Lecture 7 – Circumscription and Pref. Entailment (I)

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

Nonmonotonic Reasoning

Monotonicity:

For any sets of premises S and S':

$$S \subseteq S' \text{ implies } \{A : S \vdash A\} \subseteq \{A : S' \vdash A\}.$$

(F1) On Saturday evenings John usually visits his club.

(F2) It is Saturday evening.

(C) John is at the club.

(F3) Yesterday John had a car accident.

Definition 1 *By* nonmonotonic reasoning *we understand the drawing of conclusions which may be invalidated in the light of new information. A logical system is called* nonmonotonic *iff its provability relation violates the property of nonmonotonicity.*

(F1) On Saturday evenings John usually visits his club.

(F2) It is Saturday evening.

(C) John is at the club.

Is (C) based on the rule:

(R) From (F1) and (F2) infer (C) ?

or

(R) From (F1) and (F2) in the absence of evidence to suspect otherwise, infer (C).

Nonmonotonic Reasoning and Reasoning about Action

Frame Problem: The problem of representing which properties persist and which properties change when actions are performed.

Corresponding Nonmonotonic Rule: "All aspects of the world remain invariant except for those explicitly changed."

Qualification Problem: How to represent all qualifications which should be satisfied to assure the success of the performed action.

Corresponding Nonmonotonic Rule: "In the absence of evidence to the contrary, assume that an action succeeds."

Abnormality Theories

– Introduced by McCarthy for doing default reasoning with Circumscription.

Idea:

– Associate a number of *aspects* with every object. Generally, an object is *normal* relative to most aspects, *abnormal* to some.

– Language signature includes one or more Ab_x predicate constants which will be *minimized*.

– Ab takes three arguments, an aspect of the entity involved, the event or entity the aspect is abnormal to, and finally a situation.

$$\forall x, l, s. Ab(aspect1(x), Move(x, l), s)$$

$$\forall x, e, s, l. Holds(at(x, l), s) \wedge \neg Ab(aspect1(x), e, s) \rightarrow Holds(at(x, l), result(e, s))$$

– Minimizing the extension of Ab in effect achieves a maximization of normal aspects. Why is this? Well, any aspect not shown to be abnormal can be assumed not to be, $\neg Ab(aspect_z(x), e, s)$.

– Result: Maximize persistence, Maximize action performance.

Example: Blocks World – Paint and Move

Laws of Motion

$$\forall x, l, s. \neg Ab(aspect3(x), l, s) \rightarrow Holds(at(x, l), result(move(x, l), s))$$

$$\forall x, c, s. \neg Ab(aspect4(x), c, s) \rightarrow Holds(color(x, c), result(paint(x, c), s))$$

Frame Axioms

$$\forall x, e, s, l. Holds(at(x, l), s) \wedge \neg Ab(aspect1(x), e, s) \rightarrow Holds(at(x, l), result(e, s))$$

$$\forall x, e, s. Holds(color(x, c), s) \wedge \neg Ab(aspect2(x), e, s) \rightarrow Holds(color(x, c), result(e, s))$$

$$\forall x, l, s. Ab(aspect1(x), move(x, l), s)$$

$$\forall x, c, s. Ab(aspect2(x), paint(x, c), s)$$

Qualification Axioms

$$\forall x,l,s.\neg Holds(clear(top(x)),s)\vee\neg Holds(clear(l),s)$$
$$\vee Holds(l=top(x),s)\rightarrow Ab(aspect3(x),move(x,l),s)$$

State Constraints

$$\forall l,s.Holds(clear(l),s)\equiv\ \neg\exists x.\neg Holds(trivial(x),s)\wedge Holds(at(x,l),s)$$

$$\forall x,s.\neg Ab(aspect7(x),s)\rightarrow\neg Holds(trivial(x),s)$$

– Conjecturing $\neg Ab(aspect_z(x),e,s)$ statements **not** the same as deriving them classically.

