Computational complexity of reasoning in autoepistemic logics

- For closed FOL formulas in a logic $L$ holds: if satisfiability in $L$ is decidable, then so are nonmonotonic reasoning tasks for $L$.

- deciding whether an AE-theory has a stable expansion: $\Sigma_2^P$-complete

- credulous reasoning is $\Sigma_2^P$-complete
  cautious reasoning is $\Pi_2^P$-complete
Section 5

Circumscription

Subsection 5.1

Introducing Circumscription
Circumscription

- 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
- often circumscription uses second order logic (we concentrate here on FOL for simplicity)
- simple form of circumscription: uses FOL theory $T$, a set of designated formulas $\text{circ}(T)$ and classical reasoning
Introductory example

Consider the FOL theory $T$:

\[
\forall X (\text{bird}(X) \land \neg \text{abnormal}(X) \rightarrow \text{flies}(X))
\]

\[\text{bird(tweety)}\]

‘All birds that are not abnormal fly.’

\[\text{Wanted consequence: flies(tweety)}\]

‘Tweety is a bird.’

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

Idea of circumscription:
minimize the set of objects for which the predicate $\text{abnormal}$ is true to those objects $a$ for which there is definite information that $\text{abnormal}(a)$ is true.
Introductory example cont.

We add $\forall X \neg \text{abnormal}(X)$ to the set $\text{circ}(T)$.
Now, from $T \cup \text{circ}(T)$ it follows that $\text{flies(tweety)}$.

By adding $\forall X \neg \text{abnormal}(X)$ to $\text{circ}(T)$, all models of $T$ that have non-empty interpretations of $\text{abnormal}$ get eliminated and only models with minimal interpretations of $\text{abnormal}$ 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
Replacement of predicate symbols

Circumscription minimizes the interpretation of certain predicates. We consider the minimization of one predicate first.

**Example 5.1**
Given the formula \( isBlock(a) \land isBlock(b) \), we want to minimize the predicate \( isBlock \) and thus expect \( a \) and \( b \) to be the only blocks. Essentially, formula \( (X = a \lor X = b) \) should replace the predicate \( isBlock(X) \).

**Definition 5.2**
A predicate expression of arity \( n \) consist of a formula \( \psi \) and the distinguished variables \( X_1, \ldots, X_n \).
Intuitively, such expressions are possible candidates for replacing an \( n \)-ary predicate symbol.
Substitutions by predicate expressions

Definition 5.3
Let \( \varphi \) be a closed formula, \( p \) and \( n \)-ary predicate symbol, and \( \psi \) a predicate expression of arity \( n \) with distinguished variables \( X_1, \ldots, X_n \).

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

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

\( \varphi[p/\psi] \) is admissible iff no occurrence of a variable of \( \psi \) other than \( X_1, \ldots, X_n \) is replaced in the scope of a quantifier in \( \varphi \).

Let \( T \) be a finite first-order theory, \( T[p/\psi] \) denotes the set \( \{\varphi[p/\psi] \mid \varphi \in T\} \).
Considerations for defining circumscription

1. If a predicate expression $\psi_p$ is known to be ‘smaller’ than a predicate $p$ (i.e. $\psi_p \rightarrow p$), then $\psi_p$ is a candidate to minimize $p$.

2. In Example 5.1, the result of substituting $X = a \lor X = b$ for $isBlock(X)$ in the formula $isBlock(a) \land isBlock(b)$ is $(a = a \lor a = b) \land (b = a \lor b = b)$. This formula is valid.

Suppose, $isBlock$ is radically minimized and nothing is a block. Then the result of substituting $false$ for $isBlock(X)$ in the formula $isBlock(a) \land isBlock(b)$ is $false \land false$.

Generally, minimization of a predicate should not violate the given information!

If $\psi_p$ ‘satisfies’ the given information (from formula $\varphi$), then one may restrict $p$ in $\varphi$ to $\psi_p$, i.e., $p$ is not allowed to satisfy more tuples than $\psi_p$ does.
Circumscription

**Definition 5.4**

Let $\varphi$ be a closed first-order formula containing an $n$-ary predicate $p$.
Let $\psi_p$ 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 $\varphi$ by $\psi_p$ is the following formula:

\[
(\varphi[p/\psi_p] \land \forall X_1 \cdots \forall X_n(\psi_p \rightarrow p(X_1, \ldots, X_n))) \\
\rightarrow \forall X_1 \cdots \forall X_n(p(X_1, \ldots, X_n) \rightarrow \psi_p).
\]

If $\psi_p$ can vary, then this formula is a schema called the *circumscription* of $p$ in $\varphi$.

