Belief

Recall the axioms that characterize belief:

K1. $(K_i\varphi \wedge K_i(\varphi \Rightarrow \psi)) \Rightarrow K\psi$ (Distribution Axiom).

- or equivalently $K(\varphi \wedge \psi) \Leftrightarrow (K\varphi \wedge K\psi)$

K3. $\neg K$ *false* (Consistency Axiom).

K4. $K\varphi \Rightarrow KK\varphi$ (Positive Introspection Axiom).

K5. $\neg K\varphi \Rightarrow K\neg K\varphi$ (Negative Introspection Axiom).

Belief

Recall the axioms that characterize belief:

K1. $(K_i\varphi \wedge K_i(\varphi \Rightarrow \psi)) \Rightarrow K\psi$ (Distribution Axiom).

- or equivalently $K(\varphi \wedge \psi) \Leftrightarrow (K\varphi \wedge K\psi)$

K3. $\neg K\,\mathit{false}$ (Consistency Axiom).

K4. $K\varphi \Rightarrow KK\varphi$ (Positive Introspection Axiom).

K5. $\neg K\varphi \Rightarrow K\neg K\varphi$ (Negative Introspection Axiom).

These properties are sound in an epistemic structure if $\mathcal{K}$ is serial (K3), transitive (K4), and Euclidean (K5)

- Easy to check: $\mathcal{K}$ is Euclidean and transitive iff $w' \in \mathcal{K}(w) \Rightarrow \mathcal{K}(w) = \mathcal{K}(w')$ for all s, t
  - This looks just like the uniformity property!

Defining belief using probability

K1. $(K\varphi \wedge K(\varphi \Rightarrow \psi)) \Rightarrow K\psi$ (Distribution Axiom).

- or equivalently $K(\varphi \wedge \psi) \Leftrightarrow (K\varphi \wedge K\psi)$

K3. $\neg K\mathit{false}$ (Consistency Axiom).

K4. $K\varphi \Rightarrow KK\varphi$ (Positive Introspection Axiom).

K5. $\neg K\varphi \Rightarrow K\neg K\varphi$ (Negative Introspection Axiom).

In a probability structure, take $(M, w) \models B\varphi$ if $\Pr_w([\![\varphi]\!]_M) = 1$.

- To get K4 and K5, assuming UNIF:
  - if $w' \in W_{w'}$, then $\mathcal{PR}(w) = \mathcal{PR}(w')$

Defining belief using probability

K1. $(K\varphi \wedge K(\varphi \Rightarrow \psi)) \Rightarrow K\psi$ (Distribution Axiom).

- or equivalently $K(\varphi \wedge \psi) \Leftrightarrow (K\varphi \wedge K\psi)$

K3. $\neg K\mathit{false}$ (Consistency Axiom).

K4. $K\varphi \Rightarrow KK\varphi$ (Positive Introspection Axiom).

K5. $\neg K\varphi \Rightarrow K\neg K\varphi$ (Negative Introspection Axiom).

In a probability structure, take $(M, w) \models B\varphi$ if $\Pr_w([\![\varphi]\!]_M) = 1$.

- To get K4 and K5, assuming UNIF:
  - if $w' \in W_{w'}$, then $\mathcal{PR}(w) = \mathcal{PR}(w')$

So know we have two semantics for belief:

1. Using an accessibility relation in a Kripke structure
2. Defining belief as probability 1

The next slide gives a third semantics, which will play a major role in our discussions, that uses plausibility measures.

Defining belief using plausibility

In a plausibility space $M = (W, \mathcal{PL}, \pi)$, define

$$(M, w) \models B\varphi \text{ iff } \mathrm{Pl}_w([\![\varphi]\!]_M) > \mathrm{Pl}_w([\![\neg\varphi]\!]_M).$$

- I believe φ if φ is more plausible than not

Again, if UNIF holds, then we get K4 and K5. But what do we need to assume to get $B(\varphi \wedge \psi) \Leftrightarrow (B\varphi \wedge B\psi)$?

Defining belief using plausibility

In a plausibility space $M = (W, \mathcal{PL}, \pi)$, define

$$(M, w) \models B\varphi \text{ iff } \mathrm{Pl}_w([\![\varphi]\!]_M) > \mathrm{Pl}_w([\![\neg\varphi]\!]_M).$$

- I believe φ if φ is more plausible than not

Again, if UNIF holds, then we get K4 and K5. But what do we need to assume to get $B(\varphi \wedge \psi) \Leftrightarrow (B\varphi \wedge B\psi)$?

Reverse engineering shows that the following does the trick:

Pl4″. If $\mathrm{Pl}(U_1) > \mathrm{Pl}(\overline{U_1})$ and $\mathrm{Pl}(U_2) > \mathrm{Pl}(\overline{U_2})$, then $\mathrm{Pl}(U_1 \cap U_2) > \mathrm{Pl}(\overline{U_1 \cap U_2})$.

Proposition: $B_i(\varphi \wedge \psi) \Leftrightarrow B_i\varphi \wedge B_i\psi$ holds iff Pl satisfies Pl4″.

Pl4″. If $\mathrm{Pl}(U_1) > \mathrm{Pl}(\overline{U_1})$ and $\mathrm{Pl}(U_2) > \mathrm{Pl}(\overline{U_2})$, then $\mathrm{Pl}(U_1 \cap U_2) > \mathrm{Pl}(\overline{U_1 \cap U_2})$.

For reasons that will shortly become clear, I want a version of Pl4″ conditional on V (i.e., if we intersect everything with V):

Pl4′. If $\mathrm{Pl}(U_1 \cap V) > \mathrm{Pl}(\overline{U_1} \cap V)$ and $\mathrm{Pl}(U_2 \cap V) > \mathrm{Pl}(\overline{U_2} \cap V)$, then $\mathrm{Pl}(U_1 \cap U_2 \cap V) > \mathrm{Pl}(\overline{U_1 \cap U_2} \cap V)$.

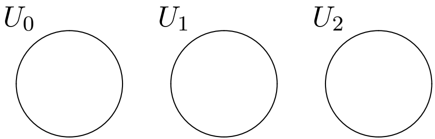
Pl4″. If $\mathrm{Pl}(U_1) > \mathrm{Pl}(\overline{U_1})$ and $\mathrm{Pl}(U_2) > \mathrm{Pl}(\overline{U_2})$, then $\mathrm{Pl}(U_1 \cap U_2) > \mathrm{Pl}(\overline{U_1 \cap U_2})$.

For reasons that will shortly become clear, I want a version of Pl4″ conditional on V (i.e., if we intersect everything with V):

Pl4′. If $\mathrm{Pl}(U_1 \cap V) > \mathrm{Pl}(\overline{U_1} \cap V)$ and $\mathrm{Pl}(U_2 \cap V) > \mathrm{Pl}(\overline{U_2} \cap V)$, then $\mathrm{Pl}(U_1 \cap U_2 \cap V) > \mathrm{Pl}(\overline{U_1 \cap U_2} \cap V)$.

Pl4′ is a little ugly. (Given Pl3) PL4′ is equivalent to

Pl4. If U_0, U_1, and U_2 are pairwise disjoint, $\mathrm{Pl}(U_0 \cup U_1) > \mathrm{Pl}(U_2)$, and $\mathrm{Pl}(U_0 \cup U_2) > \mathrm{Pl}(U_1)$, then $\mathrm{Pl}(U_0) > \mathrm{Pl}(U_1 \cup U_2)$.



