardner [3])

Discussion: Representation of the Mrs B Update

Discussion: Representation of the Mrs B Update

- Call the first white marble w and the second one w'. Mrs B does not know whether she is drawing from $b + w'$ or from $w + w'$.

Discussion: Representation of the Mrs B Update

- Call the first white marble w and the second one w'. Mrs B does not know whether she is drawing from $b + w'$ or from $w + w'$.
- For the drawing event there are four cases that are all equally likely:
  1. $b \wedge w'$: b is taken out,
  2. $b \wedge w'$: w' is taken out,
  3. $w \wedge w'$: w is taken out,
  4. $w \wedge w'$: w' is taken out.

Discussion: Representation of the Mrs B Update

- Call the first white marble w and the second one w'. Mrs B does not know whether she is drawing from $b + w'$ or from $w + w'$.
- For the drawing event there are four cases that are all equally likely:
  1. $b \wedge w'$: b is taken out,
  2. $b \wedge w'$: w' is taken out,
  3. $w \wedge w'$: w is taken out,
  4. $w \wedge w'$: w' is taken out.
- Revealing the colour boils down to the public observation of $\neg(b \text{ out})$.

Discussion: Representation of the Mrs B Update

- Call the first white marble w and the second one w'. Mrs B does not know whether she is drawing from $b + w'$ or from $w + w'$.
- For the drawing event there are four cases that are all equally likely:
  1. $b \wedge w'$: b is taken out,
  2. $b \wedge w'$: w' is taken out,
  3. $w \wedge w'$: w is taken out,
  4. $w \wedge w'$: w' is taken out.
- Revealing the colour boils down to the public observation of $\neg(b \text{ out})$.
- Lumping the two parts of the action together, and normalizing gives:
  - $b \wedge w'$: w' is taken out, $\frac{1}{3}$,
  - $w \wedge w'$: w is taken out, $\frac{1}{3}$,
  - $w \wedge w'$: w' is taken out, $\frac{1}{3}$.

Initial Situation $b : \frac{1}{2}, w : \frac{1}{2}.$

Initial Situation $b : \frac{1}{2}, w : \frac{1}{2}$.

After Mr A Update $b + w' : \frac{1}{2}, w + w' : \frac{1}{2}$.

Initial Situation $b : \frac{1}{2}, w : \frac{1}{2}$.

After Mr A Update $b + w' : \frac{1}{2}, w + w' : \frac{1}{2}$.

Mrs B Update “takes a white marble out”

- $b \wedge w'$: take w' out : $\frac{1}{3}$,
- $w \wedge w'$: take w out : $\frac{1}{3}$,
- $w \wedge w'$: take w' out : $\frac{1}{3}$.

Initial Situation $b : \frac{1}{2}, w : \frac{1}{2}$.

After Mr A Update $b + w' : \frac{1}{2}, w + w' : \frac{1}{2}$.

Mrs B Update "takes a white marble out"

- $b \wedge w'$: take w' out : $\frac{1}{3}$,
- $w \wedge w'$: take w out : $\frac{1}{3}$,
- $w \wedge w'$: take w' out : $\frac{1}{3}$.

After Mrs B Update

$$w', w \text{ out} : \frac{1}{3} \times \frac{1}{2}$$

$$w, w' \text{ out} : \frac{1}{3} \times \frac{1}{2}$$

$$b, w' \text{ out} : \frac{1}{3} \times \frac{1}{2}$$

Initial Situation $b : \frac{1}{2}, w : \frac{1}{2}$.

After Mr A Update $b + w' : \frac{1}{2}, w + w' : \frac{1}{2}$.

Mrs B Update "takes a white marble out"

- $b \wedge w'$: take w' out : $\frac{1}{3}$,
- $w \wedge w'$: take w out : $\frac{1}{3}$,
- $w \wedge w'$: take w' out : $\frac{1}{3}$.

After Mrs B Update

$$w', w \text{ out} : \frac{1}{3} \times \frac{1}{2}$$

$$w, w' \text{ out} : \frac{1}{3} \times \frac{1}{2}$$

$$b, w' \text{ out} : \frac{1}{3} \times \frac{1}{2}$$

Normalized

$$w', w \text{ out} : \frac{1}{3}$$

$$w, w' \text{ out} : \frac{1}{3}$$

$$b, w' \text{ out} : \frac{1}{3}$$

Initial Situation $b : \frac{1}{2}$, $w : \frac{1}{2}$.

After Mr A Update $b + w' : \frac{1}{2}$, $w + w' : \frac{1}{2}$.

Mrs B Update "takes a white marble out"

- $b \wedge w'$: take w' out : $\frac{1}{3}$,
- $w \wedge w'$: take w out : $\frac{1}{3}$,
- $w \wedge w'$: take w' out : $\frac{1}{3}$.

After Mrs B Update

$$w', w \text{ out} : \frac{1}{3} \times \frac{1}{2}$$

$$w, w' \text{ out} : \frac{1}{3} \times \frac{1}{2}$$

$$b, w' \text{ out} : \frac{1}{3} \times \frac{1}{2}$$

Normalized

$$w', w \text{ out} : \frac{1}{3}$$

$$w, w' \text{ out} : \frac{1}{3}$$

$$b, w' \text{ out} : \frac{1}{3}$$

Probability of $w \vee w'$: $\frac{2}{3}$.

