Lecture 8 – Circ. and Pref. Entailment (II)

1. Nonmonotonic Reasoning
2. Abnormality Theories
3. Model Preference Logics
4. Circumscription
   (a) Parallel Predicate Circumscription
   (b) Prioritized Predicate Circumscription
   (c) Pointwise Circumscription
5. Model Preference Logics Revisited

Prioritized Predicate Circumscription

– Many reasoning tasks involve minimization of several Ab predicates. If minimizing one Ab predicate affects the extension of another, then conflicts may arise.

– The solution is to introduce a priority ordering on predicates where those with the highest priority are minimized at the expense of those with less priority.

Example: Inheritence Hierarchies

1. *Normally*, adults are employed.

2. *Normally*, high school dropouts are not employed.

3. All high school dropouts are adults.

4. Mary is a high school dropout.

– Intuitively, one might suppose that Mary is unemployed, given this information and application of a "rule of thumb" where

> **Values of properties about more specific concepts override those of less specific concepts.**

Example (con't)

$$
\begin{gathered}
\Gamma(\{Ab_1, Ab_2\}, \{Employed\}) = \\
\{\forall x. Adult(x) \wedge \neg Ab_1(x) \rightarrow Employed(x), \\
\{\forall x. Dropout(x) \wedge \neg Ab_2(x) \rightarrow \neg Employed(x), \\
\{\forall x. Dropout(x) \rightarrow Adult(x), \\
Dropout(Mary)
\end{gathered}
$$

– Applying parallel predicate circumscription where

- Ab_1 and Ab_2 are minimized,
- $Employed$ is varied, and
- $Dropout$ and $Adult$ are fixed

results in two sets of preferred models:

1. A set where Ab_1 is minimal and $Mary$ is employed.
2. A set where Ab_2 is minimal and $Mary$ is unemployed.

– Using the specificity criteria, we prefer (2), yet PP-circumscription is too weak to do this.

Example (con't)

– The problem becomes clearer by first manipulating formulas in our premise set to logically equivalent forms:

– Consider the two formulas

$$\forall x. Adult(x) \wedge \neg Ab_1(x) \rightarrow Employed(x), \tag{1}$$

$$\forall x. Dropout(x) \wedge \neg Ab_2(x) \rightarrow \neg Employed(x). \tag{2}$$

– The premise (4) can be rewritten as the equivalent

$$\forall x. Dropout(x) \wedge Employed(x) \rightarrow Ab_2(x). \tag{3}$$

– Premises (3) and (5) together give

$$\forall x. Adult(x) \wedge Dropout(x) \wedge \neg Ab_1(x) \rightarrow Ab_2(x). \tag{4}$$

which is equivalent to both

$$\forall x. Adult(x) \wedge Dropout(x) \wedge \neg Ab_2(x) \rightarrow Ab_1(x) \tag{5}$$

and

$$\forall x. Adult(x) \wedge Dropout(x) \rightarrow (Ab_1(x) \vee Ab_2(x)). \tag{6}$$

– From (8), it is clear that if both Ab predicates are minimized there will be a conflict. In some preferred models Dropouts that are adults will be Ab_1, and in others, Ab_2. The best we can do is to derive

$$\neg Ab_1(Mary) \vee \neg Ab_2(Mary).$$

Priority Orderings

– Clearly we would like to prefer minimizing Ab_2 at the expense of Ab_1.

– In terms of circumscription, the following policy captures the idea:

1. Minimize Ab_2; let Ab_1 and $Employed$ vary; keep $Dropout$ and $Adult$ fixed.

   Take the resulting preferred model classes and

2. Minimize Ab_1; let $Employed$ vary ;keep Ab_2, $Dropout$ and $Adult$ fixed.

Informal Semantic Characterization:

– Informally in terms of comparing extensions:

Given two models M and N, M is preferred to N if

3. $Value(Ab_2, M) \subseteq Value(Ab_2, N)$, or
   If $Value(Ab_2, M) = Value(Ab_2, N)$
   then
   $Value(Ab_1, M) \subseteq Value(Ab_1, N)$.