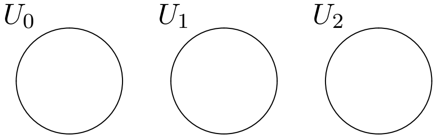
Pl4″. If $\mathrm{Pl}(U_1) > \mathrm{Pl}(\overline{U_1})$ and $\mathrm{Pl}(U_2) > \mathrm{Pl}(\overline{U_2})$, then $\mathrm{Pl}(U_1 \cap U_2) > \mathrm{Pl}(\overline{U_1 \cap U_2})$.

For reasons that will shortly become clear, I want a version of Pl4″ conditional on V (i.e., if we intersect everything with V):

Pl4′. If $\mathrm{Pl}(U_1 \cap V) > \mathrm{Pl}(\overline{U_1} \cap V)$ and $\mathrm{Pl}(U_2 \cap V) > \mathrm{Pl}(\overline{U_2} \cap V)$, then $\mathrm{Pl}(U_1 \cap U_2 \cap V) > \mathrm{Pl}(\overline{U_1 \cap U_2} \cap V)$.

Pl4′ is a little ugly. (Given Pl3) PL4′ is equivalent to

Pl4. If U_0, U_1, and U_2 are pairwise disjoint, $\mathrm{Pl}(U_0 \cup U_1) > \mathrm{Pl}(U_2)$, and $\mathrm{Pl}(U_0 \cup U_2) > \mathrm{Pl}(U_1)$, then $\mathrm{Pl}(U_0) > \mathrm{Pl}(U_1 \cup U_2)$.



Sometimes mathematics can be more beautiful than you have any right to expect :-)

Pl4

Pl4. If U_0, U_1, and U_2 are pairwise disjoint, $\mathrm{Pl}(U_0 \cup U_1) > \mathrm{Pl}(U_2)$, and $\mathrm{Pl}(U_0 \cup U_2) > \mathrm{Pl}(U_1)$, then $\mathrm{Pl}(U_0) > \mathrm{Pl}(U_1 \cup U_2)$.

Proposition:

(a) Probability does not in general satisfy Pl4.
(b) Possibility measures satisfy Pl4.
(c) Ranking functions satisfy Pl4.

Pl4

Pl4. If U_0, U_1, and U_2 are pairwise disjoint, $\mathrm{Pl}(U_0 \cup U_1) > \mathrm{Pl}(U_2)$, and $\mathrm{Pl}(U_0 \cup U_2) > \mathrm{Pl}(U_1)$, then $\mathrm{Pl}(U_0) > \mathrm{Pl}(U_1 \cup U_2)$.

Proposition:

(a) Probability does not in general satisfy Pl4.
(b) Possibility measures satisfy Pl4.
(c) Ranking functions satisfy Pl4.

Proof: (a) If $\Pr(U_0) = \Pr(U_1) = \Pr(U_2) = 1/3$, then $\Pr(U_0 \cup U_1) > \Pr(U_2)$, $\Pr(U_0 \cup U_2) > \Pr(U_1)$, but $\Pr(U_0) < \Pr(U_1 \cup U_2)$.

Pl4

Pl4. If U_0, U_1, and U_2 are pairwise disjoint, $\mathrm{Pl}(U_0 \cup U_1) > \mathrm{Pl}(U_2)$, and $\mathrm{Pl}(U_0 \cup U_2) > \mathrm{Pl}(U_1)$, then $\mathrm{Pl}(U_0) > \mathrm{Pl}(U_1 \cup U_2)$.

Proposition:

(a) Probability does not in general satisfy Pl4.
(b) Possibility measures satisfy Pl4.
(c) Ranking functions satisfy Pl4.

Proof: (a) If $\Pr(U_0) = \Pr(U_1) = \Pr(U_2) = 1/3$, then $\Pr(U_0 \cup U_1) > \Pr(U_2)$, $\Pr(U_0 \cup U_2) > \Pr(U_1)$, but $\Pr(U_0) < \Pr(U_1 \cup U_2)$.

(b) If $\mathrm{Poss}(U_0 \cup U_1) > \mathrm{Poss}(U_2)$ and $\mathrm{Poss}(U_0 \cup U_2) > \mathrm{Poss}(U_1)$, then the element(s) of greatest possibility in $U_0 \cup U_1 \cup U_2$ must all be in U_0

- Recall: $\mathrm{Poss}(V) = \max_{x \in V} \mathrm{Poss}(x)$.

Thus, $\mathrm{Poss}(U_0) > \mathrm{Poss}(U_1 \cup U_2)$.

Pl4

Pl4. If U_0, U_1, and U_2 are pairwise disjoint, $\mathrm{Pl}(U_0 \cup U_1) > \mathrm{Pl}(U_2)$, and $\mathrm{Pl}(U_0 \cup U_2) > \mathrm{Pl}(U_1)$, then $\mathrm{Pl}(U_0) > \mathrm{Pl}(U_1 \cup U_2)$.

Proposition:

(a) Probability does not in general satisfy Pl4.
(b) Possibility measures satisfy Pl4.
(c) Ranking functions satisfy Pl4.

Proof: (a) If $\Pr(U_0) = \Pr(U_1) = \Pr(U_2) = 1/3$, then $\Pr(U_0 \cup U_1) > \Pr(U_2)$, $\Pr(U_0 \cup U_2) > \Pr(U_1)$, but $\Pr(U_0) < \Pr(U_1 \cup U_2)$.

(b) If $\mathrm{Poss}(U_0 \cup U_1) > \mathrm{Poss}(U_2)$ and $\mathrm{Poss}(U_0 \cup U_2) > \mathrm{Poss}(U_1)$, then the element(s) of greatest possibility in $U_0 \cup U_1 \cup U_2$ must all be in U_0

- Recall: $\mathrm{Poss}(V) = \max_{x \in V} \mathrm{Poss}(x)$.

Thus, $\mathrm{Poss}(U_0) > \mathrm{Poss}(U_1 \cup U_2)$.

(c) The argument for ranking functions is the same; just replace max by min.

Default Reasoning

Defaults are statements of the form $\varphi \to \psi$ that are read "if φ then typically/normally/by default ψ". Defaults appear naturally in many applications.

An (in)famous example:

- *bird* $\to$ *fly*
- *penguin* $\to \neg$*fly*
- *penguin* $\to$ *bird*

Default reasoning is not *monotonic*:

- If $A \to C$, then we may not have $A \wedge B \to C$
- We have *bird* $\to$ *fly*, but not *bird* $\wedge$ *penguin* $\to$ *fly*

Nevertheless, it seems that it should obey certain properties. But which ones? While there is no consensus, there is an agreement on a common "core" of default reasoning.

- Kraus, Lehmann and Magidor suggest a set of properties (the "KLM properties", also known as "System $\mathbf{P}$")

The KLM Properties/System $\mathbf{P}$

(LLE) If $\varphi \equiv \varphi'$, then from $\varphi \to \psi$ infer $\varphi' \to \psi$.

The KLM Properties/System $\mathbf{P}$

(LLE) If $\varphi \equiv \varphi'$, then from $\varphi \to \psi$ infer $\varphi' \to \psi$.

(REF) Always infer $\varphi \to \varphi$.

The KLM Properties/System $\mathbf{P}$

(LLE) If $\varphi \equiv \varphi'$, then from $\varphi \to \psi$ infer $\varphi' \to \psi$.

(REF) Always infer $\varphi \to \varphi$.

(RW) If $\psi \Rightarrow \psi'$, then from $\varphi \to \psi$ infer $\varphi \to \psi'$.

The KLM Properties/System $\mathbf{P}$

(LLE) If $\varphi \equiv \varphi'$, then from $\varphi \to \psi$ infer $\varphi' \to \psi$.

(REF) Always infer $\varphi \to \varphi$.

