TECHNISCHE UNIVERSITÄT DRESDEN

DRESDEN concept

Faculty of Computer Science Chair of Automata Theory

INTRODUCTION TO NONMONOTONIC REASONING

Anni-Yasmin Turhan

Dresden, WS 2019/20

Section 5

Circumscription

Subsection 5.1

Introducing Circumscription

Circumscription

- was developed by John McCarthy, refined by Vladimir Lifschitz in the eighties
- circumscription does not extend the underlying logic syntactically to provide nonmonotonic reasoning—unlike Default logic or autoepistemic logic
- uses designated predicates, whose interpretation is to be minimized
- for reasoning not all models of a sentence are regarded;
  Circumscription uses preferred models
- often circumscription uses Second-order logic
  (we concentrate here on First-order logic for simplicity)
- simple form of circumscription:
  uses First-order logic theory T, a set of designated formulae $circ(T)$ and classical reasoning

Example 5.1 (Introductory example)

Consider the FOL theory T:

$\forall X(bird(X) \land \neg abnormal(X) \longrightarrow flies(X))$, 'All birds that are not abnormal fly.'
$bird(tweety)$ 'Tweety is a bird.'

Wanted consequence: $flies(tweety)$

In classical FOL this does not follow, since $\neg abnormal(tweety)$ cannot be derived from T. (Tweety could be abnormal.)

Idea of circumscription:
minimize the set of objects for which the predicate *abnormal* is true to those objects a for which there is definite information that $abnormal(a)$ is true.

Approach of Circumscription —Example 5.1 cont.

We add $\forall X \neg abnormal(X)$ to the set $circ(T)$. Now, from

$$\begin{aligned} T \cup circ(T) \quad &= \quad \{\forall X(bird(X) \wedge \neg abnormal(X) \longrightarrow flies(X)), bird(tweety)\} \\ &\qquad \cup \ \{\forall X \neg abnormal(X), \ldots\} \end{aligned}$$

it follows that $flies(tweety)$.

The effect of adding $\forall X \neg abnormal(X)$ to $circ(T)$ is the following:
all models of T that have non-empty interpretations of *abnormal* get eliminated and only models with minimal interpretations of the predicate *abnormal* remain.

Approach of circumscription:
minimize the interpretations of certain predicates, thereby eliminating many models of T and thus enabling more logical conclusions.

Subsection 5.2

Predicate circumscription

Circumscription minimizes the interpretation of certain predicates.

We consider the minimization of one predicate first.

Example 5.2

Consider the formula: $isBlock(a) \wedge isBlock(b)$.

We want to minimize the predicate *isBlock* and thus expect a and b to be the only blocks. Essentially, formula $(X = a \vee X = b)$ should replace the predicate $isBlock(X)$.

Definition 5.3

A predicate expression of arity n consist of a formula ψ and the distinguished variables $X_1, \ldots, X_n$.

Intuitively, such expressions are possible candidates for “replacing” an n-ary predicate symbol.

Definition 5.4

Let φ be a closed formula, p an n-ary predicate symbol, and ψ a predicate expression of arity n with distinguished variables $X_1, \ldots, X_n$.

The result of substituting ψ for p in φ (denoted as $\varphi[p/\psi]$) is defined inductively for FOL formulae:

- $q(t_1, \ldots, t_k)[p/\psi] = q(t_1, \ldots, t_k)$, if q is a predicate name and $q \neq p$,
- $p(t_1, \ldots, t_n)[p/\psi] = \psi\{X_1/t_1, \ldots, X_n/t_n\}$,
- $(\varphi_1 * \varphi_2)[p/\psi] = (\varphi_1[p/\psi] * \varphi_2[p/\psi])$, for $* \in \{\wedge, \vee, \longrightarrow\}$,
- $(\neg\varphi)[p/\psi] = \neg(\varphi[p/\psi])$, and
- $(Q\,X\varphi)[p/\psi] = Q\,X(\varphi[p/\psi])$, for $Q \in \{\forall, \exists\}$.

A substitution $\varphi[p/\psi]$ is admissible iff no occurrence of a variable of ψ other than $X_1, \ldots, X_n$ is placed into the scope of a quantifier in φ.

Let T be a finite first-order theory, $T[p/\psi]$ denotes the set $\{\varphi[p/\psi] \mid \varphi \in T\}$.

Intuition for admissibility of a predicate substitution:
It prevents unwanted semantic effects by excluding $X_1, \ldots, X_n$ and thus ensures that they are treated as the n arguments of p.