The Puzzle of Monty Hall

Form of the Monty Hall Update

precondition	action	probability
$A \wedge \text{choice} = A$	$!\neg B$	$\frac{1}{2}$
$A \wedge \text{choice} = A$	$!\neg C$	$\frac{1}{2}$
$A \wedge \text{choice} = B$	$!\neg C$	1
$A \wedge \text{choice} = C$	$!\neg B$	1

Form of the Monty Hall Update

precondition	action	probability
$A \wedge \text{choice} = A$	$!\neg B$	$\frac{1}{2}$
$A \wedge \text{choice} = A$	$!\neg C$	$\frac{1}{2}$
$A \wedge \text{choice} = B$	$!\neg C$	1
$A \wedge \text{choice} = C$	$!\neg B$	1
$B \wedge \text{choice} = A$	$!\neg C$	1
$B \wedge \text{choice} = B$	$!\neg A$	$\frac{1}{2}$
$B \wedge \text{choice} = B$	$!\neg C$	$\frac{1}{2}$
$B \wedge \text{choice} = C$	$!\neg A$	1

Form of the Monty Hall Update

precondition	action	probability
$A \wedge \text{choice} = A$	$!\neg B$	$\frac{1}{2}$
$A \wedge \text{choice} = A$	$!\neg C$	$\frac{1}{2}$
$A \wedge \text{choice} = B$	$!\neg C$	1
$A \wedge \text{choice} = C$	$!\neg B$	1
$B \wedge \text{choice} = A$	$!\neg C$	1
$B \wedge \text{choice} = B$	$!\neg A$	$\frac{1}{2}$
$B \wedge \text{choice} = B$	$!\neg C$	$\frac{1}{2}$
$B \wedge \text{choice} = C$	$!\neg A$	1
$C \wedge \text{choice} = A$	$!\neg B$	1
$C \wedge \text{choice} = B$	$!\neg A$	1
$C \wedge \text{choice} = C$	$!\neg A$	$\frac{1}{2}$
$C \wedge \text{choice} = C$	$!\neg B$	$\frac{1}{2}$

Run of the Monty Hall Update Scenario

Run of the Monty Hall Update Scenario

- Initial situation: suppose my choice $= A$.

Run of the Monty Hall Update Scenario

- Initial situation: suppose my choice $= A$.
- Suppose Monty Hall announces $!\neg B$.

Run of the Monty Hall Update Scenario

- Initial situation: suppose my choice $= A$.
- Suppose Monty Hall announces $!\neg B$.
- Then the Monty Hall update boils down to:
  - $A \wedge \text{choice} = A$: $!\neg B$: $\frac{1}{3}$
  - $B \wedge \text{choice} = A$: $!\neg B$: 0
  - $C \wedge \text{choice} = A$: $!\neg B$: $\frac{2}{3}$

Run of the Monty Hall Update Scenario

- Initial situation: suppose my choice $= A$.
- Suppose Monty Hall announces $!\neg B$.
- Then the Monty Hall update boils down to:
  - $A \wedge \text{choice} = A$: $!\neg B$: $\frac{1}{3}$
  - $B \wedge \text{choice} = A$: $!\neg B$: 0
  - $C \wedge \text{choice} = A$: $!\neg B$: $\frac{2}{3}$
- Result of update with this:
  $A, \text{choice} = A, \frac{1}{3}$, $B, \text{choice} = A, 0$, $C, \text{choice} = A, \frac{2}{3}$.

Run of the Monty Hall Update Scenario

- Initial situation: suppose my choice $= A$.
- Suppose Monty Hall announces $!\neg B$.
- Then the Monty Hall update boils down to:
  - $A \wedge \text{choice} = A$: $!\neg B$: $\frac{1}{3}$
  - $B \wedge \text{choice} = A$: $!\neg B$: 0
  - $C \wedge \text{choice} = A$: $!\neg B$: $\frac{2}{3}$
- Result of update with this: $A, \text{choice} = A, \frac{1}{3}$, $B, \text{choice} = A, 0$, $C, \text{choice} = A, \frac{2}{3}$.
- It is clear I should reconsider my choice.

Semantic Concept Learning

Semantic Concept Learning

- Learning events: presentation of a new object, use of a new word.

Semantic Concept Learning

- Learning events: presentation of a new object, use of a new word.

Semantic Concept Learning

- Learning events: presentation of a new object, use of a new word.

- "This is a rose"

Semantic Concept Learning

- Learning events: presentation of a new object, use of a new word.

- “This is a rose”
- “Dit is een roos”

Semantic Concept Learning

- Learning events: presentation of a new object, use of a new word.

- “This is a rose”
- “Dit is een roos”
- Do we learn the use of a word? Do we update our knowledge about roses? Or both?

Modelling Uncertainty About Basic Predications

Modelling Uncertainty About Basic Predications

- Let a propositional language over a set of basic predications be given, as follows.

$$
\begin{array}{rcl}
t & ::= & x \mid a_1 \mid a_2 \mid \cdots \mid a_m \\
Q & ::= & Q_1 \mid Q_2 \mid \cdots \mid Q_n \\
\phi & ::= & Qt \mid \neg\phi \mid \phi \wedge \phi \mid \phi \vee \phi.
\end{array}
$$

Modelling Uncertainty About Basic Predications

- Let a propositional language over a set of basic predications be given, as follows.

$$\begin{aligned} t &::= x \mid a_1 \mid a_2 \mid \cdots \mid a_m \\ Q &::= Q_1 \mid Q_2 \mid \cdots \mid Q_n \\ \phi &::= Qt \mid \neg\phi \mid \phi \wedge \phi \mid \phi \vee \phi. \end{aligned}$$

- Here we assume a single variable x, a finite number of proper names $a_1, a_2, \ldots, a_m$ and a finite number of basic unary predicates $Q_1, Q_2, \ldots, Q_n$.

Modelling Uncertainty About Basic Predications

- Let a propositional language over a set of basic predications be given, as follows.

$$
\begin{array}{rcl}
t & ::= & x \mid a_1 \mid a_2 \mid \cdots \mid a_m \\
Q & ::= & Q_1 \mid Q_2 \mid \cdots \mid Q_n \\
\phi & ::= & Qt \mid \neg\phi \mid \phi \land \phi \mid \phi \lor \phi.
\end{array}
$$

- Here we assume a single variable x, a finite number of proper names $a_1, a_2, \ldots, a_m$ and a finite number of basic unary predicates $Q_1, Q_2, \ldots, Q_n$.
- Any ϕ that contains occurrences of x is called a *predication*. Use $\phi(x)$ for predications, and $\phi(a/x)$ for the result of replacing x by a everywhere in a predication.