(RW) If $\psi \Rightarrow \psi'$, then from $\varphi \to \psi$ infer $\varphi \to \psi'$.

(AND) From $\varphi \to \psi_1$ and $\varphi \to \psi_2$ infer $\varphi \to \psi_1 \wedge \psi_2$.

- Since birds typically fly and birds typically have have wings, birds typically (fly and have wings).

The KLM Properties/System $\mathbf{P}$

(LLE) If $\varphi \equiv \varphi'$, then from $\varphi \to \psi$ infer $\varphi' \to \psi$.

(REF) Always infer $\varphi \to \varphi$.

(RW) If $\psi \Rightarrow \psi'$, then from $\varphi \to \psi$ infer $\varphi \to \psi'$.

(AND) From $\varphi \to \psi_1$ and $\varphi \to \psi_2$ infer $\varphi \to \psi_1 \wedge \psi_2$.

- Since birds typically fly and birds typically have have wings, birds typically (fly and have wings).

(OR) If $\varphi_1 \to \psi$ and $\varphi_2 \to \psi$ infer $\varphi_1 \vee \varphi_2 \to \psi$.

- Since birds typically fly and insects typically fly, things that are either birds or insects typically fly.

The KLM Properties/System $\mathbf{P}$

(LLE) If $\varphi \equiv \varphi'$, then from $\varphi \rightarrow \psi$ infer $\varphi' \rightarrow \psi$.

(REF) Always infer $\varphi \rightarrow \varphi$.

(RW) If $\psi \Rightarrow \psi'$, then from $\varphi \rightarrow \psi$ infer $\varphi \rightarrow \psi'$.

(AND) From $\varphi \rightarrow \psi_1$ and $\varphi \rightarrow \psi_2$ infer $\varphi \rightarrow \psi_1 \wedge \psi_2$.

- Since birds typically fly and birds typically have have wings, birds typically (fly and have wings).

(OR) If $\varphi_1 \rightarrow \psi$ and $\varphi_2 \rightarrow \psi$ infer $\varphi_1 \vee \varphi_2 \rightarrow \psi$.

- Since birds typically fly and insects typically fly, things that are either birds or insects typically fly.

(CM) From $\varphi \rightarrow \psi_1$ and $\varphi \rightarrow \psi_2$ infer $\varphi \wedge \psi_1 \rightarrow \psi_2$.

- Since birds typically fly and birds typically have have wings, birds that fly typically have wings.

The KLM Properties/System $\mathbf{P}$

(LLE) If $\varphi \equiv \varphi'$, then from $\varphi \to \psi$ infer $\varphi' \to \psi$.

(REF) Always infer $\varphi \to \varphi$.

(RW) If $\psi \Rightarrow \psi'$, then from $\varphi \to \psi$ infer $\varphi \to \psi'$.

(AND) From $\varphi \to \psi_1$ and $\varphi \to \psi_2$ infer $\varphi \to \psi_1 \wedge \psi_2$.

- Since birds typically fly and birds typically have have wings, birds typically (fly and have wings).

(OR) If $\varphi_1 \to \psi$ and $\varphi_2 \to \psi$ infer $\varphi_1 \vee \varphi_2 \to \psi$.

- Since birds typically fly and insects typically fly, things that are either birds or insects typically fly.

(CM) From $\varphi \to \psi_1$ and $\varphi \to \psi_2$ infer $\varphi \wedge \psi_1 \to \psi_2$.

- Since birds typically fly and birds typically have have wings, birds that fly typically have wings.

The latter three properties get at the heart of default reasoning.

- All these properties hold if $\to$ is interpreted as $\Rightarrow$, but we can have other interpretations.
  - Necessary for the bird/penguin example to be consistent!

The KLM Properties/System $\mathbf{P}$

(LLE) If $\varphi \equiv \varphi'$, then from $\varphi \to \psi$ infer $\varphi' \to \psi$.

(REF) Always infer $\varphi \to \varphi$.

(RW) If $\psi \Rightarrow \psi'$, then from $\varphi \to \psi$ infer $\varphi \to \psi'$.

(AND) From $\varphi \to \psi_1$ and $\varphi \to \psi_2$ infer $\varphi \to \psi_1 \wedge \psi_2$.

- Since birds typically fly and birds typically have have wings, birds typically (fly and have wings).

(OR) If $\varphi_1 \to \psi$ and $\varphi_2 \to \psi$ infer $\varphi_1 \vee \varphi_2 \to \psi$.

- Since birds typically fly and insects typically fly, things that are either birds or insects typically fly.

(CM) From $\varphi \to \psi_1$ and $\varphi \to \psi_2$ infer $\varphi \wedge \psi_1 \to \psi_2$.

- Since birds typically fly and birds typically have have wings, birds that fly typically have wings.

The latter three properties get at the heart of default reasoning.

- All these properties hold if $\to$ is interpreted as $\Rightarrow$, but we can have other interpretations.
  - Necessary for the bird/penguin example to be consistent!

Many different proposals have been made in the literature for giving semantics to defaults.

Defaults as Extreme Probabilities

One interpretation:

- "*bird* $\rightarrow$ *fly*" means that $\Pr([\![\mathit{fly}]\!] \mid [\![\mathit{bird}]\!])$ is high.
  - But how high should it be?
    - If we require $\Pr([\![\mathit{fly}]\!] \mid [\![\mathit{bird}]\!]) \geq 1 - \epsilon$ for $\epsilon > 0$, then the And rule will not hold.
    - We can have $\Pr(A \mid C) > 1 - \epsilon$, $\Pr(B \mid C) > 1 - \epsilon$, but $\Pr(A \wedge B \mid C) < 1 - \epsilon$.
  - But if we take the probability to be 1, then we can't capture the bird/penguin example: $\{\Pr([\![\mathit{fly}]\!] \mid [\![\mathit{bird}]\!]) = 1, \Pr([\![\mathit{bird}]\!] \mid [\![\mathit{penguin}]\!]) = 1, \Pr([\![\mathit{fly}]\!] | [\![\mathit{penguin}]\!]) = 0\}$ is inconsistent.
- We'd like to say $\Pr([\![\mathit{fly}]\!] \mid [\![\mathit{bird}]\!]) \approx 1$, but how?

Defaults as Extreme Probabilities

One interpretation:

- "*bird* $\rightarrow$ *fly*" means that $\Pr([\![\mathit{fly}]\!] \mid [\![\mathit{bird}]\!])$ is high.
  - But how high should it be?
    - If we require $\Pr([\![\mathit{fly}]\!] \mid [\![\mathit{bird}]\!]) \geq 1 - \epsilon$ for $\epsilon > 0$, then the And rule will not hold.
    - We can have $\Pr(A \mid C) > 1 - \epsilon$, $\Pr(B \mid C) > 1 - \epsilon$, but $\Pr(A \wedge B \mid C) < 1 - \epsilon$.
  - But if we take the probability to be 1, then we can't capture the bird/penguin example: $\{\Pr([\![\mathit{fly}]\!] \mid [\![\mathit{bird}]\!]) = 1, \Pr([\![\mathit{bird}]\!] \mid [\![\mathit{penguin}]\!]) = 1, \Pr([\![\mathit{fly}]\!] | [\![\mathit{penguin}]\!]) = 0\}$ is inconsistent.
- We'd like to say $\Pr([\![\mathit{fly}]\!] \mid [\![\mathit{bird}]\!]) \approx 1$, but how?
- Idea (due independently to Judea Pearl, a computer scientist, and Earnest Adams, a philosopher): use, not one probability distribution, but a *probability sequence* $(\Pr_1, \Pr_2, \ldots)$ and consider limiting probability
  - These are all probability distribution on the same space W.