Considerations for defining circumscription

1. In Example 5.2, the result of substituting $(X = a \lor X = b)$ for $\mathit{isBlock}(X)$ in the formula $(\mathit{isBlock}(a) \land \mathit{isBlock}(b))$ is $((a = a \lor a = b) \land (b = a \lor b = b))$.
   This formula is valid.

2. Suppose, *isBlock* is radically minimized and nothing is a block.
   Then the result of substituting *false* for $\mathit{isBlock}(X)$ in the formula $(\mathit{isBlock}(a) \land \mathit{isBlock}(b))$ is $(\mathit{false} \land \mathit{false})$—a false formula!

   Minimization of a predicate should not violate the given information!

3. If a predicate expression ψ_p is known to be "smaller" than a predicate p (i.e. $\psi_p \longrightarrow p$), then ψ_p is a candidate to minimize p.

   If ψ_p "satisfies" the given information (from formula φ), then one may restrict p in φ to ψ_p.

   $\leadsto$ p is not allowed to satisfy more tuples than ψ_p does!

Definition 5.5

Let φ be a closed first-order formula containing an n-ary predicate p.
Let ψ_p be a predicate expression of arity n with distinguished variables $X_1, \ldots, X_n$ such that $\varphi[p/\psi_p]$ is admissible.

The circumscription of p in φ by ψ_p is the following formula:

$$\Big(\varphi[p/\psi_p] \ \wedge\ \forall X_1 \cdots \forall X_n \big(\psi_p \longrightarrow p(X_1, \ldots, X_n)\big)\Big) \\ \longrightarrow \forall X_1 \cdots \forall X_n \big(p(X_1, \ldots, X_n) \longrightarrow \psi_p\big).$$

If ψ_p can vary, then this formula is a schema called the circumscription of p in φ.
The set of all formulae of the form above for varying ψ_p is denoted $Circum(\varphi, p)$.

A formula χ is derivable from φ with circumscription of p (denoted $\{\varphi\} \vdash_{Circ(p)} \chi$) iff $\{\varphi\} \cup Circum(\varphi, p) \models \chi$.

The generalization of these notions to finite sets of closed predicate logic formulae is straightforward (and is left as an exercise).

Applying the definition of circumscription to Example 5.2

In Example 5.2, we have:

$$\begin{array}{rcl} \varphi & = & \big(isBlock(a) \wedge isBlock(b)\big) \\ \psi_p & = & (X = a \vee X = b) \\ p & = & isBlock(X) \end{array}$$

Circumscription of *isBlock* in $(isBlock(a) \wedge isBlock(b))$ yields the schema (for general ψ):

$$\Big(\big(\psi(a) \wedge \psi(b)\big) \wedge \forall X\big(\psi(X) \longrightarrow isBlock(X)\big)\Big) \longrightarrow \forall X\big(isBlock(X) \longrightarrow \psi(X)\big).$$

The conclusion is in our case: $\forall X\big(isBlock(X) \longrightarrow (X = a \vee X = b)\big)$.

We therefore have:
$\{isBlock(a) \wedge isBlock(b)\} \vdash_{Circ(isBlock)} \forall X(isBlock(X) \longrightarrow (X = a \vee X = b))$.

Now, *a* and *b* are the only blocks!

What happens if $block(c)$ is added? Then $\forall X(isBlock(X) \longrightarrow (X = a \vee X = b))$ can no longer be derived! Yes, circumscription is nonmonotonic!

Example: treating missing information

Consider the formula: $\varphi = \neg p(a)$.

It is impossible to derive $p(t)$ for any term t and thus the minimization should yield $\forall X \neg p(X)$.

Circumscription of p in $\neg p(a)$ produces the schema:

$$\big(\neg\psi_p(a) \wedge \forall X\big(\psi_p(X) \longrightarrow p(X)\big)\big) \;\longrightarrow\; \forall X\big(p(X) \longrightarrow \psi_p(X)\big).$$

Since p should not be true for any argument, we chose: $\psi_p \equiv \mathit{false}$ and get

$$\begin{aligned}\big(\neg \mathit{false} \wedge \forall X\big(\mathit{false} \longrightarrow p(X)\big)\big) \;\longrightarrow\; \forall X\big(p(X) \longrightarrow \mathit{false}\big) \quad &\equiv \quad \forall X\big(p(X) \longrightarrow \mathit{false}\big)\\ &\equiv \quad \forall X \neg p(X)\end{aligned}$$

as desired!

Reprise: Closed world assumption

Closed world assumption (CWA) is another formalism
based on the idea of minimizing interpretations of predicates.

According to CWA, $\neg p(t)$ is obtained for every ground term t such that
$p(t)$ does not follow from the given knowledge.