– Adding additional axioms to the KB can affect their status as consequences, as well as those consequences inferred using the conjectures.

Model Preference Logics

– Based on idea of specifying criteria for choosing a preferred set of models among the models for a theory or premise set.

– For abnormality theories: Given a premise set Γ, prefer those models of Γ whose extensions for Ab are set minimal.

– Smaller set of models generally implies more consequences. Extra consequences may be viewed as conjectures.

Example:

A room contains three blocks. Our language contains three constant symbols $b1$, $b2$, and $b3$, and several predicate symbols including $Block$. Γ consists of

$$Block(b1), Block(b2), Block(b3)$$
$$b1 \neq b2 \wedge b1 \neq b3 \wedge b2 \neq b3$$

– Want to derive the conjecture

"The only blocks in the room are those I am aware of"

$$\forall x. Block(x) \rightarrow x = b1 \vee x = b2 \vee x = b3$$

For finite case, we could just add the axiom, but this will not work in general. Instead, we could prefer those models that minimize the extension of $Block$, written $M(\Gamma)_{\prec}$.

$$M(\Gamma)_{\prec} \models \forall x. Block(x) \rightarrow x = b1 \vee x = b2 \vee x = b3$$

Desired:

- Principled means of specifying criteria for preferring one model over another.

- Syntactic characterization of preferential entailment.

Circumscription

– "Rule of Conjecture" that can be used for jumping to certain conclusions:

> **"The objects that can be shown to have a certain property P by reasoning from certain facts Γ are all the objects that satisfy P ."**

– More technically: The tuples $\langle x, y, \ldots, z \rangle$ that can be shown to satisfy a relation $P(x, y, \ldots, z)$ are all the tuples satisfying this relation.

– Thus the name *Circumscription* and technique of *circumscribing*.

– Many types of circumscription:

- Parallel Predicate Circumscription
- Prioritized Circumscription
- Pointwise Circumscription

Parallel Predicate Circumscription

– Let $\mathcal{L}$ be our logical language, P a set of predicate symbols to minimize, and Z another set of symbols(predicate and/or function symbols), which can vary across models.

– $\Gamma(P, Z)$ denotes a first-order sentence containing symbols P and Z.

Semantic Characterization

Let M_1 and M_2 be two models of $\Gamma(P, Z)$, where D_{M_1} and D_{M_2} are the domains of discourse for M_1 and M_2, respectively.

Definition 2 (Minimal Models)
We say that the model M_1 is P-minimal to M_2, written as $M_1 \preceq_{P;Z} M_2$, iff

1. $D_{M_1} = D_{M_2}$.
2. $Value(U, M_1) = Value(U, M_2)$ *for all* U *not in* P *or* Z.
3. $Value(U, M_1) \subseteq Value(U, M_2)$ *for all* $P_i \in P$.

The strict ordering, $\prec_{P;Z}$, on models is defined as

$$M_1 \preceq_{P;Z} M_2 \text{ and not } M_2 \preceq_{P;Z} M1.$$

The P-preferred models are those minimal relative to $\prec_{P;Z}$.

Technical Minimization Criteria

– Note that in order to use this definition that predicates and functions in our axiomatization must be classified into three sets:

1. Those that we want to minimize, P (only predicates),

2. those that we want to vary, Z (predicates and functions),

3. and those that we want to keep fixed.

Each combination represents a different minimization policy.

Example

– $\Lambda = \langle\{Tweety, Fred\}, \{\}, \{Bird, Ab, Flies\}\rangle$ is L's signature.

$$\begin{gathered}\Gamma(Ab, Flies) = \\ \{\forall x.Bird(x) \wedge \neg Ab(x) \rightarrow Flies(x), \\ Bird(Tweety) \\ \forall x.x = Tweety \vee x = Fred \\ Tweety \neq Fred\}\end{gathered} \tag{1}$$

Minimization Policy: (1) Minimize: Ab, (2) Vary: $Flies$, (3) Fix: $Bird$.