Simple PS structures

A *simple PS structure* is a tuple $M = (W, (\Pr_1, \Pr_2, ...), \pi)$.

$$M \models \varphi \to \psi \text{ if } \lim_{n\to\infty} \Pr_n([\![\psi]\!]_M \mid [\![\varphi]\!]_M) = 1$$

(where $\Pr_n(U \mid V) = 1$ if $\Pr_n(V) = 0$).

- Intuitively, this means that (according to $\vec{\Pr} = (\Pr_1, \Pr_2, \ldots)$) the probability of the default is arbitrarily close to 1.
- This is independent of the world w (I wrote $M \models$, not $(M, w) \models$), since $\vec{\Pr}$ is independent of w.
- Last class, Meir asked where the sequence of probability measures comes from in real-world settings?
  - I'm not sure that Pearl or Adams had real worlds settings in mind, but in a few weeks, I'll give you a relatively recent concrete application (involving security and cryptography).
    - In security applications, there is a *security parameter* – the length of the key used to do encryption.
    - Longer keys provide greater security.
    - $\Pr_n$ describes the probability of breaking the encryption using a key of length n.
    - This lets us make sense of statements like "typically, the encryption is hard to break".

To repeat, in a PS structure M,

$$M \models \varphi \to \psi \text{ if } \lim_{n\to\infty} \Pr_n(\llbracket\psi\rrbracket_M \mid \llbracket\varphi\rrbracket_M) = 1$$

This semantics satisfies the KLM properties. In fact, the KLM properties characterize this semantics.

To repeat, in a PS structure M,

$$M \models \varphi \to \psi \text{ if } \lim_{n\to\infty} \Pr_n(\llbracket\psi\rrbracket_M \mid \llbracket\varphi\rrbracket_M) = 1$$

This semantics satisfies the KLM properties. In fact, the KLM properties characterize this semantics. Some notation: if Σ is a finite set of defaults, then write

- $M \models \Sigma$ if $M \models \sigma$ for every formula $\sigma \in \Sigma$.

To repeat, in a PS structure M,

$$M \models \varphi \rightarrow \psi \text{ if } \lim_{n\to\infty} \Pr_n([\![\psi]\!]_M \mid [\![\varphi]\!]_M) = 1$$

This semantics satisfies the KLM properties. In fact, the KLM properties characterize this semantics. Some notation: if Σ is a

finite set of defaults, then write

- $M \models \Sigma$ if $M \models \sigma$ for every formula $\sigma \in \Sigma$.
- $\Sigma \models_{PS} \varphi$ holds if, for all PS structures M, if $M \models \Sigma$ then $M \models \varphi$.

To repeat, in a PS structure M,

$$M \models \varphi \to \psi \text{ if } \lim_{n\to\infty} \Pr_n([\![\psi]\!]_M \mid [\![\varphi]\!]_M) = 1$$

This semantics satisfies the KLM properties. In fact, the KLM properties characterize this semantics. Some notation: if Σ is a finite set of defaults, then write

- $M \models \Sigma$ if $M \models \sigma$ for every formula $\sigma \in \Sigma$.
- $\Sigma \models_{PS} \varphi$ holds if, for all PS structures M, if $M \models \Sigma$ then $M \models \varphi$.
- $\Sigma \vdash_{\mathbf{P}} \varphi$ holds if φ is provable from Σ using the KLM properties.
  - e.g., $\{p \wedge q \to r, p \to \neg r\} \vdash_{\mathbf{P}} p \to \neg q$.

To repeat, in a PS structure M,

$$M \models \varphi \to \psi \text{ if } \lim_{n\to\infty} \Pr_n([\![\psi]\!]_M \mid [\![\varphi]\!]_M) = 1$$

This semantics satisfies the KLM properties. In fact, the KLM properties characterize this semantics. Some notation: if Σ is a

finite set of defaults, then write

- $M \models \Sigma$ if $M \models \sigma$ for every formula $\sigma \in \Sigma$.
- $\Sigma \models_{PS} \varphi$ holds if, for all PS structures M, if $M \models \Sigma$ then $M \models \varphi$.
- $\Sigma \vdash_{\mathbf{P}} \varphi$ holds if φ is provable from Σ using the KLM properties.
  - e.g., $\{p \land q \to r, p \to \neg r\} \vdash_{\mathbf{P}} p \to \neg q$.

Theorem: The KLM properties are sound and complete for simple PS structures:

$$\Sigma \models_{PS} \varphi \text{ iff } \Sigma \vdash_{\mathbf{P}} \varphi$$

Example

Suppose that $W = \{w_1, w_2, w_3, w_4\}$ and π is such that

- $(M_2, w_1) \models$ *bird* $\wedge$ *fly* $\wedge \neg$*penguin*,
- $(M_2, w_2) \models$ *bird* $\wedge \neg$*fly* $\wedge$ *penguin*,
- $(M_2, w_3) \models$ *bird* $\wedge$ *fly* $\wedge$ *penguin*,
- $(M_2, w_4) \models$ *bird* $\wedge \neg$*fly* $\wedge \neg$*penguin*,

Suppose that

- $\Pr_n(w_1) = 1 - 1/n - 2/n^2$
- $\Pr_n(w_2) = 1/n$
- $\Pr_n(w_3) = 1/n^2$
- $\Pr_n(w_4) = 1/n^2$

Note that $\Pr_n([\![\neg fly]\!]_M \mid [\![\mathit{penguin}]\!]_M) = (1/n)/(1/n + 1/n^2)$.

Example

Suppose that $W = \{w_1, w_2, w_3, w_4\}$ and π is such that

- $(M_2, w_1) \models$ *bird* $\wedge$ *fly* $\wedge \neg$*penguin*,
- $(M_2, w_2) \models$ *bird* $\wedge \neg$*fly* $\wedge$ *penguin*,
- $(M_2, w_3) \models$ *bird* $\wedge$ *fly* $\wedge$ *penguin*,
- $(M_2, w_4) \models$ *bird* $\wedge \neg$*fly* $\wedge \neg$*penguin*,

Suppose that

- $\Pr_n(w_1) = 1 - 1/n - 2/n^2$
- $\Pr_n(w_2) = 1/n$
- $\Pr_n(w_3) = 1/n^2$
- $\Pr_n(w_4) = 1/n^2$

Note that $\Pr_n([\![\neg fly]\!]_M \mid [\![\textit{penguin}]\!]_M) = (1/n)/(1/n + 1/n^2)$.
Let $M = (W, (\Pr_n), \pi)$. It is easy to check that

$M \models$ (*bird* $\rightarrow$ *fly*) $\wedge$ (*penguin* $\rightarrow$ *bird*) $\wedge$ (*penguin* $\rightarrow \neg$*fly*)

Semantics for defaults using preferential structures

(This was KLM's semantics for $\rightarrow$:)

Suppose that we have a partial order $\succeq$ on worlds:

- $w \succeq w'$ means that w is "at least as normal" as w'
  - The worlds where penguins don't fly is more normal than the world where penguins fly.

Let $M = (W, \succeq, \pi)$ be a *simple preferential structure*:

- Given $U \subseteq W$, define $best_M(U)$ to be the most normal worlds in U (according to $\succeq$):

$$\mathrm{best}_M(U) = \{w \in U : \text{ for all } w' \in U,\ w' \not\succ w\}.$$

Semantics for defaults using preferential structures

(This was KLM's semantics for $\rightarrow$:)

Suppose that we have a partial order $\succeq$ on worlds:

- $w \succeq w'$ means that w is "at least as normal" as w'
  - The worlds where penguins don't fly is more normal than the world where penguins fly.