New Names, New Predicates

New Names, New Predicates

- Call this language L_n^m.

New Names, New Predicates

- Call this language L_n^m.
- If we extend L_n^m with one name a_{m+1}, the new language is called L_n^{m+1}.

New Names, New Predicates

- Call this language L_n^m.
- If we extend L_n^m with one name a_{m+1}, the new language is called L_n^{m+1}.
- If we extend L_n^m with one new predicate Q_{n+1}, the new language is called L_{n+1}^m

Representing Partial Knowledge

Representing Partial Knowledge

- For convenience, we identify names and objects, so we assume a domain $D_m = \{a_1, a_2, \ldots, a_m\}$.

Representing Partial Knowledge

- For convenience, we identify names and objects, so we assume a domain $D_m = \{a_1, a_2, \ldots, a_m\}$.
- The type of a world w with respect to language L_n^m is given by $w : \{Q_1, \ldots, Q_n\} \rightarrow \mathcal{P}(D_m)$.

Representing Partial Knowledge

- For convenience, we identify names and objects, so we assume a domain $D_m = \{a_1, a_2, \ldots, a_m\}$.
- The type of a world w with respect to language L_n^m is given by $w : \{Q_1, \ldots, Q_n\} \to \mathcal{P}(D_m)$.
- $w(Q_i)$ is the interpretation of Q_i in w, for L_n^m.

Representing Partial Knowledge

- For convenience, we identify names and objects, so we assume a domain $D_m = \{a_1, a_2, \dots, a_m\}$.
- The type of a world w with respect to language L_n^m is given by $w : \{Q_1, \dots, Q_n\} \to \mathcal{P}(D_m)$.
- $w(Q_i)$ is the interpretation of Q_i in w, for L_n^m.
- A probabilistic model M is a tuple (D, W, P) with D a domain, W a set of worlds for that domain (predicate interpretations in that domain), and P a probability function over W, i.e., for all $w \in W$, $P(w) \in [0, 1]$, and $\sum_{w \in W} P(w) = 1$.

Representing Partial Knowledge

- For convenience, we identify names and objects, so we assume a domain $D_m = \{a_1, a_2, \ldots, a_m\}$.
- The type of a world w with respect to language L_n^m is given by $w : \{Q_1, \ldots, Q_n\} \to \mathcal{P}(D_m)$.
- $w(Q_i)$ is the interpretation of Q_i in w, for L_n^m.
- A probabilistic model M is a tuple (D, W, P) with D a domain, W a set of worlds for that domain (predicate interpretations in that domain), and P a probability function over W, i.e., for all $w \in W$, $P(w) \in [0, 1]$, and $\sum_{w \in W} P(w) = 1$.
- The probabilities in a model M represent the priors of an idealized semantic learner.

Interpretation

Interpretation

- An interpretation of L_n^m in an L_n^m-model $M = (D, W, P)$ is given in terms of the standard notion $w \models \phi$, as follows:

$$[\![\phi]\!]^M \quad := \quad \sum \{P(w) \mid w \in W, w \models \phi\}$$

Interpretation

- An interpretation of L_n^m in an L_n^m-model $M = (D, W, P)$ is given in terms of the standard notion $w \models \phi$, as follows:

$$[\![\phi]\!]^M \quad := \quad \sum\{P(w) \mid w \in W, w \models \phi\}$$

- It is straightforward to verify that this yields $[\![\neg\phi]\!]^M = 1 - [\![\phi]\!]^M$.

Interpretation

- An interpretation of L_n^m in an L_n^m-model $M = (D, W, P)$ is given in terms of the standard notion $w \models \phi$, as follows:

$$[\![\phi]\!]^M \quad := \quad \sum \{P(w) \mid w \in W, w \models \phi\}$$

- It is straightforward to verify that this yields $[\![\neg\phi]\!]^M = 1 - [\![\phi]\!]^M$.
- Also, if $\phi \models \neg\psi$, i.e., if $W_\phi \cap W_\psi = \emptyset$, then $[\![\phi \vee \psi]\!]^M = \sum_{w \in W_{\phi\vee\psi}} P(w) = \sum_{w \in W_\phi} P(w) + \sum_{w \in W_\psi} P(w) = [\![\phi]\!]^M + [\![\psi]\!]^M$.

Interpretation

- An interpretation of L_n^m in an L_n^m-model $M = (D, W, P)$ is given in terms of the standard notion $w \models \phi$, as follows:

$$[\![\phi]\!]^M \quad := \quad \sum\{P(w) \mid w \in W, w \models \phi\}$$

- It is straightforward to verify that this yields $[\![\neg\phi]\!]^M = 1 - [\![\phi]\!]^M$.
- Also, if $\phi \models \neg\psi$, i.e., if $W_\phi \cap W_\psi = \emptyset$, then $[\![\phi \vee \psi]\!]^M = \sum_{w \in W_{\phi\vee\psi}} P(w) = \sum_{w \in W_\phi} P(w) + \sum_{w \in W_\psi} P(w) = [\![\phi]\!]^M + [\![\psi]\!]^M$.
- Tautologies have probability 1, contradictions probability 0.

Example 1

Example 1

- Assume there are just two predicates Q_1 and Q_2, and two objects a, b.

Example 1

- Assume there are just two predicates Q_1 and Q_2, and two objects a, b.
- Complete ignorance about how the predicates are applied is represented by a model with 16 worlds, because for each object x and each predicate Q there are two cases: Q applies to x or not.

Example 1

- Assume there are just two predicates Q_1 and Q_2, and two objects a, b.
- Complete ignorance about how the predicates are applied is represented by a model with 16 worlds, because for each object x and each predicate Q there are two cases: Q applies to x or not.
- If the probability of each of the cases is completely unknown, each of these worlds has probability $\frac{1}{16}$.