– $Tweety \neq Fred$ means that any model for $\Gamma(Ab, Flies)$ will have at least two entities in the domain of discourse.

– $\forall x.x = Tweety \vee x = Fred$ means that there are at most two entities in any model for $\Gamma(Ab, Flies)$.

– The following table describes the models for $\Gamma(Ab, Flies)$. For any two objects, a table isomorphic to the table shown is also a model.

	$Tweety_I$	$Fred_I$	$Bird_I$	Ab_I	$Flies_I$
1	$\star$	$\bullet$	$\star$	$\star$	$\star$
2	$\star$	$\bullet$	$\star$	$\star$	$\star, \bullet$
3	$\star$	$\bullet$	$\star$	$\star$	$-$
4	$\star$	$\bullet$	$\star$	$-$	$\star$
5	$\star$	$\bullet$	$\star$	$-$	$\star, \bullet$
6	$\star$	$\bullet$	$\star$	$\bullet$	$\star$
7	$\star$	$\bullet$	$\star$	$\bullet$	$\star, \bullet$
8	$\star$	$\bullet$	$\star, \bullet$	$\star$	$\bullet$
9	$\star$	$\bullet$	$\star, \bullet$	$\star$	$\star, \bullet$
10	$\star$	$\bullet$	$\star, \bullet$	$\bullet$	$\star$
11	$\star$	$\bullet$	$\star, \bullet$	$\bullet$	$\star, \bullet$
12	$\star$	$\bullet$	$\star, \bullet$	$\star, \bullet$	$\star$
13	$\star$	$\bullet$	$\star, \bullet$	$\star, \bullet$	$\bullet$
14	$\star$	$\bullet$	$\star, \bullet$	$\star, \bullet$	$\star, \bullet$
15	$\star$	$\bullet$	$\star, \bullet$	$\star, \bullet$	$-$
16	$\star$	$\bullet$	$\star, \bullet$	$-$	$\star, \bullet$

– Ab-Preferred models of $\Gamma(Ab, Flies)$ are 4,5, and 16.

– Let $Mod_{\prec}(\Gamma(Ab, Flies))$ denote the Ab-preferred models of $\Gamma(Ab, Flies)$. The following formulas are preferentially entailed:

$$Mod_{\prec}(\Gamma(Ab, Flies)) \models \neg Ab(Tweety), Flies(Tweety).$$

– From the example, one can see why we must vary certain predicates. If our minimization policy minimized Ab without varying $Flies$ then the Ab-preferred models would be 4,5,8,10,15, and 16.

– In this case, we could not conjecture any new positive or negative atoms. The original form of circumscription did not allow varying predicates.

– The nonmonotonic behavior of the preferred models approach may be seen by adding $Ab(Tweety)$ to $\Gamma(Ab, Flies)$. The preferred models are then 1,2,3,8, and 9. It is clear that we can no longer entail $Flies(Tweety)$.

Syntactic Characterization

– Given a set Γ of wff's and a minimization policy, the circumscription formalism adds a 1st-order schema or 2nd-order formula to Γ where the extended set of axioms is sound and complete (generally) relative to the preferred models.

– Our task is to determine what the circumscription axiom looks like. First, some preliminaries:

Comparing Extensions:

Given a model M and two predicates U and V of the same arity, we denote their extensions in M by U_M and V_M, respectively.

- If $U_M \subseteq V_M$ then

  $Value(\forall x.U(x) \rightarrow V(x), M) = 1$.

- If $U_M \subset V_M$ then

  $Value(\forall x.U(x) \rightarrow V(x) \wedge \neg \forall x.V(x) \rightarrow U(x), M) = 1$.

- If $U_M \equiv V_M$ then

  $Value(\forall x.U(x) \rightarrow V(x) \wedge \forall x.V(x) \rightarrow U(x), M) = 1$.