Let $M = (W, \succeq, \pi)$ be a *simple preferential structure*:

- Given $U \subseteq W$, define $best_M(U)$ to be the most normal worlds in U (according to $\succeq$):

$$\mathrm{best}_M(U) = \{w \in U : \text{ for all } w' \in U,\ w' \not\succ w\}.$$

- $M \models \varphi \rightarrow \psi$ iff $\mathrm{best}_M([\![\varphi]\!]_M) \subseteq [\![\psi]\!]_M$.
  - The most normal worlds satisfying φ also satisfy ψ

Write $\Sigma \models_{Pref} \varphi$ if, for all preferential models M, if $M \models \Sigma$ then $M \models \varphi$.

Semantics for defaults using preferential structures

(This was KLM's semantics for $\rightarrow$:)

Suppose that we have a partial order $\succeq$ on worlds:

- $w \succeq w'$ means that w is "at least as normal" as w'
  - The worlds where penguins don't fly is more normal than the world where penguins fly.

Let $M = (W, \succeq, \pi)$ be a *simple preferential structure*:

- Given $U \subseteq W$, define $best_M(U)$ to be the most normal worlds in U (according to $\succeq$):

$$\mathrm{best}_M(U) = \{w \in U : \text{ for all } w' \in U,\ w' \not\succ w\}.$$

- $M \models \varphi \rightarrow \psi$ iff $\mathrm{best}_M([\![\varphi]\!]_M) \subseteq [\![\psi]\!]_M$.
  - The most normal worlds satisfying φ also satisfy ψ

Write $\Sigma \models_{Pref} \varphi$ if, for all preferential models M, if $M \models \Sigma$ then $M \models \varphi$.

Theorem: The KLM properties are sound and complete for simple preferential structures:

$$\Sigma \models_{Pref} \varphi \text{ iff } \Sigma \vdash_{\mathbf{P}} \varphi$$

Example

Suppose that $W = \{w_1, w_2, w_3, w_4\}$ and π is such that

- $(M_2, w_1) \models \mathit{bird} \land \mathit{fly} \land \neg \mathit{penguin}$,
- $(M_2, w_2) \models \mathit{bird} \land \neg \mathit{fly} \land \mathit{penguin}$,
- $(M_2, w_3) \models \mathit{bird} \land \mathit{fly} \land \mathit{penguin}$,
- $(M_2, w_4) \models \mathit{bird} \land \neg \mathit{fly} \land \neg \mathit{penguin}$,

Suppose that $w_1 \succ w_2 \succ w_3$, $w_2 \succ w_4$, and w_3 and w_4 are incomparable. Let $M = (W, \succeq, \pi)$. Then

$M \models (\mathit{bird} \to \mathit{fly}) \land (\mathit{penguin} \to \mathit{bird}) \land (\mathit{penguin} \to \neg \mathit{fly})$

Semantics for defaults using possibility measures

A *simple possibility structure* is a tuple $M = (W, \mathrm{Poss}, \pi)$.

Define $(M, w) \models \varphi \to \psi$ iff
$\mathrm{Poss}([\![\varphi]\!]_M) = 0$ or $\mathrm{Poss}([\![\varphi \wedge \psi]\!]_M) > \mathrm{Poss}([\![\varphi \wedge \neg\psi]\!]_M)$

- Intuitively, we are saying that $\mathrm{Poss}([\![psi]\!]_M \mid [\![\varphi]\!]_M) > \mathrm{Poss}([\![\neg\psi]\!]_M \mid [\![)]\!]_\varphi$
- the first clause $\mathrm{Poss}([\![\varphi]\!]_M) = 0$ deals with the case that conditioning isn't well defined.
  - In this case we take the default to be vacuously true

Semantics for defaults using possibility measures

A *simple possibility structure* is a tuple $M = (W, \mathrm{Poss}, \pi)$.

Define $(M, w) \models \varphi \to \psi$ iff
$\mathrm{Poss}([\![\varphi]\!]_M) = 0$ or $\mathrm{Poss}([\![\varphi \wedge \psi]\!]_M) > \mathrm{Poss}([\![\varphi \wedge \neg\psi]\!]_M)$

- Intuitively, we are saying that $\mathrm{Poss}([\![psi]\!]_M \mid [\![\varphi]\!]_M) > \mathrm{Poss}([\![\neg\psi]\!]_M \mid [\![)]\!]_\varphi$
- the first clause $\mathrm{Poss}([\![\varphi]\!]_M) = 0$ deals with the case that conditioning isn't well defined.
  - In this case we take the default to be vacuously true

Theorem: The KLM properties are sound and complete for simple possibility structures:

$$\Sigma \models_{Poss} \varphi \text{ iff } \Sigma \vdash_{\mathbf{P}} \varphi$$

Semantics for defaults using possibility measures

A *simple possibility structure* is a tuple $M = (W, \mathrm{Poss}, \pi)$.

Define $(M, w) \models \varphi \rightarrow \psi$ iff
$\mathrm{Poss}([\![\varphi]\!]_M) = 0$ or $\mathrm{Poss}([\![\varphi \wedge \psi]\!]_M) > \mathrm{Poss}([\![\varphi \wedge \neg\psi]\!]_M)$

- Intuitively, we are saying that
  $\mathrm{Poss}([\![psi]\!]_M \mid [\![\varphi]\!]_M) > \mathrm{Poss}([\![\neg\psi]\!]_M \mid [\![)]\!]_\varphi$
- the first clause $\mathrm{Poss}([\![\varphi]\!]_M) = 0$ deals with the case that conditioning isn't well defined.
  - In this case we take the default to be vacuously true

Theorem: The KLM properties are sound and complete for simple possibility structures:

$$\Sigma \models_{Poss} \varphi \text{ iff } \Sigma \vdash_{\mathbf{P}} \varphi$$

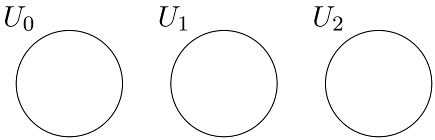
We can do the same thing with κ rankings:
Define $M \models \varphi \rightarrow \psi$ iff
$[\![\varphi]\!]_M = \emptyset$ or $\kappa([\![\varphi \wedge \psi]\!]_M) < \kappa([\![\varphi \wedge \neg\psi]\!]_M)$
And again we get that system $\mathbf{P}$ is sound and complete.

Recall the KLM properties (System $\mathbf{P}$) that are intended to characterize the core of default reasoning:

- **(LLE)** If $\varphi \equiv \varphi'$, then from $\varphi \to \psi$ infer $\varphi' \to \psi$.
- **(REF)** Always infer $\varphi \to \varphi$.
- **(RW)** If $\psi \Rightarrow \psi'$, then from $\varphi \to \psi$ infer $\varphi \to \psi'$.
- **(AND)** From $\varphi \to \psi_1$ and $\varphi \to \psi_2$ infer $\varphi \to \psi_1 \wedge \psi_2$.
- **(OR)** If $\varphi_1 \to \psi$ and $\varphi_2 \to \psi$ infer $\varphi_1 \vee \varphi_2 \to \psi$.

Also recall

Pl4. If U_0, U_1, and U_2 are pairwise disjoint, $\mathrm{Pl}(U_0 \cup U_1) > \mathrm{Pl}(U_2)$, and $\mathrm{Pl}(U_0 \cup U_2) > \mathrm{Pl}(U_1)$, then $\mathrm{Pl}(U_0) > \mathrm{Pl}(U_1 \cup U_2)$.



What's going on?