Example 2

Example 2

- Suppose again there are two objects a, b and two predicates Q_1, Q_2.

Example 2

- Suppose again there are two objects a, b and two predicates Q_1, Q_2.
- Suppose it is known that a has Q_1, and the probability that b has Q_1 is taken to be $\frac{2}{3}$.

Example 2

- Suppose again there are two objects a, b and two predicates Q_1, Q_2.
- Suppose it is known that a has Q_1, and the probability that b has Q_1 is taken to be $\frac{2}{3}$.
- Suppose it is known that no object has Q_2.

Example 2

- Suppose again there are two objects a, b and two predicates Q_1, Q_2.
- Suppose it is known that a has Q_1, and the probability that b has Q_1 is taken to be $\frac{2}{3}$.
- Suppose it is known that no object has Q_2.
- Then $W = \{w_1, w_2\}$ with $w_1(Q_1) = \{a, b\}$, $w_2(Q_1) = \{a\}$, $w_1(Q_2) = \emptyset$, $w_2(Q_2) = \emptyset$,

Example 2

- Suppose again there are two objects a, b and two predicates Q_1, Q_2.
- Suppose it is known that a has Q_1, and the probability that b has Q_1 is taken to be $\frac{2}{3}$.
- Suppose it is known that no object has Q_2.
- Then $W = \{w_1, w_2\}$ with $w_1(Q_1) = \{a, b\}$, $w_2(Q_1) = \{a\}$, $w_1(Q_2) = \emptyset$, $w_2(Q_2) = \emptyset$,
- P is given by $P(w_1) = \frac{2}{3}$, $P(w_2) = \frac{1}{3}$.

Example 2

- Suppose again there are two objects a, b and two predicates Q_1, Q_2.
- Suppose it is known that a has Q_1, and the probability that b has Q_1 is taken to be $\frac{2}{3}$.
- Suppose it is known that no object has Q_2.
- Then $W = \{w_1, w_2\}$ with $w_1(Q_1) = \{a, b\}$, $w_2(Q_1) = \{a\}$, $w_1(Q_2) = \emptyset$, $w_2(Q_2) = \emptyset$,
- P is given by $P(w_1) = \frac{2}{3}$, $P(w_2) = \frac{1}{3}$.
- In this example $\neg Q_1(b)$ is true in w_2 and not in w_1.

Example 2

- Suppose again there are two objects a, b and two predicates Q_1, Q_2.
- Suppose it is known that a has Q_1, and the probability that b has Q_1 is taken to be $\frac{2}{3}$.
- Suppose it is known that no object has Q_2.
- Then $W = \{w_1, w_2\}$ with $w_1(Q_1) = \{a, b\}$, $w_2(Q_1) = \{a\}$, $w_1(Q_2) = \emptyset$, $w_2(Q_2) = \emptyset$,
- P is given by $P(w_1) = \frac{2}{3}$, $P(w_2) = \frac{1}{3}$.
- In this example $\neg Q_1(b)$ is true in w_2 and not in w_1.
- Therefore $[\![\neg Q_1(b)]\!] = \frac{1}{3}$.

Learning Semantic Concepts: Dictionary Learning

Learning Semantic Concepts: Dictionary Learning

- Learning a new semantic concept Q_{n+1} is learning how (or: to what extent) predicate Q_{n+1} applies to the objects one knows about.

Learning Semantic Concepts: Dictionary Learning

- Learning a new semantic concept Q_{n+1} is learning how (or: to what extent) predicate Q_{n+1} applies to the objects one knows about.
- The simplest way to model such a learning event is as a pair $(Q_{n+1}, \phi(x))$ where $\phi(x)$ is an L_n^m predication.

Learning Semantic Concepts: Dictionary Learning

- Learning a new semantic concept Q_{n+1} is learning how (or: to what extent) predicate Q_{n+1} applies to the objects one knows about.
- The simplest way to model such a learning event is as a pair $(Q_{n+1}, \phi(x))$ where $\phi(x)$ is an L_n^m predication.
- The effect of the learning event could then be modelled in a way that is very similar to the manner in which factual change is modelled in epistemic update logic.

Learning Semantic Concepts: Dictionary Learning

- Learning a new semantic concept Q_{n+1} is learning how (or: to what extent) predicate Q_{n+1} applies to the objects one knows about.
- The simplest way to model such a learning event is as a pair $(Q_{n+1}, \phi(x))$ where $\phi(x)$ is an L_n^m predication.
- The effect of the learning event could then be modelled in a way that is very similar to the manner in which factual change is modelled in epistemic update logic.
- The result of updating a model $M = (D, W, P)$ with concept learning event $(Q_{n+1}, \phi(x))$ is the model that is like M except for the fact that the interpretation in each world of Q_{n+1} is given by

$$w(Q_{n+1}) := \{a \mid a \in D_m, w \models \phi(a/x)\}.$$

Learning Semantic Concepts: Dictionary Learning

- Learning a new semantic concept Q_{n+1} is learning how (or: to what extent) predicate Q_{n+1} applies to the objects one knows about.
- The simplest way to model such a learning event is as a pair $(Q_{n+1}, \phi(x))$ where $\phi(x)$ is an L_n^m predication.
- The effect of the learning event could then be modelled in a way that is very similar to the manner in which factual change is modelled in epistemic update logic.
- The result of updating a model $M = (D, W, P)$ with concept learning event $(Q_{n+1}, \phi(x))$ is the model that is like M except for the fact that the interpretation in each world of Q_{n+1} is given by

$$w(Q_{n+1}) := \{a \mid a \in D_m, w \models \phi(a/x)\}.$$

- Note that the probability function P of the model does not change.

Learning Example

Learning Example

- Let's return to example 1.

Learning Example

- Let's return to example 1.
- This is the model where there are two objects and two predicates, and nothing is known about the properties of the objects.