Let U and V be predicates of the same arity. We will use the following abbreviations:

- $U \leq V$ stands for $\forall x.U(x) \rightarrow V(x)$.
- $U < V$ stands for $\forall x.U(x) \rightarrow V(x) \wedge \neg \forall x.V(x) \rightarrow U(x)$.
- $U = V$ stands for $\forall x.U(x) \rightarrow V(x) \wedge \forall x.V(x) \rightarrow U(x)$.

– $U \leq V$ expresses the fact that the extension of U is a subset of the extension of V.

– $U < V$ expresses the fact that the extension of U is a strict subset of the extension of V.

– $U = V$ expresses the fact that the extension of U is the same as the extension of V.

If $\overline{U} = \{U_1, \ldots, U_i\}$ and $\overline{V} = \{V_1, \ldots, V_i\}$ are similar tuples of predicate expressions, then $\overline{U} \leq \overline{V}$ is used as an abbreviation for

$$\bigwedge_{i=1}^{n} [U_i \leq V_i]$$

.

– $\overline{U} < \overline{V}$ is an abbreviation for $\overline{U} \leq \overline{V}$ and $\neg \overline{V} \leq \overline{U}$
– $\overline{U} \equiv \overline{V}$ is an abbreviation for $\overline{U} \leq \overline{V}$ and $\overline{V} \leq \overline{U}$.

– For example, if $\overline{U} = \{P, Q\}$ and $\overline{V} = \{P', Q'\}$ then $\overline{U} \leq \overline{V}$ is equivalent to $P \leq P' \wedge Q \leq Q'$.

– Given Γ, suppose we would like to minimize P and let Z vary.

– We denote the axiom set as $\Gamma(P, Z)$ and the variable circumscription of P in Γ with predicate variables Z as $Circ(\Gamma(P, Z); P; Z)$.

–The idea is that the class of formulas

$$\{\alpha \mid Circ(\Gamma(P, Z); P; Z) \models \alpha\} \tag{2}$$

is the class of formulas minimally entailed by Γ. In other words, those formulas true in all P-preferred models of Γ.

What $Circ(\Gamma(P, Z); P; Z)$?

What $Circ(\Gamma(P,Z);P;Z)$?

– Should include $\Gamma(P,Z)$ because all preferred models must satisfy the original axiom set.

In addition, we would like to add a set of premises which together state the following:

> "Given $\Gamma(P,Z)$, it is not the case that there are extensions, P' and Z' for P and Z where if substituted for P and Z in Γ will both
>
> - satisfy what holds for P and Z in Γ,
>
>   and
>
> - at the same time the extension of P' will be smaller than that of P."

In other words, any substitution we make will guarantee that P has the smallest extension possible that still staisfies the constraints placed on P by Γ.

Here is the second order formula characterizing the statement:

$$\Gamma(P,Z) \wedge \neg \exists P', Z'.[\Gamma(P',Q') \wedge P' < P].$$

– $\Gamma(P',Z')$ denotes the axiomatization $\Gamma(P,Z)$ where P and Z are replaced by P' and Z', respectively.

– Predicate expressions such as P' and Z' are represented as lambda expressions $(\lambda\bar{x}.F(\bar{x}))$, where F is an arbitrary formula.

– Here is an example: $(\lambda x.Bird(x) \wedge \neg Ostrich(x))$.

Some Equivalent Forms:

$$\Gamma(P,Z)\wedge\neg\exists P',Z'.[\Gamma(P',Q')\wedge P' < P].$$

$$\neg\exists P',Z'.[\Gamma(P',Z')\wedge(\forall\bar{x}.P'(\bar{x})\rightarrow P(\bar{x}))\wedge(\neg\forall\bar{x}.P(\bar{x})\rightarrow P'(\bar{x}))]$$

$$\forall P',Z'.\neg[\Gamma(P',Z')\wedge(\forall\bar{x}.P'(\bar{x})\rightarrow P(\bar{x}))\wedge(\neg\forall\bar{x}.P(\bar{x})\rightarrow P'(\bar{x}))]$$