Informal Syntactic Characterization:

$$\overline{P}' \ll \overline{P} \equiv \{\{Ab'_2\}, \{Ab'_1\}\} \ll \{\{Ab_2\}, \{Ab_1\}\} \equiv$$

$$[Ab'_2 \leq Ab_2 \wedge (Ab'_2 \equiv Ab_2 \rightarrow Ab'_1 \leq Ab_1)] \rightarrow$$
$$[Ab_2 \leq Ab'_2 \wedge (Ab_2 \equiv Ab'_2 \rightarrow Ab_1 \leq Ab'_1)] \equiv$$

$$[(\forall x. Ab'_2(x) \rightarrow Ab_2(x)) \wedge$$
$$((\forall x. Ab'_2(x) \leftrightarrow Ab_2(x)) \rightarrow (\forall x. Ab'_1(x) \rightarrow Ab_1(x)))] \rightarrow$$
$$[(\forall x. Ab_2(x) \rightarrow Ab'_2(x)) \wedge ((\forall x. Ab_2(x) \leftrightarrow$$
$$Ab'_2(x)) \rightarrow (\forall x. Ab_1(x) \rightarrow Ab'_1(x)))]$$

Formal Syntactic Characterization

– Suppose, we would like to minimize a pair of predicate constants P_1 and P_2 in Γ. In parallel circumscription this is expressed as

$$\Gamma(\{P_1, P_2\}, \emptyset) \wedge \forall P_1', P_2' [\Gamma(\{P_1', P_2'\}, \emptyset) \wedge (P_1', P_2') \leq (P_1, P_2)] \rightarrow (P_1, P_2) \leq (P_1', P_2') \quad (7)$$

– For arbitrary pairs of similar predicate expressions (U_1, U_2) and (V_1, V_2),

$$(U_1, U_2) \leq (V_1, V_2) \text{ iff } (U_1 \leq V_1) \wedge (U_2 \leq V_2).$$

– To minimize P_1 at a higher priority than P_2, use the circumscription axiom (9), but with a different ordering relation.

– Instead of $\leq$, we will replace it by the ordering $\lll$, where

$$(U_1, U_2) \lll (V_1, V_2) \text{ iff } (U_1 \leq V_1) \wedge ((U_1 \equiv V_1) \rightarrow (U_2 \leq V_2)).$$

– The relation can be extended to similar tuples of predicate expressions

$$(U_1, \ldots, U_n) \lll (V_1, \ldots, V_n) \text{ iff}$$
$$U_1 \lll V_1 \wedge \bigwedge_{i=2}^{n} ((\bigwedge_{j=1}^{i-1} U_j \equiv V_j) \rightarrow U_i \leq V_i)$$

and to similar lists of tuples of predicate expressions

$$(\overline{U_1}, \ldots, \overline{U_n}) \underline{\ll} (\overline{V_1}, \ldots, \overline{V_n}) \text{ iff}$$
$$\overline{U_1} \underline{\ll} \overline{V_1} \wedge \bigwedge_{i=2}^{n} ((\bigwedge_{j=1}^{i-1} \overline{U_j} \equiv \overline{V_j}) \rightarrow \overline{U_i} \leq \overline{V_i}).$$

We write $(\overline{U_1}, \ldots, \overline{U_n}) \ll (\overline{V_1}, \ldots, \overline{V_n})$ for

$$(\overline{U_1}, \ldots, \overline{U_n}) \underline{\ll} (\overline{V_1}, \ldots, \overline{V_n}) \wedge \neg (\overline{V_1}, \ldots, \overline{V_n}) \underline{\ll} (\overline{U_1}, \ldots, \overline{U_n}).$$

Classifying Predicates into Priority Classes

– Given: a tuple of predicate constants $\overline{P}$, to minimize.

– Divide $\overline{P}$ into disjoint sublists of predicate constants $\overline{P_1}, \ldots, \overline{P_n}$.

– Predicates in $\overline{P_1}$ are minimized at highest priority, Predicates in $\overline{P_2}$ are minimized at next highest priority, etc.

– This is expressed as

$$\overline{P^1} > \overline{P^2} >, \ldots > \overline{P^n}.$$

Definition 1 (Prioritized Circumscription)
Let $\overline{P}$ be a tuple of distinct predicate constants, $\overline{S}$ be a tuple of distinct predicate constants and/or function constants disjoint from $\overline{P}$, and $\Gamma(\overline{P}, \overline{S})$, a theory. The prioritized circumscription *of $\overline{P}$ in $\Gamma(\overline{P}, \overline{S})$ with variable $\overline{S}$ with respect to priorities $\overline{P^1} > \overline{P^2} >, \ldots > \overline{P^n}$, written $Circ_{pc}(\Gamma; \overline{P^1} > \overline{P^2} >, \ldots > \overline{P^n}; \overline{S})$, is the sentence*

$$\Gamma(\overline{P}, \overline{S}) \wedge \forall \overline{P}', \overline{S}'. \neg[\Gamma(\overline{P}, \overline{S}) \wedge \overline{P}' \ll \overline{P}]$$

Here, $\overline{P}' \ll \overline{P}$ is
$(\overline{P^{1\prime}} > \overline{P^{2\prime}} >, \ldots > \overline{P^{n\prime}}) \ll (\overline{P^1} > \overline{P^2} >, \ldots > \overline{P^n})$.