CWA and circumscription do behave differently!

Closed world assumption vs. circumscription

To see the difference, consider $\varphi = \mathit{isBlock}(a) \vee \mathit{isBlock}(b)$
Expected conclusion: "there is one block, and it is either *a* or *b*".

1. applying circumscription:
   By use of $\psi_{\mathit{isBlock}}(X) \equiv (X = a)$ in the circumscription schema of *isBlock* in φ we get:
   $$\mathit{isBlock}(a) \longrightarrow \forall X\big(\mathit{isBlock}(X) \longrightarrow (X = a)\big)$$
   Analogous formula is obtained for $\psi_{\mathit{isBlock}}(X) \equiv (X = b)$.
   Together with φ this yields:
   $$\forall X\big(\mathit{isBlock}(X) \longrightarrow X = a\big) \;\vee\; \forall X\big(\mathit{isBlock}(X) \longrightarrow X = b\big).$$

2. applying CWA:
   Neither $\mathit{isBlock}(a)$ nor $\mathit{isBlock}(b)$ follows from φ, thus $\neg\mathit{isBlock}(a)$ and $\neg\mathit{isBlock}(b)$ is implied. But together this yields a contradiction!

While CWA "misbehaves", circumscription yields the expected result!

Generalization of circumscription to several predicates

Predicate circumscription can easily be generalized to allow minimization of several predicates simultaneously.

For example, circumscription of p and q in φ is given by the schema:

$$\Big(\varphi[p/\psi_p, q/\psi_q] \ \wedge \\ \forall X_1,\ldots,X_n\big(\psi_p \longrightarrow p(X_1,\ldots,X_n)\big) \wedge \forall Y_1,\ldots,Y_m\big(\psi_q \longrightarrow q(Y_1,\ldots,Y_m)\big)\Big) \longrightarrow \\ \Big(\forall X_1,\ldots,X_n\big(p(X_1,\ldots,X_n) \longrightarrow \psi_p\big) \wedge \forall Y_1,\ldots,Y_m\big(q(Y_1,\ldots,Y_m) \longrightarrow \psi_q\big)\Big),$$

where ψ_p, ψ_q are suitable predicate expressions of the same arity as p and q, respectively, and such that $\varphi[p/\psi_p, q/\psi_q]$ is admissible.

For a finite set P of predicate symbols, $\vdash_{Circ(P)}$ is defined in the obvious way.

Subsection 5.3

Minimal models

Semantic aspects of minimizing predicates

Consider Example 5.2 again: $\varphi = (\mathit{isBlock}(a) \wedge \mathit{isBlock}(b))$.
Circumscription of *isBlock* in φ derives

$$\forall X\big(\mathit{isBlock}(X) \longrightarrow (X = a \vee X = b)\big) \equiv \forall X\big((\neg(X = a) \wedge \neg(X = b)) \longrightarrow \neg \mathit{isBlock}(X)\big).$$

Thus from all models $\mathcal{I}$ of φ only those that interpret *isBlock* as being true for $a^{\mathcal{I}}$ and $b^{\mathcal{I}}$ only, are models of $\{\varphi\} \cup \mathit{Circum}(\varphi, \mathit{isBlock})$.

Consider the interpretation $\mathcal{J}$ defined as:

- $\mathit{dom}(\mathcal{J}) = \{1, 2, 3, 4\}$,
- $a^{\mathcal{J}} = 1, \quad b^{\mathcal{J}} = 2$,
- $\mathit{isBlock}^{\mathcal{J}} = \{(1), (2), (3)\}$

$\mathcal{J}$ is a model of φ, but not of $\{\varphi\} \cup \mathit{Circum}(\varphi, \mathit{isBlock})$.

Now, $\mathcal{J}$ can be made smaller: define $\mathcal{J}'$ as $\mathcal{J}$, but $\mathit{isBlock}^{\mathcal{J}'} = \{(1), (2)\}$.
Obviously: $\mathit{isBlock}^{\mathcal{J}'} \subset \mathit{isBlock}^{\mathcal{J}}$.
$\mathcal{J}'$ cannot be minimized further and still be a model of φ!

Definition 5.6

Let T be a finite first-order theory in a signature containing the predicates symbols $P = \{p_1, \ldots p_k\}$. Let $\mathcal{I}$ and $\mathcal{J}$ be models of T.