$$\forall P',Z'.\neg\Gamma(P',Z')\vee(\neg\forall\bar{x}.P'(\bar{x})\rightarrow P(\bar{x}))\vee(\forall\bar{x}.P(\bar{x})\rightarrow P'(\bar{x}))$$

$$\forall P',Z'.\Gamma(P',Z')\rightarrow((\forall\bar{x}.P'(\bar{x})\rightarrow P(\bar{x}))\rightarrow(\forall\bar{x}.P(\bar{x})\rightarrow P'(\bar{x})))$$

$$\forall P',Z'.[\Gamma(P',Z')\wedge(\forall\bar{x}.P'(\bar{x})\rightarrow P(\bar{x}))]\rightarrow(\forall\bar{x}.P(\bar{x})\rightarrow P'(\bar{x})).\ (*)$$

It's straightforward to show that (*) is equivalent to

$$\forall P',Z'.[\Gamma(P',Z')\wedge(\forall\bar{x}.P'(\bar{x})\rightarrow P(\bar{x}))]\rightarrow(\forall\bar{x}.P(\bar{x})\equiv P'(\bar{x})).\ (**)$$

$$\forall P', Z'.[\Gamma(P', Z') \wedge (\forall \bar{x}.P'(\bar{x}) \rightarrow P(\bar{x}))] \rightarrow (\forall \bar{x}.P(\bar{x}) \rightarrow P'(\bar{x})). \quad (*)$$

$$\forall P', Z'.[\Gamma(P', Z') \wedge (\forall \bar{x}.P'(\bar{x}) \rightarrow P(\bar{x}))] \rightarrow (\forall \bar{x}.P(\bar{x}) \equiv P'(\bar{x})). \quad (**)$$

– (*) and (**) provide the key to understanding circumscription.

– Find two predicate expressions P' and Z' that make the antecedents of the universally quantified sentence (*) true.

1. $\Gamma(P', Z')$ guarantees that everything that holds for P and Z in Γ also holds for the substituted predicate expressions P' and Z'.

2. $\forall \bar{x}.P'(\bar{x}) \rightarrow P(\bar{x})$ guarantees that only those objects proven to be in the extension of P belong to the extension of P'.

– If the two antecedents are true then by (**), P is in fact equivalent to the substituted predicate expression P'.

Definition 3 (Parallel 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* parallel circumscription *of* $\overline{P}$ *in* $\Gamma(\overline{P},\overline{S})$ *with variable* $\overline{S}$*, written* $Circ(\Gamma;\overline{P};\overline{S})$*, is the sentence*

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

where $\overline{P}'$ *and* $\overline{S}'$ *are tuples of variables similar to* $\overline{P}$ *and* $\overline{S}$*, respectively.*

Example

We would like to minimize Ab, vary $Flies$, with $Bird$, $Tweety$ and $Fred$ remaining fixed. Our premise set is

$$\begin{array}{c} \Gamma(Ab, Flies) = \\ \{\forall x.Bird(x) \wedge \neg Ab(x) \rightarrow Flies(x), \\ Bird(Tweety) \\ \forall x.x = Tweety \vee x = Fred \\ Tweety \neq Fred\} \end{array}$$

$\Gamma(Ab', Flies')$ is

$$\begin{array}{c} \Gamma(Ab', Flies') = \\ \{\forall x.Bird(x) \wedge \neg Ab'(x) \rightarrow Flies'(x), \\ Bird(Tweety) \\ \forall x.x = Tweety \vee x = Fred \\ Tweety \neq Fred\} \end{array}$$

$Circ(\Gamma(Ab, Flies); Ab; Flies)$ is

$$\begin{array}{l} \Gamma(Ab, Flies) \wedge \forall Ab', Flies'.[\Gamma(Ab', Flies') \wedge \\ \quad \forall \bar{x}.Ab'(\bar{x}) \rightarrow Ab(\bar{x})] \rightarrow \quad \forall \bar{x}.Ab(\bar{x}) \equiv Ab'(\bar{x}). \end{array}$$