We've seen four quite different semantics for defaults:

- One uses sequences of probability masures
- One uses a partial order on worlds: on probability in sight
- One uses a single possibilty measure
- The fourth uses a single ranking function

What's going on?

We've seen four quite different semantics for defaults:

- One uses sequences of probability masures
- One uses a partial order on worlds: on probability in sight
- One uses a single possibilty measure
- The fourth uses a single ranking function

[Geffner/Pearl:]

- Why do we always end up with System $\mathbf{P}$?
- Why can't we go beyond System $\mathbf{P}$ (i.e., get additional axioms/properties)
  - That's what the authors of these approaches were hoping to do!

Plausibility measures can help answer these questions

- Key insight: with plausiblity measures we can isolate exactly what properties are needed for the axioms in System $\mathbf{P}$

Semantics for defaults using plausibility measures

We'll start by giving semantics for defaults using plausibility measures:

A *simple plausibility structure* is a tuple $M = (W, \mathrm{Pl}, \pi)$:
$M \models \varphi \rightarrow \psi$ iff $\mathrm{Pl}([\![\varphi]\!]_M) = \perp$ or $\mathrm{Pl}([\![\varphi \land \psi]\!]_M) > \mathrm{Pl}([\![\varphi \land \neg\psi]\!]_M)$

- Again, we're really saying that $\mathrm{Pl}([\![\psi]\!]_M \mid [\![\varphi]\!]_M) > \mathrm{Pl}([\![\psi]\!]_M \mid [\![\varphi]\!]_M)$
- (and if we can't condition because $\mathrm{Pl}([\![\varphi]\!]_M) = \perp$, then the default holds vacuously)
- Note that *true* $\rightarrow \varphi$ can be taken as an abbreviation for $B\varphi$.

Semantics for defaults using plausibility measures

We'll start by giving semantics for defaults using plausibility measures:

A *simple plausibility structure* is a tuple $M = (W, \mathrm{Pl}, \pi)$:
$M \models \varphi \to \psi$ iff $\mathrm{Pl}([\![\varphi]\!]_M) = \bot$ or $\mathrm{Pl}([\![\varphi \land \psi]\!]_M) > \mathrm{Pl}([\![\varphi \land \neg\psi]\!]_M)$

- Again, we're really saying that $\mathrm{Pl}([\![\psi]\!]_M \mid [\![\varphi]\!]_M) > \mathrm{Pl}([\![\psi]\!]_M \mid [\![\varphi]\!]_M)$
- (and if we can't condition because $\mathrm{Pl}([\![\varphi]\!]_M) = \bot$, then the default holds vacuously)
- Note that *true* $\to \varphi$ can be taken as an abbreviation for $B\varphi$.

With this semantics, it's easy to see the LLE, REF, and RW hold for all plausibility measures. The remaining proerties (AND, OR, and CM) don't hold without extra assumptions.

- E.g. the AND rule doesn't hold for probability
- Recall that Pl must satisfy Pl4 to get the AND rule
  - We showed that Pl4″ (where we intersected with V) was equivalent to Pl4 so as to get this "conditional" AND rule)
- What about OR and CM?

You might think that we also need to reverse engineer conditions on Pl to get CM and OR to hold. But ...

- PI4 also guarantees CM! (homework)

You might think that we also need to reverse engineer conditions on Pl to get CM and OR to hold. But ...

- Pl4 also guarantees CM! (homework)
- Another surprise: Pl4 gives us OR in the case that $\mathrm{Pl}([\![\varphi_1]\!]_M) \neq \bot$ or $\mathrm{Pl}([\![\varphi_2]\!]_M) \neq \bot$ (also homework)

The OR rule may not hold if $\mathrm{Pl}([\![\varphi_1]\!]_M) = \mathrm{Pl}([\![\varphi_2]\!]_M) = \bot$.

- In that case, $\varphi_1 \to \psi$ and $\varphi_2 \to \psi$ both hold vacuously.
- But why should $(\varphi_1 \vee \varphi_2) \to \psi$ hold?
  - Note that it would hold (vacuously) if $\mathrm{Pl}([\![\varphi_1 \vee \varphi_2]\!]_M) = \bot$.
- So we assume it does!

  Pl5. If $\mathrm{Pl}(A) = \mathrm{Pl}(B) = \bot$, then $\mathrm{Pl}(A \cup B) = \bot$.
    - Pl5 doesn't hold for lower probability or belief functions.
- Pl4 and Pl5 are *exactly* what we need to get AND, OR, and CM.

A plausibility measure is *qualitative* if it satisfies Pl4 and Pl5.

You might think that we also need to reverse engineer conditions on Pl to get CM and OR to hold. But . . .

- Pl4 also guarantees CM! (homework)
- Another surprise: Pl4 gives us OR in the case that $\mathrm{Pl}([\![\varphi_1]\!]_M) \neq \bot$ or $\mathrm{Pl}([\![\varphi_2]\!]_M) \neq \bot$ (also homework)

The OR rule may not hold if $\mathrm{Pl}([\![\varphi_1]\!]_M) = \mathrm{Pl}([\![\varphi_2]\!]_M) = \bot$.

- In that case, $\varphi_1 \to \psi$ and $\varphi_2 \to \psi$ both hold vacuously.
- But why should $(\varphi_1 \vee \varphi_2) \to \psi$ hold?
  - Note that it would hold (vacuously) if $\mathrm{Pl}([\![\varphi_1 \vee \varphi_2]\!]_M) = \bot$.
- So we assume it does!

  Pl5. If $\mathrm{Pl}(A) = \mathrm{Pl}(B) = \bot$, then $\mathrm{Pl}(A \cup B) = \bot$.
    - Pl5 doesn't hold for lower probability or belief functions.
- Pl4 and Pl5 are *exactly* what we need to get AND, OR, and CM.

A plausibility measure is *qualitative* if it satisfies Pl4 and Pl5.

Theorem: The KLM properties are sound and complete for simple *qualitative* plausibility structures.

Why we get the KLM properties

- We've already observed that possibility measures and ranking functions satisfy PI4.
- They trivially satisfy PI5
- Thus, they are qualitative plausibility measures
- Conclusion: the KLM properties will be sound for possibility measures and ranking functions

What about PS structures?

Why we get the KLM properties

- We've already observed that possibility measures and ranking functions satisfy PI4.
- They trivially satisfy PI5
- Thus, they are qualitative plausibility measures
- Conclusion: the KLM properties will be sound for possibility measures and ranking functions

What about PS structures?

- A probability sequence can also be identified with a (single) qualitative plausibility measure:
  - if $\vec{\Pr} = (\Pr_1, \Pr_2, \ldots)$, define $\mathrm{Pl}_{\vec{\Pr}}$ by taking $\mathrm{Pl}_{\vec{\Pr}}(A) \leq \mathrm{Pl}_{\vec{\Pr}}(B)$ iff $\lim_{n\to\infty} \Pr_n(B \mid A \cup B) = 1$.
    - B is "almost all" of $A \cup B$.

Why we get the KLM properties

- We've already observed that possibility measures and ranking functions satisfy Pl4.
- They trivially satisfy Pl5
- Thus, they are qualitative plausibility measures
- Conclusion: the KLM properties will be sound for possibility measures and ranking functions

What about PS structures?

- A probability sequence can also be identified with a (single) qualitative plausibility measure:
  - if $\vec{\Pr} = (\Pr_1, \Pr_2, \ldots)$, define $\mathrm{Pl}_{\vec{\Pr}}$ by taking $\mathrm{Pl}_{\vec{\Pr}}(A) \leq \mathrm{Pl}_{\vec{\Pr}}(B)$ iff $\lim_{n\to\infty} \Pr_n(B \mid A \cup B) = 1$.
    - B is "almost all" of $A \cup B$.