Learning Example

- Let's return to example 1.
- This is the model where there are two objects and two predicates, and nothing is known about the properties of the objects.
- Take the learning event $(Q_3, Q_1 x \wedge \neg Q_2 x)$.

Learning Example

- Let's return to example 1.
- This is the model where there are two objects and two predicates, and nothing is known about the properties of the objects.
- Take the learning event $(Q_3, Q_1 x \wedge \neg Q_2 x)$.
- This defines Q_3 as the difference of Q_1 and Q_2.

Learning Example

- Let's return to example 1.
- This is the model where there are two objects and two predicates, and nothing is known about the properties of the objects.
- Take the learning event $(Q_3, Q_1 x \wedge \neg Q_2 x)$.
- This defines Q_3 as the difference of Q_1 and Q_2.
- The resulting model will again have 16 worlds, and in each world w_i, $w_i(Q_3)$ is given by $w_i(Q_1) \cap (D - w_i(Q_2))$.

Learning Example

- Let's return to example 1.
- This is the model where there are two objects and two predicates, and nothing is known about the properties of the objects.
- Take the learning event $(Q_3, Q_1 x \wedge \neg Q_2 x)$.
- This defines Q_3 as the difference of Q_1 and Q_2.
- The resulting model will again have 16 worlds, and in each world w_i, $w_i(Q_3)$ is given by $w_i(Q_1) \cap (D - w_i(Q_2))$.
- The probabilities of the worlds remain unchanged.

Adjustment and Approximation (1)

Adjustment and Approximation (1)

- To allow adjustment of the meaning of a concept by means of a learning event, we can use probabilistic updating.

Adjustment and Approximation (1)

- To allow adjustment of the meaning of a concept by means of a learning event, we can use probabilistic updating.
- A concept learning event now is a tuple

$$(Q, \phi, \psi(x), q)$$

where ϕ is a sentence, $\psi(x)$ is a predication, and q is a probability.

Adjustment and Approximation (1)

- To allow adjustment of the meaning of a concept by means of a learning event, we can use probabilistic updating.
- A concept learning event now is a tuple

$$(Q, \phi, \psi(x), q)$$

  where ϕ is a sentence, $\psi(x)$ is a predication, and q is a probability.
- ϕ expresses the observational circumstances of the revision.

Adjustment and Approximation (1)

- To allow adjustment of the meaning of a concept by means of a learning event, we can use probabilistic updating.
- A concept learning event now is a tuple

$$(Q, \phi, \psi(x), q)$$

  where ϕ is a sentence, $\psi(x)$ is a predication, and q is a probability.
- ϕ expresses the observational circumstances of the revision.
- q expresses the observational certainty of the new information.

Adjustment and Approximation (2)

Adjustment and Approximation (2)

- The result of updating $M = (D, W, P)$ with $(Q, \phi, \psi(x), q)$ is a new model $M = (D, W', P')$.

Adjustment and Approximation (2)

- The result of updating $M = (D, W, P)$ with $(Q, \phi, \psi(x), q)$ is a new model $M = (D, W', P')$.
- W' is given by changing the interpretation of Q in members w of W_ϕ to $\{a \mid w \models \psi(a/x)\}$ while leaving the interpretation of Q in members of $W_{\neg\phi}$ unchanged.

Adjustment and Approximation (2)

- The result of updating $M = (D, W, P)$ with $(Q, \phi, \psi(x), q)$ is a new model $M = (D, W', P')$.
- W' is given by changing the interpretation of Q in members w of W_ϕ to $\{a \mid w \models \psi(a/x)\}$ while leaving the interpretation of Q in members of $W_{\neg\phi}$ unchanged.
- P' is given by $P'(w) = \frac{P(w) \times q}{X}$ for members of W_ϕ,

Adjustment and Approximation (2)

- The result of updating $M = (D, W, P)$ with $(Q, \phi, \psi(x), q)$ is a new model $M = (D, W', P')$.
- W' is given by changing the interpretation of Q in members w of W_ϕ to $\{a \mid w \models \psi(a/x)\}$ while leaving the interpretation of Q in members of $W_{\neg\phi}$ unchanged.
- P' is given by $P'(w) = \frac{P(w) \times q}{X}$ for members of W_ϕ,
- and by $P'(w) = \frac{P(w) \times (1-q)}{X}$ for members of $W_{\neg\phi}$.

Adjustment and Approximation (2)

- The result of updating $M = (D, W, P)$ with $(Q, \phi, \psi(x), q)$ is a new model $M = (D, W', P')$.
- W' is given by changing the interpretation of Q in members w of W_ϕ to $\{a \mid w \models \psi(a/x)\}$ while leaving the interpretation of Q in members of $W_{\neg\phi}$ unchanged.
- P' is given by $P'(w) = \frac{P(w) \times q}{X}$ for members of W_ϕ,
- and by $P'(w) = \frac{P(w) \times (1-q)}{X}$ for members of $W_{\neg\phi}$.
- $\frac{1}{X}$ (the normalization factor) is given by

$$X = \sum_{w \in W_\phi} P(w) \times q + \sum_{w \in W_{\neg\phi}} P(w) \times (1 - q).$$

Back to an Example

Back to an Example

- Consider again the example with the two objects and the two properties, where nothing is known. A learning event for this could be:

$$(Q_2, \neg Q_1 b, Q_1 x \vee Q_2 x, \frac{2}{3}).$$

Back to an Example

- Consider again the example with the two objects and the two properties, where nothing is known. A learning event for this could be:

$$(Q_2, \neg Q_1 b, Q_1 x \vee Q_2 x, \frac{2}{3}).$$

- Then the resulting model has again 2 worlds, but now the probability of w_2 has gone up from $\frac{1}{3}$ to

$$\frac{\frac{2}{3} \times \frac{1}{3}}{\frac{1}{3} \times \frac{2}{3} + \frac{2}{3} \times \frac{1}{3}} = \frac{1}{2}$$

Back to an Example

- Consider again the example with the two objects and the two properties, where nothing is known. A learning event for this could be:

$$(Q_2, \neg Q_1 b, Q_1 x \vee Q_2 x, \frac{2}{3}).$$

- Then the resulting model has again 2 worlds, but now the probability of w_2 has gone up from $\frac{1}{3}$ to

$$\frac{\frac{2}{3} \times \frac{1}{3}}{\frac{1}{3} \times \frac{2}{3} + \frac{2}{3} \times \frac{1}{3}} = \frac{1}{2}$$

- The probability of w_1 has gone down from $\frac{2}{3}$ to

$$\frac{\frac{1}{3} \times \frac{2}{3}}{\frac{4}{9}} = \frac{1}{2}.$$

Learning Concepts by Example (1)

Learning Concepts by Example (1)

- You are given something of which you are told that it is called a “rose”, and you observe that it is thorny, red and a flower.