Formal Semantic Characterization

Definition 2 (Minimal Models)
Let $\overline{P^1}, \overline{P^n}$ be disjoint sublists of all predicates constants from $\overline{P}$, $\overline{S}$ be a tuple of distinct predicate constants and/or function constants disjoint from $\overline{P}$, and $\Gamma(\overline{P}, \overline{S})$, a theory. Suppose that M and N are models of Γ. M is said to be a $(\overline{P^1} >, \ldots, > \overline{P^n}); \overline{S})$-preferred model of N, written $M \underline{\ll}^{(\overline{P^1} >, \ldots, > \overline{P^n}; \overline{S})} N$, iff

1. $D_M = D_N$.

2. $Value(U, M) = Value(U, N)$ *for all constants* U *not in* $\overline{P} \cup \overline{S}$.

3. *For every* $j = 1, \ldots, n$, *if*

   $Value(P_l, M) = Value(P_l, N)$, *for every* $P_l \in (\overline{P}^1, \ldots, \overline{P}^{j-1})$,

   then

   $Value(P_l, M) \subseteq Value(P_l, N)$, *for evry* $P_l \in \overline{P}^j$.

We write $M \ll^{(\overline{P^1} >, \ldots, > \overline{P^n}; \overline{S})} N$ if $M \underline{\ll}^{(\overline{P^1} >, \ldots, > \overline{P^n}; \overline{S})} N$, but not $N \underline{\ll}^{(\overline{P^1} >, \ldots, > \overline{P^n}; \overline{S})} M$.

A model of Γ is $(\overline{P^1} >, \ldots, > \overline{P^n}; \overline{S})$-minimal iff Γ has no model N such that $N \ll^{(\overline{P^1} >, \ldots, > \overline{P^n}; \overline{S})} M$. The set of all $(\overline{P^1} >, \ldots, > \overline{P^n}; \overline{S})$-minimal models of Γ will be denoted by $MOD^{(\overline{P^1} >, \ldots, > \overline{P^n}; \overline{S})}(\Gamma)$.

Pointwise Circumscription

There are two alternative ways to interpret an n-ary predicate constant in a model M with domain D_M:

1. As a subset of ${D_M}^n$.

2. As a function D_M into $\{True, False\}$.

(1) leads to the usual specification of circumscription: to circumscribe a predicate P in a theory Γ is to assume the extension of P can not be made smaller (in the sense of set inclusion) without violating Γ.

(2) offers another possibility. We can think of amaking a predicate "smaller" at point $x \in {D_M}^n$ as changing its value at this point from $True$ to $False$. The predicate is then viewed as minimal wrt Γ if it cannot be made smaller (in this sense) witout violating Γ.

Basic Case

We begin with minimizing one predicate P in $\Gamma(P, \bar{S})$ with predicate and/or function constants $\bar{S}$ allowed to vary. We want to express the following:

> It is impossible to change the value of P at any one point from $True$ to $False$ without violating Γ, even if the values of elements of $\bar{S}$ are arbitrarily changed.

$$A = \exists \bar{x} \exists \bar{\Psi}.[P(\bar{x}) \wedge \Gamma(\lambda \bar{y}(P(\bar{y}) \wedge \bar{y} \neq \bar{x}), \bar{\Psi})]. \tag{8}$$

A states: There is a point,$\bar{x}$, and a tuple of predicate and/or function symbols,$\bar{\Psi}$, such that the value of P at $\bar{x}$ can be changed from $True$ to $False$ without violating Γ, provided that the elements of $\bar{S}$ are replaced by those of $\bar{\Psi}$.

$$A = \exists\bar{x}\exists\bar{\Psi}.[P(\bar{x})\wedge\Gamma(\underbrace{\lambda\bar{y}(P(\bar{y})\wedge\bar{y} \neq \bar{x})}, \bar{\Psi})]. \tag{9}$$

$\lambda\bar{y}(P(\bar{y})\wedge\bar{y} \neq \bar{x})$ – The set of points, excluding $\bar{x}$, at which P is true.

Definition 3 (Pointwise Circumscription) *Let* $\Gamma(P, \bar{S})$ *be a theory, where* P *is a predicate constant and* $\bar{S}$ *is a tuple of predicate and/or function constants not containing* P. *The* pointwise circumscription *of* $P \in \Gamma(P, \bar{S})$ *with variable* $\bar{S}$, *denoted by* $CIRC_{PW}(\Gamma; P; \bar{S})$, *is the sentence*

$$\begin{aligned}&\Gamma(P, \bar{S})\\ &\quad\forall\bar{x}\forall\bar{\Psi}.\neg[P(\bar{x})\wedge\Gamma(\lambda\bar{y}(P(\bar{y})\wedge\bar{y} \neq \bar{x}), \bar{\Psi})].\end{aligned} \tag{10}$$

The set of theorems derivable by pointwise circumscription of P in Γ with variable $\bar{S}$ is the set of formulas entailed by $CIRC_{PW}(\Gamma; P; \bar{S})$.