$\mathcal{I}$ is called a P-submodel of $\mathcal{J}$ (denoted $\mathcal{I} \leq^P \mathcal{J}$), iff the following conditions hold:

- $dom(\mathcal{I}) = dom(\mathcal{J})$,
- $f^{\mathcal{I}} = f^{\mathcal{J}}$, for all function symbols f,
- $p^{\mathcal{I}} = p^{\mathcal{J}}$, for all predicate symbols $p \notin P$
- $p^{\mathcal{I}} \subseteq p^{\mathcal{J}}$, for all predicate symbols $p \in P$

A model $\mathcal{I}$ of T is called P-minimal iff every model of T which is a P-submodel of $\mathcal{I}$ is identical with $\mathcal{I}$.

Soundness of predicate circumscription

Theorem 5.7

Let T be a finite set of closed first-order formulae, $P = \{p_1, \ldots, p_k\}$ a set of predicate symbols, and χ a formula.
If $T \vdash_{Circ(P)} \chi$ then every P-minimal model of T is a model of χ.

Proof: blackboard.

Completeness of circumscription need not hold

$T \cup \{Circum(T, P)\}$ is too weak, since it admits too many models. The relations in the models are expressible in FOL, but there exist relations (on the interpretation domain) not expressible in FOL.

In the circumscription schema $Circum(T, P)$ describes P-minimality only for relations expressible in predicate logic.

How to re-gain completeness?

Use second order logic to quantify over relations.

Then the circumscription schema becomes:

$$\forall \psi \Big(\varphi[p/\psi] \ \wedge \ \forall X_1, \ldots, X_n(\psi \longrightarrow p(X_1, \ldots, X_n)) \Big) \longrightarrow \ \Big(\forall X_1, \ldots, X_n(p(X_1, \ldots, X_n) \longrightarrow \psi) \Big),$$

Now ψ is not limited to predicate expressions anymore!

Subsection 5.4

Consistency and expressive power

Consistency preservation:

if T is consistent, then $T \cup \{Circum(T, P)\}$ consistent as well.

Predicate circumscription does not preserve consistency.

By Theorem 5.7, inconsistency can only occur if T does not have a minimal model.

$\rightsquigarrow$ Restrict to theories for which minimal models always exist!

A special case: finite universal sets

A set of closed formulae is called universal iff the prenex normal form of all of its formulae does not contain any existential quantifier.

Sets of universal formulae always have a minimal model!

Theorem 5.8

Let T be a finite, consistent, universal set of closed formulae, and P a finite set of predicate symbols.
Then there exists a P-minimal model of T. Consequently, $T \cup Circum(T, P)$ is consistent.

Expressive power of circumscription

By (predicate) circumscription no new facts regarding the predicates not being circumscribed can be derived about ground terms.

$\leadsto$ New information regarding ground terms can only be obtained for circumscribed predicates!

Theorem 5.9

Let T be a finite, universal set of closed formulae, P a finite set of predicate symbols, p an n-ary predicate symbol with $p \notin P$ and $t_1, \ldots, t_n$ ground terms. Then

1. $T \vdash_{Circ(P)} p(t_1, \ldots, t_n)$ *iff* $T \models p(t_1, \ldots, t_n)$.
2. $T \vdash_{Circ(P)} \neg p(t_1, \ldots, t_n)$ *iff* $T \models \neg p(t_1, \ldots, t_n)$.

Proof: blackboard

Applying Theorem 5.9 to the Tweety example

Consider again:
$\forall X(\mathit{bird}(X) \wedge \neg \mathit{abnormal}(X) \longrightarrow \mathit{flies}(X))$,
$\mathit{bird}(\mathit{tweety})$

By Theorem 5.9 the circumscription of *abnormal* cannot derive $\mathit{flies}(\mathit{tweety})$!

To see this, consider the interpretation $\mathcal{I}$ with

- $\mathit{dom}(\mathcal{I}) = \{1\}$,
- $\mathit{tweety}^{\mathcal{I}} = 1$
- $\mathit{bird}^{\mathcal{I}} = \mathit{abnormal}^{\mathcal{I}} = \{1\}$, and
- $\mathit{flies}^{\mathcal{I}} = \emptyset$

Now, $\mathcal{I}$ is a model of T and $\mathit{flies}(\mathit{tweety})$ is not true in $\mathcal{I}$.

The predicate $\mathit{abnormal}^{\mathcal{I}}$ cannot be reduced while keeping $\mathit{flies}^{\mathcal{I}}$ and validity of T, thus $\mathcal{I}$ is $\{\mathit{abnormal}\}$-minimal. Then by Theorem 5.7 it follows that

$$T \not\models_{\mathit{Circ}(\mathit{abnormal})} \mathit{flies}(\mathit{tweety}).$$

Predicate circumscription does not suffice to realize default reasoning!