Learning Concepts by Example (1)

- You are given something of which you are told that it is called a "rose", and you observe that it is thorny, red and a flower.
- A learning example is an encounter with a new object a_{m+1}.

Learning Concepts by Example (1)

- You are given something of which you are told that it is called a "rose", and you observe that it is thorny, red and a flower.
- A learning example is an encounter with a new object a_{m+1}.
- Suppose you learn that predicate Q applies to a_{m+1}.

Learning Concepts by Example (1)

- You are given something of which you are told that it is called a “rose”, and you observe that it is thorny, red and a flower.
- A learning example is an encounter with a new object a_{m+1}.
- Suppose you learn that predicate Q applies to a_{m+1}.
- The properties you observe of a_{m+1} are given by $\theta(a_{m+1})$, where $\theta(a_{m+1})$ is a conjunction of $\pm Q_i(a_{m+1})$ for all known predicates.

Learning Concepts by Example (1)

- You are given something of which you are told that it is called a "rose", and you observe that it is thorny, red and a flower.
- A learning example is an encounter with a new object a_{m+1}.
- Suppose you learn that predicate Q applies to a_{m+1}.
- The properties you observe of a_{m+1} are given by $\theta(a_{m+1})$, where $\theta(a_{m+1})$ is a conjunction of $\pm Q_i(a_{m+1})$ for all known predicates.
- Update event: $(a_{m+1}, Q, \theta(a_{m+1}))$.

Learning Concepts by Example (1)

- You are given something of which you are told that it is called a "rose", and you observe that it is thorny, red and a flower.
- A learning example is an encounter with a new object a_{m+1}.
- Suppose you learn that predicate Q applies to a_{m+1}.
- The properties you observe of a_{m+1} are given by $\theta(a_{m+1})$, where $\theta(a_{m+1})$ is a conjunction of $\pm Q_i(a_{m+1})$ for all known predicates.
- Update event: $(a_{m+1}, Q, \theta(a_{m+1}))$.
- You learn that a_{m+1} is called a Q, and you observe that a_{m+1} satisfies the properties $\theta(a_{m+1})$.

Learning Concepts by Example (2)

Learning Concepts by Example (2)

- Updating a model $M = (D, W, P)$ for L_n^m with this will create a new model $M' = (D \cup \{a_{m+1}\}, W', P)$ for L_n^{m+1}.

Learning Concepts by Example (2)

- Updating a model $M = (D, W, P)$ for L_n^m with this will create a new model $M' = (D \cup \{a_{m+1}\}, W', P)$ for L_n^{m+1}.
- New model has domain $\{a_1, \dots, a_{m+1}\}$.

Learning Concepts by Example (2)

- Updating a model $M = (D, W, P)$ for L_n^m with this will create a new model $M' = (D \cup \{a_{m+1}\}, W', P)$ for L_n^{m+1}.
- New model has domain $\{a_1, \dots, a_{m+1}\}$.
- W' is given by assigning, in each w, to a_{m+1} the properties specified by $\theta(a_{m+1})$.

Learning Concepts by Example (2)

- Updating a model $M = (D, W, P)$ for L_n^m with this will create a new model $M' = (D \cup \{a_{m+1}\}, W', P)$ for L_n^{m+1}.
- New model has domain $\{a_1, \ldots, a_{m+1}\}$.
- W' is given by assigning, in each w, to a_{m+1} the properties specified by $\theta(a_{m+1})$.
- Interpretation of Q is given by setting

$$w(Q) = \{a \mid w \models \theta(a/a_{m+1})\}.$$

Learning Concepts by Example (2)

- Updating a model $M = (D, W, P)$ for L_n^m with this will create a new model $M' = (D \cup \{a_{m+1}\}, W', P)$ for L_n^{m+1}.
- New model has domain $\{a_1, \ldots, a_{m+1}\}$.
- W' is given by assigning, in each w, to a_{m+1} the properties specified by $\theta(a_{m+1})$.
- Interpretation of Q is given by setting

$$w(Q) = \{a \mid w \models \theta(a/a_{m+1})\}.$$

- This resets the interpretation Q on the basis of the new observation.

Learning Concepts by Example (2)

- Updating a model $M = (D, W, P)$ for L_n^m with this will create a new model $M' = (D \cup \{a_{m+1}\}, W', P)$ for L_n^{m+1}.
- New model has domain $\{a_1, \ldots, a_{m+1}\}$.
- W' is given by assigning, in each w, to a_{m+1} the properties specified by $\theta(a_{m+1})$.
- Interpretation of Q is given by setting

$$w(Q) = \{a \mid w \models \theta(a/a_{m+1})\}.$$

- This resets the interpretation Q on the basis of the new observation.
- Probability distribution remains unchanged.

Learning Concepts by Example (3)

Learning Concepts by Example (3)

- Account can be refined for cases where the observation is less precise.