Theorem: (a) $\mathrm{Pl}_{\vec{\Pr}}$ is a qualitative plausibility measure.
(b) $(W, \mathrm{Pl}_{\vec{\Pr}}, \pi) \models \varphi \to \psi$ iff $(W, \vec{\Pr}, \pi) \models \varphi \to \psi$.

- As far as default statement go, $\vec{\Pr}$ and $\mathrm{Pl}_{\vec{\Pr}}$ are saying "the same thing".
- That's why we get the KLM properties for PS structures

What about preferential structures?

- A preference order $\succeq$ can also be identified with a qualitative plausibility measure:
  - Given a preference order $\succeq$ on worlds, define $\mathrm{Pl}_{\succeq}$ by taking $\mathrm{Pl}_{\succeq}(A) \leq \mathrm{Pl}_{\succeq}(B)$ if for all $w \in A$, there exists $w' \in B$ such that $w' \succeq w$.

What about preferential structures?

- A preference order $\succeq$ can also be identified with a qualitative plausibility measure:
  - Given a preference order $\succeq$ on worlds, define $\mathrm{Pl}_{\succeq}$ by taking $\mathrm{Pl}_{\succeq}(A) \leq \mathrm{Pl}_{\succeq}(B)$ if for all $w \in A$, there exists $w' \in B$ such that $w' \succeq w$.

Theorem: (a) $\mathrm{Pl}_{\succeq}$ is a qualitative plausibility measure.
(b) $(W, \mathrm{Pl}_{\succeq}, \pi) \models \varphi \rightarrow \psi$ iff $(W, \succeq, \pi) \models \varphi \rightarrow \psi$.

- As far as default statement go, $\vec{\mathrm{Pr}}$ and $\mathrm{Pl}_{\vec{\mathrm{Pr}}}$ are saying "the same thing".
- That's why we get the KLM properties for preferential structures.

Why is it so hard to beyond KLM

Let $\mathcal{M}^{qual}$ denote the set of all qualitative plausibility structures.

As long as $\mathcal{M} \subseteq \mathcal{M}^{qual}$, the KLM properties are sound in $\mathcal{M}$. If we take "too small" a subset of $\mathcal{M}^{qual}$, we may get extra properties. When is a set of plausibility structures not "too small"?

Why is it so hard to beyond KLM

Let $\mathcal{M}^{qual}$ denote the set of all qualitative plausibility structures.

As long as $\mathcal{M} \subseteq \mathcal{M}^{qual}$, the KLM properties are sound in $\mathcal{M}$. If we take "too small" a subset of $\mathcal{M}^{qual}$, we may get extra properties. When is a set of plausibility structures not "too small"?

A set $\mathcal{M}$ is *rich* if for every sequence of mutually exclusive formulas $\varphi_1, \ldots, \varphi_n$, there is a plausibility structure $M = (W, \mathrm{Pl}, \pi) \in \mathcal{M}$ such that

$$\mathrm{Pl}([\![\varphi_1]\!]_M) > \cdots > \mathrm{Pl}([\![\varphi_n]\!]_M) = \mathrm{Pl}(\emptyset).$$

Theorem: A set $\mathcal{M}$ of qualitative plausibility structures is rich iff the KLM properties provide a sound and complete axiomatization of entailment w.r.t. $\mathcal{M}$.

Going beyond KLM I

The KLM properties are useful, but we'd like more:

- From $\Sigma_1 = \{\textit{bird} \rightarrow \textit{fly}, \textit{penguin} \rightarrow \neg\textit{fly}\}$, we'd like to conclude $(red \wedge \textit{penguin}) \rightarrow \textit{fly}$. But this does not follow from KLM.
  - redness should be independent of penguins and flying!
  - It's easy to construct a qualitative plausibility structure where it's false
- From $\Sigma_1 \cup \{\textit{bird} \rightarrow \textit{have\_wings}\}$, we'd like to conclude $\textit{penguin} \rightarrow \textit{have\_wings}$, but again, this doesn't follow from KLM.
  - Since penguins are atypical in one respect, they might be atypical in others ...
- From $\Sigma_1 \cup \{yellow \rightarrow easy\text{-}to\text{-}see\}$, we'd like to conclude $(\textit{penguin} \wedge yellow) \rightarrow easy\text{-}to\text{-}see$, but we can't.
- From $\Sigma_1 \cup \{robin \rightarrow \textit{bird}\}$, we'd like to conclude $robin \rightarrow \textit{fly}$, but we can't.

The previous theorem shows that it will be hard to get a principled approach that allows us to draw conclusions over and above those of KLM. People tried and (mostly) failed.

Going beyond KLM II

One general idea for going beyond KLM:

- Instead of having a set of structures $\mathcal{M}$ and taking $\Sigma \models \varphi$ to mean "for all $M \in \mathcal{M}$, if $M \models \Sigma$ then $M \models \varphi$", we now take it to mean "for the 'best' structures $M \in \mathcal{M}$ such that $M \models \Sigma$, we have $M \models \varphi$.

We'll next consider two instantiations of this approach, one using ranking functions, and one using maximum entropy.

System Z [Goldszmidt and Pearl]

- Fix a set Φ of primitive propositions; let W_Φ consist of all truth assignments to the primitive propositions in Φ.
  - So we are identifying a world with a truth assignment
- Define a partial order $\succeq$ on ranking functions on W_Φ by defining $\kappa_1 \succeq \kappa_2$ if $\kappa_1(w) \leq \kappa_2(w)$ for all $w \in W_\Phi$.
  - κ_1 is preferred to κ_2 if every world is no more surprising according to κ_1 than it is according to κ_2.
- let $\pi_\Phi(w, p) = w(p)$
  - This makes sense because each world in W_Φ is a truth assignment.
- If $M_i = (W_\Phi, \kappa_i, \pi_\Phi)$ for $i = 1, 2$, the $M_1 \succeq M_2$ iff $\kappa_1 \succeq \kappa_2$.

System Z [Goldszmidt and Pearl]

- Fix a set Φ of primitive propositions; let W_Φ consist of all truth assignments to the primitive propositions in Φ.
  - So we are identifying a world with a truth assignment
- Define a partial order $\succeq$ on ranking functions on W_Φ by defining $\kappa_1 \succeq \kappa_2$ if $\kappa_1(w) \leq \kappa_2(w)$ for all $w \in W_\Phi$.
  - κ_1 is preferred to κ_2 if every world is no more surprising according to κ_1 than it is according to κ_2.
- let $\pi_\Phi(w, p) = w(p)$
  - This makes sense because each world in W_Φ is a truth assignment.
- If $M_i = (W_\Phi, \kappa_i, \pi_\Phi)$ for $i = 1, 2$, the $M_1 \succeq M_2$ iff $\kappa_1 \succeq \kappa_2$.

M is a "best" (most preferred) structure satisfying Σ if

- $M \models \Sigma$
- there is no M' such that $M' \models \Sigma$ and $M' \succ \mathcal{M}$.

System Z (cont'd)

Theorem: Given a finite set Σ of defaults that is satisfiable, there is a unique best structure M_Σ satisfying Σ (even though κ is a partial order on ranking functions).

System Z (cont'd)

Theorem: Given a finite set Σ of defaults that is satisfiable, there is a unique best structure M_Σ satisfying Σ (even though κ is a partial order on ranking functions).

Define $\Sigma \mid\!\approx^Z \varphi$ if either Σ is not satisfiable or $M_\Sigma \models \varphi$.