The set of all formulas of the form above for varying $\psi_p$ is denoted $\text{Circ}(\varphi, p)$.

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

The generalization of these notions to finite sets of closed predicate logic formulas is straightforward and is left as an exercise.
Applying the definition of circumscription to Example 5.1

In Example 5.1, we have:

\[ \varphi = (\text{isBlock}(a) \land \text{isBlock}(b)) \]
\[ \psi_p = (X = a \lor X = b) \]
\[ p = \text{isBlock}(X) \]

Circumscription of \text{isBlock} in \text{isBlock}(a) \land \text{isBlock}(b) yields the schema (for general \( \psi \)):

\[(\psi(a) \land \psi(b)) \land \forall X (\psi(X) \rightarrow \text{isBlock}(X)) \rightarrow \forall X (\text{isBlock}(X) \rightarrow \psi(X)).\]

The conclusion is in our case: \( \forall X (\text{isBlock}(X) \rightarrow (X = a \lor X = b)) \).

We therefore have:
\[ \{ \text{isBlock}(a) \land \text{isBlock}(b) \} \models_{\text{Circ(isBlock)}} \forall X (\text{isBlock}(X) \rightarrow (X = a \lor X = b)). \]

Now, \( a \) and \( b \) are the only blocks!
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:

$$(\neg \psi_p(a) \land \forall X(\psi_p(X) \rightarrow p(X))) \rightarrow \forall X(p(X) \rightarrow \psi_p(X)).$$

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

$$(\neg false \land \forall X(false \rightarrow p(X))) \rightarrow \forall X(p(X) \rightarrow false) \equiv \forall X(p(X) \rightarrow false) \equiv \forall X \neg p(X)$$

as desired!
Closed world assumption vs. circumscription

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!

To see this, consider \( \varphi = isBlock(a) \vee isBlock(b) \)

Expected conclusion: ‘there is one block, and it is either \( a \) or \( b \).

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

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

While CWA ‘misbehaves’, circumscription yields the expected result.
Generalization to several predicates

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

For example, circumscription of $p$ and $q$ in $\varphi$ is given by the schema:

$$\left( \varphi[p/\psi_p, q/\psi_q] \land 
\forall X_1, \ldots, X_n(\psi_p \rightarrow p(X_1, \ldots, X_n)) \land \forall Y_1, \ldots, Y_m(\psi_q \rightarrow q(Y_1, \ldots, Y_m)) \right) \rightarrow 
\left( \forall X_1, \ldots, X_n(p(X_1, \ldots, X_n) \rightarrow \psi_p) \land \forall Y_1, \ldots, Y_m(q(Y_1, \ldots, Y_m) \rightarrow \psi_q) \right),$$

where $\psi_p, \psi_q$ 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, $\models_{\text{Circ}(P)}$ is defined in the obvious way.
Subsection 5.3
Minimal models
Semantic aspects of minimizing predicates

Consider Example 5.1 again: $\varphi = \text{isBlock}(a) \land \text{isBlock}(b)$. Circumscription of $\text{isBlock}$ in $\varphi$ derives

$$\forall X (\text{isBlock}(X) \rightarrow (X = a \lor X = b)) \equiv \forall X ((\neg (X = a) \land \neg (X = b)) \rightarrow \neg \text{isBlock}(X)).$$

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

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

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

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

Now, $\mathcal{J}$ can be made smaller: define $\mathcal{J}'$ as $\mathcal{J}$, but $\text{isBlock}^{\mathcal{J}'} = \{(1), (2)\}$. Obviously: $\text{isBlock}^{\mathcal{J}' \mathcal{J}} \subset \text{isBlock}^\mathcal{J}$.

$\mathcal{J}'$ cannot be minimized further and still be a model of $\varphi$!
Definition 5.5
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 by $\mathcal{I} \leq^P \mathcal{J}$, iff the following conditions hold:

- $\text{dom}(\mathcal{I}) = \text{dom}(\mathcal{J})$,
- $f^\mathcal{I} = f^\mathcal{J}$, 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 5.6**

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

Proof: blackboard.