Learning Concepts by Example (3)

- Account can be refined for cases where the observation is less precise.
- Learning event:

$$(a_{m+1}, Q, \{(\theta_1(a_{m+1}), q_1), \ldots, (\theta_k(a_{m+1}), q_k)\})$$

Learning Concepts by Example (3)

- Account can be refined for cases where the observation is less precise.
- Learning event:

$$(a_{m+1}, Q, \{(\theta_1(a_{m+1}), q_1), \ldots, (\theta_k(a_{m+1}), q_k)\})$$

- Here q_i gives the observational probability that the new object satisfies θ_i.

Learning Concepts by Example (3)

- Account can be refined for cases where the observation is less precise.
- Learning event:

$$(a_{m+1}, Q, \{(\theta_1(a_{m+1}), q_1), \ldots, (\theta_k(a_{m+1}), q_k)\})$$

- Here q_i gives the observational probability that the new object satisfies θ_i.
- The probabilities should observe $\sum_{i=1}^{k} q_i = 1$.

Learning Concepts by Example (3)

- Account can be refined for cases where the observation is less precise.
- Learning event:

$$(a_{m+1}, Q, \{(\theta_1(a_{m+1}), q_1), \ldots, (\theta_k(a_{m+1}), q_k)\})$$

- Here q_i gives the observational probability that the new object satisfies θ_i.
- The probabilities should observe $\sum_{i=1}^{k} q_i = 1$.
- The update can be defined in such way that the probability of the new predicate applying to the old objects will get recomputed.

A Toy Fragment (1)

A Toy Fragment (1)

- Basic types are e (entities), s (worlds), t (truth values), d (domain size & predicate number restriction) and $[0, 1]$ (the space of probabilities).

A Toy Fragment (1)

- Basic types are e (entities), s (worlds), t (truth values), d (domain size & predicate number restriction) and $[0, 1]$ (the space of probabilities).
- Abbreviate $d \to s \to t$ as i (intensions).

A Toy Fragment (1)

- Basic types are *e* (entities), *s* (worlds), *t* (truth values), *d* (domain size & predicate number restriction) and $[0, 1]$ (the space of probabilities).
- Abbreviate $d \rightarrow s \rightarrow t$ as *i* (intensions).
- Types for S, N, VP, NP, DET get lifted to the level of intensions, by substituting *i* for *t* in all types.

A Toy Fragment (1)

- Basic types are e (entities), s (worlds), t (truth values), d (domain size & predicate number restriction) and $[0, 1]$ (the space of probabilities).
- Abbreviate $d \rightarrow s \rightarrow t$ as i (intensions).
- Types for S, N, VP, NP, DET get lifted to the level of intensions, by substituting i for t in all types.
- This gives, e.g., $\tau(\text{DET}) = (e \rightarrow i) \rightarrow (e \rightarrow i) \rightarrow i$.

A Toy Fragment (1)

- Basic types are e (entities), s (worlds), t (truth values), d (domain size & predicate number restriction) and $[0, 1]$ (the space of probabilities).
- Abbreviate $d \to s \to t$ as i (intensions).
- Types for S, N, VP, NP, DET get lifted to the level of intensions, by substituting i for t in all types.
- This gives, e.g., $\tau(\text{DET}) = (e \to i) \to (e \to i) \to i$.
- The lifting rules for the interpretation functions are completely straightforward.

A Toy Fragment (1)

- Basic types are e (entities), s (worlds), t (truth values), d (domain size & predicate number restriction) and $[0, 1]$ (the space of probabilities).
- Abbreviate $d \to s \to t$ as i (intensions).
- Types for S, N, VP, NP, DET get lifted to the level of intensions, by substituting i for t in all types.
- This gives, e.g., $\tau(\text{DET}) = (e \to i) \to (e \to i) \to i$.
- The lifting rules for the interpretation functions are completely straightforward.
- $I(\text{Some}) = \lambda p \lambda q \lambda d \lambda w.\text{some}(\lambda x.p\ x\ d\ w)(\lambda y.q\ y\ d\ w)$.

A Toy Fragment (1)

- Basic types are e (entities), s (worlds), t (truth values), d (domain size & predicate number restriction) and $[0, 1]$ (the space of probabilities).
- Abbreviate $d \rightarrow s \rightarrow t$ as i (intensions).
- Types for S, N, VP, NP, DET get lifted to the level of intensions, by substituting i for t in all types.
- This gives, e.g., $\tau(\text{DET}) = (e \rightarrow i) \rightarrow (e \rightarrow i) \rightarrow i$.
- The lifting rules for the interpretation functions are completely straightforward.
- $I(\text{Some}) = \lambda p \lambda q \lambda d \lambda w.\textbf{some}(\lambda x.p\ x\ d\ w)(\lambda y.q\ y\ d\ w)$.
- Here some is the familiar constant function for existential quantification, of type $(e \rightarrow t) \rightarrow (e \rightarrow t) \rightarrow t$.

A Toy Fragment (2)

A Toy Fragment (2)

- Sentences get interpretations of type i, i.e., $d \to s \to t$.

A Toy Fragment (2)

- Sentences get interpretations of type i, i.e., $d \to s \to t$.
- Intensions are mapped to probabilities by means of a function prob of type $i \to m \to [0.1]$, where m is the type of models with their domains.

A Toy Fragment (2)

- Sentences get interpretations of type i, i.e., $d \to s \to t$.
- Intensions are mapped to probabilities by means of a function prob of type $i \to m \to [0.1]$, where m is the type of models with their domains.
- The function prob is given by:

$$\text{prob } f\ (D, W, P) = \sum \{P(w) \mid w \in W, f\ D\ w\}.$$

A Toy Fragment (2)

- Sentences get interpretations of type i, i.e., $d \to s \to t$.
- Intensions are mapped to probabilities by means of a function prob of type $i \to m \to [0.1]$, where m is the type of models with their domains.
- The function prob is given by:

$$\text{prob } f\ (D, W, P) = \sum \{P(w) \mid w \in W, f\ D\ w\}.$$

- This assigns to every sentence of the fragment a probability, on the basis of the prior probabilities encoded by (D, W, P).