- That is, $\Sigma \mid\!\approx^Z \varphi$ if φ is true in the most preferred structure of all the structures satisfying Σ.
- This was called the "System Z" approach by Goldszmidt and Pearl.

Examples: Recall that $\Sigma_1 = \{\textit{bird} \rightarrow \textit{fly}, \textit{penguin} \rightarrow \neg \textit{fly}\}$. Then

- $\Sigma_1 \mid\!\approx^Z \textit{penguin} \wedge red \rightarrow \neg \textit{fly}$.

System Z (cont'd)

Theorem: Given a finite set Σ of defaults that is satisfiable, there is a unique best structure M_Σ satisfying Σ (even though κ is a partial order on ranking functions).

Define $\Sigma \mathrel{|\!\!\approx^Z} \varphi$ if either Σ is not satisfiable or $M_\Sigma \models \varphi$.

- That is, $\Sigma \mathrel{|\!\!\approx^Z} \varphi$ if φ is true in the most preferred structure of all the structures satisfying Σ.
- This was called the "System Z" approach by Goldszmidt and Pearl.

Examples: Recall that $\Sigma_1 = \{\mathit{bird} \to \mathit{fly}, \mathit{penguin} \to \neg \mathit{fly}\}$. Then

- $\Sigma_1 \mathrel{|\!\!\approx^Z} \mathit{penguin} \land red \to \neg \mathit{fly}$.

More generally: **Lemma:** Let $\Sigma_a = \{\varphi_1 \to \varphi_2, \varphi_2 \to \varphi_3\}$ and let $\Sigma_b = \Sigma_a \cup \{\varphi_1 \to \neg\varphi_3, \varphi_1 \to \varphi_4\}$.

(a) $\Sigma_a \mathrel{|\!\!\approx}^Z \varphi_1 \land \psi \to \varphi_3$ if $\varphi_1 \land \varphi_2 \land \varphi_3 \land \psi$ is satisfiable.

(b) $\Sigma_b \mathrel{|\!\!\approx}^Z \varphi_1 \land \psi \to \neg\varphi_3 \land \varphi_4$ if $\varphi_1 \land \varphi_2 \land \neg\varphi_3 \land \varphi_4 \land \psi$ is satisfiable.

   - if penguins have wings, then red penguins have wings but do not fly

System Z gives some of the results we want, but not all of them.

- Not only do penguins not inherit properties of birds such as flying (which, intuitively, they should not inherit), they also do not inherit properties of birds like having wings (which, intuitively, there is no reason for them not to inherit).
- We do not have $\Sigma_1 \cup \{\mathit{bird} \rightarrow \mathit{have\_wings}\} \not\approx^{Z} (\mathit{penguin} \wedge \mathit{bird}) \rightarrow \mathit{have\_wings}$ nor $\Sigma_1 \cup \{yellow \rightarrow easy\text{-}to\text{-}see\} \not\approx^{Z} (\mathit{penguin} \wedge yellow) \rightarrow easy\text{-}to\text{-}see$.

The next approach seems to have all the properties we want.

Preference via maximum entropy [Goldszmidt-Morris-Pearl]

This approach uses PS structures.

- Given a collection Σ of defaults, let Σ^k consist of the statements that result by replacing each default $\varphi \rightarrow \psi$ in Σ by the formula $\ell(\psi \mid \varphi) \geq 1 - 1/k$.
- Let $\mathcal{P}^k$ be the set of probability measures that satisfy Σ^k.
- If $\mathcal{P}^k \neq \emptyset$, let $\Pr_k^{me}$ be the probability measure of maximum entropy in $\mathcal{P}^k$.
  - There is a unique probability measure of maximum entropy in this set, since it is defined by linear inequalities.
- As long as $\mathcal{P}^k \neq \emptyset$ for all $k \geq 1$, this procedure gives a probability sequence $(\Pr_1^{me}, \Pr_2^{me}, \ldots)$.
- Let $M_\Sigma^{me} = (W_\Phi, (\Pr_1^{me}, \Pr_2^{me}, \ldots), \pi_\Phi)$.
- $\Sigma \mid\!\approx^{me} \varphi$ if either there is some k such that $\mathcal{P}^k = \emptyset$ (in which case $\mathcal{P}^{k'} = \emptyset$ for all $k' \geq k$) or $M_\Sigma^{me} \models \varphi$.

This approach seems to get all the desired properties:

- Intuition: the distribution that maximizes entropy makes this "as independent as possible", subject to the constraints.

Conditional Logic

Default logic just contains formulas of the form $\varphi \rightarrow \psi$

- No conjunctions, negations, implications, nested defaults, multiple agents, ...

Conditional logic generalizes to allow all of the above.

- It is completely straightforward to extend the earlier approaches (preferential semantics, PS structures, possibility measures, ranking functions, qualitative plausibility measures) to giving semantics to conditional logic.
- Whereas all the earlier approaches had the same sound and complete axiomatizations in default logic (system $\mathbf{P}$), with conditional logic, we can distinguish them.
  - $(\varphi \rightarrow_i \psi_1) \land \neg(\varphi \rightarrow_i \neg\psi_2) \Rightarrow ((\varphi \land \psi_2) \rightarrow_i \psi_1)$ is true if we have a totally ordered plausibility measure
    - true for possibility measures, ranking functions, total orders $\succeq$
    - not true if $\succeq$ is a partial order or for PS structures.

Conditional Logic

Default logic just contains formulas of the form $\varphi \to \psi$

- No conjunctions, negations, implications, nested defaults, multiple agents, ...

Conditional logic generalizes to allow all of the above.

- It is completely straightforward to extend the earlier approaches (preferential semantics, PS structures, possibility measures, ranking functions, qualitative plausibility measures) to giving semantics to conditional logic.
- Whereas all the earlier approaches had the same sound and complete axiomatizations in default logic (system $\mathbf{P}$), with conditional logic, we can distinguish them.
  - $(\varphi \to_i \psi_1) \land \neg(\varphi \to_i \neg\psi_2) \Rightarrow ((\varphi \land \psi_2) \to_i \psi_1)$ is true if we have a totally ordered plausibility measure
    - true for possibility measures, ranking functions, total orders $\succeq$
    - not true if $\succeq$ is a partial order or for PS structures.
  - $\neg(\mathit{true} \to_i \mathit{false})$
    - This basically says that $\mathrm{Pl}(W) \neq \bot$
    - Holds for possibility measures, ranking functions, PS structures, but not for preferential orders.

Reasoning about counterfactuls

We can use conditional logic for counterfactual reasoning:

- Now we interpret $\varphi \to \psi$ as "if φ were the case, then ψ would be true"
- If my match were dry (which it's not) then it would light.

Counterfactual reasoning is critical in game theory:

- What would happen if my opponent were to move left ...

It is also critical in reasoning about causality:

- A is a cause of B if it is the case that if A hadn't happened, then B wouldn't have happend
- My hitting you is a a cause of you falling over since if I hadn't hit you, you wouldn't have fallen

Semantics for Counterfactuals

Intuition for $\varphi \to \psi$ (due to David Lewis/Robert Stalnaker):

- $(M, w) \models \varphi \to \psi$ if, at the closest world to w/most plausible world from the perspective of w where φ is true, ψ is also true
  - At the world most like this one where the match is dry, it would light.
  - This is not a world where there is no oxygen!
- We capture this by using a preference order/plausibility measure that is world-dependent
- We require that the closest world to w is w itself.
  - This gives us the property: $\varphi \Rightarrow ((\varphi \to \psi) \Leftrightarrow \psi)$

In the text, there's lots of discussion of sound and complete axiomatizations