Choose a Substitution:

– Find two suitable predicate expressions to substitute for Ab' and $Flies'$. Using some wishful thinking, we can make the following educated guess:

$$Ab'(x) = (\lambda x.\bot),$$

where $\bot$ denotes falsity, and

$$Flies'(x) = (\lambda x.Bird(x)).$$

Show that Antecedents of CIRC axiom are true:

– Next, show using $\Gamma(Ab, Flies)$ that the antecedents $\Gamma(Ab', Flies')$ and $\forall\bar{x}.Ab'(\bar{x}) \rightarrow Ab(\bar{x})$ in the universal formula are true.

$\underline{\Gamma(Ab', Flies')}$:

– Substitute our choice for Ab' and $Flies'$:

$$\begin{aligned}
&\Gamma(Ab', Flies') = \\
&\quad \{\forall x.Bird(x) \wedge \neg\bot \rightarrow Bird(x), \\
&\quad\quad Bird(Tweety) \\
&\quad \forall x.x = Tweety \vee x = Fred \\
&\quad\quad Tweety \neq Fred\}
\end{aligned}$$

– It is easy to see that all premises in $\Gamma(Ab', Flies')$ are true.

$\underline{\forall\bar{x}.Ab'(\bar{x}) \to Ab(\bar{x})}$:

– Substitute our choice of Ab' into the second antecedent

$$\forall\bar{x}.\bot \to Ab(\bar{x}),$$

which again is always true.

Modus Ponens Application:

Since we have shown that the antcedents in the universal formula are true, we may derive

$$\forall\bar{x}.Ab(\bar{x}) \equiv Ab'(\bar{x})$$

using our choice of Ab'.

– Substituting in gives $\forall\bar{x}.Ab(\bar{x}) \equiv \bot$ which is equivalent to $\forall\bar{x}.\neg Ab(\bar{x})$.

Using the Conjecture:

– We can now use this conjecture to derive $\neg Ab(Tweety)$.

– This may then be used to derive that $Flies(Tweety)$, using the fact that $Bird(Tweety)$ is one of the premises.

Nonmonotonicity

– Add $Ab(Tweety)$ to my original premise set.

New premise set:

$$\begin{array}{c} \Gamma(Ab, Flies) = \\ \{\forall x.Bird(x)\wedge\neg Ab(x)\to Flies(x), \\ Bird(Tweety) \\ Ab(Tweety) \\ \forall x.x = Tweety \vee x = Fred \\ Tweety \neq Fred\} \end{array}$$

$\Gamma(Ab', Flies')$ is

$$\begin{array}{c} \Gamma(Ab', Flies') = \\ \{\forall x.Bird(x)\wedge\neg Ab'(x)\to Flies'(x), \\ Bird(Tweety) \\ Ab'(Tweety) \\ \forall x.x = Tweety \vee x = Fred \\ Tweety \neq Fred\} \end{array}$$

$Circ(\Gamma(Ab, Flies); Ab; Flies)$ is

$$\Gamma(Ab, Flies)\wedge\forall Ab', Flies'.[\Gamma(Ab', Flies')\wedge \forall\bar{x}.Ab'(\bar{x})\to Ab(\bar{x})]\to\forall\bar{x}.Ab(\bar{x}) \equiv Ab'(\bar{x}).$$

Choose a Substitution:

– Find two suitable predicate expressions to substitute for Ab' and $Flies'$. Using some wishful thinking, we can make the following educated guess:

$$Ab'(x) = (\lambda x.x = Tweety),$$

where $\perp$ denotes falsity, and

$$Flies'(x) = (\lambda x.Bird(x) \wedge x \neq Tweety).$$

Show that Antecedents of CIRC axiom are true:

– Substituting in as before, we can show that the antecedents of the universal formula are true.

Using the Conjecture:

– In addition, we can no longer derive that $Flies(Tweety)$.