A Toy Fragment (2)

- Sentences get interpretations of type i, i.e., $d \to s \to t$.
- Intensions are mapped to probabilities by means of a function prob of type $i \to m \to [0.1]$, where m is the type of models with their domains.
- The function prob is given by:

$$\text{prob } f\ (D, W, P) = \sum \{P(w) \mid w \in W, f\ D\ w\}.$$

- This assigns to every sentence of the fragment a probability, on the basis of the prior probabilities encoded by (D, W, P).
- Predicates have type $e \to i$, predicate interpretation functions type String $\to e \to i$.

A Toy Fragment (2)

- Sentences get interpretations of type i, i.e., $d \to s \to t$.
- Intensions are mapped to probabilities by means of a function prob of type $i \to m \to [0.1]$, where m is the type of models with their domains.
- The function prob is given by:

$$\text{prob } f\ (D, W, P) = \sum\{P(w) \mid w \in W, f\ D\ w\}.$$

- This assigns to every sentence of the fragment a probability, on the basis of the prior probabilities encoded by (D, W, P).
- Predicates have type $e \to i$, predicate interpretation functions type String $\to e \to i$.
- Concept definitions have type (String $\to e \to i) \to$ String $\to e \to i$.

A Toy Fragment (2)

- Sentences get interpretations of type i, i.e., $d \to s \to t$.
- Intensions are mapped to probabilities by means of a function prob of type $i \to m \to [0.1]$, where m is the type of models with their domains.
- The function prob is given by:

$$\text{prob } f\ (D, W, P) = \sum\{P(w) \mid w \in W, f\ D\ w\}.$$

- This assigns to every sentence of the fragment a probability, on the basis of the prior probabilities encoded by (D, W, P).
- Predicates have type $e \to i$, predicate interpretation functions type String $\to e \to i$.
- Concept definitions have type (String $\to e \to i$) $\to$ String $\to e \to i$.
- Learning events have type $m \to m$.

Experiments with Semantic Concept Learning

Refinement: Hierarchical Concept Learning

Refinement: Hierarchical Concept Learning

- We encounter a member of a new tribe.

Refinement: Hierarchical Concept Learning

- We encounter a member of a new tribe.
- He looks like this:

Refinement: Hierarchical Concept Learning

- We encounter a member of a new tribe.
- He looks like this:

- Do we conclude that members of the tribe are dark-haired? Yes.

Refinement: Hierarchical Concept Learning

- We encounter a member of a new tribe.
- He looks like this:

- Do we conclude that members of the tribe are dark-haired? Yes.
- Do we conclude that members of the tribe are obese? Most learners do not: Kemp, Perfors and Tenenbaum 2007 [6].

Refinement: Hierarchical Concept Learning

- We encounter a member of a new tribe.
- He looks like this:

- Do we conclude that members of the tribe are dark-haired? Yes.
- Do we conclude that members of the tribe are obese? Most learners do not: Kemp, Perfors and Tenenbaum 2007 [6].
- Program: Hierarchical version of concept learning, using hierarchical Bayesian models; see Kemp, Perfors and Tenenbaum [6].

Conclusions

Conclusions

- It is difficult to make a distinction between learning semantic concepts and learning facts about the world.

Conclusions

- It is difficult to make a distinction between learning semantic concepts and learning facts about the world.
- Try to explain in DEL: Difference between factual change and change in the interpretation of the language.

Conclusions

- It is difficult to make a distinction between learning semantic concepts and learning facts about the world.
- Try to explain in DEL: Difference between factual change and change in the interpretation of the language.
- Solution: Fregean senses? Distinction between (fixed) language of thought and (flexible) concept language?

Conclusions

- It is difficult to make a distinction between learning semantic concepts and learning facts about the world.
- Try to explain in DEL: Difference between factual change and change in the interpretation of the language.
- Solution: Fregean senses? Distinction between (fixed) language of thought and (flexible) concept language?
- To explain: How does semantic concept learning get off the ground? What is the language in which the learner makes her first distinctions?

Conclusions

- It is difficult to make a distinction between learning semantic concepts and learning facts about the world.
- Try to explain in DEL: Difference between factual change and change in the interpretation of the language.
- Solution: Fregean senses? Distinction between (fixed) language of thought and (flexible) concept language?
- To explain: How does semantic concept learning get off the ground? What is the language in which the learner makes her first distinctions?
- To do: Use this model as basis for semantic learning experiments. Cf Alexandru's talk this Friday.

Conclusions

- It is difficult to make a distinction between learning semantic concepts and learning facts about the world.
- Try to explain in DEL: Difference between factual change and change in the interpretation of the language.
- Solution: Fregean senses? Distinction between (fixed) language of thought and (flexible) concept language?
- To explain: How does semantic concept learning get off the ground? What is the language in which the learner makes her first distinctions?
- To do: Use this model as basis for semantic learning experiments. Cf Alexandru's talk this Friday.
- To do: Close the gap between the computational semantics tradition based on type theory, logic, and Montague grammar, and the statistical tradition in natural language processing.

References

Alexandru Baltag and Sonja Smets.
Probabilistic dynamic belief revision.
Synthese, 165(2):179–202, 2008.

J. van Benthem, J. Gerbrandy, and B. Kooi.
Dynamic update with probabilities.
Studia Logica, 93:67–96, 2009.

Martin Gartner.
Mathematical Circus.
Vintage, 1981.

Nina Gierasimszuk.
Bridging learning theory and dynamic epistemic logic.
Synthese, 169:371–374, 2009.

Joseph Y. Halpern.
Reasoning about Uncertainty.
MIT Press, 2003.

Charles Kemp, Amy Perfors, and Joshua B. Tenenbaum.
Learning overhypotheses with hierarchical Bayesian models.
Developmental Science, 10(3):307–321, 2007.

Barteld P. Kooi.
Knowledge, Chance, and Change.
PhD thesis, Groningen University, 2003.