Example1

Let $\Gamma = \{P(a)\}$.

$$\Gamma(\lambda y(P(y) \wedge y \neq x)) \Leftrightarrow P(a) \wedge a \neq x.$$

$$CIRC_{PW}(\Gamma; P; ()) = \Gamma \wedge \forall x \neg [P(x) \wedge P(a) \wedge a \neq x]$$

It is easily seen that

$$CIRC_{PW}(\Gamma; P; ()) \Leftrightarrow P(a) \wedge \forall x . P(x) \supset x = a.$$

Example2

Let $\Gamma = \{P(a) \equiv P(b)\}$.

It can be shown that:

1.
$$CIRC_{PW}(\Gamma; P; ()) \Leftrightarrow \Gamma \wedge \forall x.P(x) \supset [(P(a) \wedge x \neq a \wedge x = b) \vee (P(b) \wedge x = a \wedge x \neq b)].$$

2.
$$CIRC_{SO}(\Gamma; P; ()) \Leftrightarrow \Gamma \wedge \forall x.\neg P(x) \qquad (11)$$

Note that P interpreted as universally false is a model for both (1) and (2) but that (1) has additional models, namely those in which P is true at precisely two points: a and b.

Generalized Pointwise Circumscription

Definition 4 (Gen. Pointwise Circumscription) *Let P be a predicate constant of arity n, $\bar{S} = \{S_1, \ldots, S_m\}$ be a tuple of distinct function and/or predicate constants different from $\bar{P}$, and let $\Gamma(P, \bar{S})$ be a theory. Let V be a $2 \times n$-ary predicate expression and suppose that $\bar{U} = \{U_1, \ldots, U_m\}$, where each $U_i (1 \leq i \leq m)$ is a predicate expression whose arity is the arity of S_i plus n. It is assumed that $V, U_1, \ldots, U_m$ have no parameters and contain none of $P, S_1, \ldots, S_m$. The* pointwise circumscription of P in $\Gamma(P, \bar{S})$ with P allowed to vary on $V(\bar{x})$ and S_i allowed to vary on $U_i(\bar{x})(1 \leq i \leq m)$, *written*

$Circ_{PW}(\Gamma; P/V; \bar{S}/\bar{U})$, *is the sentence*

$$\begin{aligned} &\Gamma(P, \bar{S}) \wedge \\ &\quad \forall \bar{x}, \Phi, \Psi. \neg [P(\bar{x}) \wedge \neg \Phi(\bar{x}) \wedge EQ_{V(\bar{x})}(P, \Phi) \\ &\qquad \wedge \textstyle\bigwedge_{i=1}^{m} [EQ_{U_i(\bar{x})}(S_i, \Psi_i)] \wedge \Gamma(\Phi, \Psi)] \end{aligned} \tag{12}$$

where $EQ_V(U, W) = \forall \bar{x}. \neg V(\bar{x}) \rightarrow [U(\bar{x}) \equiv W(\bar{x})]$ and

where Φ and $\bar{\Psi}$ correspond to P and $\bar{S}$, respectively.

– The intuition behind the use of $EQ_V(U,W)$ is that only points that have property V are allowed to vary. Consequently, when minimizing a point in a predicate P. it is possible to vary other points in P, or points in additional predicates S.

– Just as one can control the variation of predicate extensions in second-order circumscription, here one can control variation of points *within* a predicate extension.

Pros and Cons of Circumscription

1. Pros

   (a) Entirely embedded in the classical framework. first- or second- order logic.

   (b) After constructing the circumscription axiom, all deduction proceeds in monotonic logic. first-order logic plus nonmonotonicity reduces to classical second-order logic.

   (c) Possibilities of computing with fragments of theories to use logic programming etc.

2. Cons

   (a) How do we select the circumscriptive policy?

   (b) How do we choose the proper substitutions, or compute them automatically?

Model Preference Logics Revisited

Yoav Shoham provides an elegant generalization of the idea of prefered models implicit in McCarthy's work and made more explicit in Lifschitz's work, that applies to logics other than FOPC. The idea is as follows:

1. Start with any standard logic, such as FOPC, Modal S4, or Intuitionistic logic, for example.

2. Define a partial order on the interpretations of the logic. The partial order might reflect a policy of preference based on the particular phenomona being modeled, such as temporal phenomona.

3. Shift the emphasis away from the particular way of defining the ordering, such as circumscription, to the semantics. Work with the models directly.

In a sense, this is the approach that will be used in the rest of the course.