TECHNISCHE UNIVERSITÄT DRESDEN

Computational Logic ∴ Group

Foundations of Knowledge Representation

Lecture 8: Nonmonotonic Reasoning - II

Hannes Straß

based on slides of
Bernardo Cuenca Grau,
Ian Horrocks, and
Przemysław Wałęga

Datalog & Least Herbrand Models

We have seen so far:

- It is easy to formalise intuitions about preferred models if we have a least Herbrand model
- In that case, everyone agrees that the least Herbrand model is the right choice
- Datalog knowledge bases have a least Herbrand model, which can be computed deterministically using forward chaining
- We can successfully formalise the Closed World Assumption

However, we cannot express default statements:

$$\frac{\mathit{hasOrg}(x, y) \wedge \mathit{Heart}(y) \text{ \& consistent to assume } \mathit{hasLocation}(y, \mathit{left})}{\text{deduce } \mathit{hasLocation}(y, \mathit{left})}$$

Going Beyond Datalog

To overcome expressivity limitations we next

1. Extend Datalog to a more expressive logic
2. Develop a new mechanism for selecting preferred models

Idea: First, allow for negation in the body of rules

$$\forall x.(\forall y.(hasOrg(x,y) \wedge Heart(y) \wedge \neg hasLocation(y, right) \quad \rightarrow \quad hasLocation(y, left)))$$

Then, devise a preferred model selection mechanism such that negation is read non-monotonically, as follows:

- "Deduce that heart is on left unless we prove that it is on right"
- "Deduce that heart is on left if $\neg hasLocation(y, right)$ (that is, $hasLocation(y, left)$) is consistent with our knowledge"

Datalog$^\neg$-Rules

A Datalog$^\neg$ rule is a function-free, universally quantified implication of the form

$$L_1 \wedge \ldots \wedge L_n \rightarrow H$$

With L_i a literal and H either an atom or $\bot$.

A Datalog$^\neg$ knowledge base is a pair $\mathcal{K} = \langle \mathcal{R}, \mathcal{F} \rangle$ where $\mathcal{R}$ is a finite set of Datalog$^\neg$ rules and $\mathcal{F}$ is a finite set of facts.

$$\forall x.(\mathit{Heart}(x) \wedge \mathit{hasLoc}(x, \mathit{left}) \rightarrow \mathit{NormalHeart}(x))$$
$$\forall x.(\mathit{Heart}(x) \wedge \mathit{hasLoc}(x, \mathit{right}) \rightarrow \mathit{SitInvHeart}(x))$$
$$\forall x.\forall y.(\mathit{Human}(x) \wedge \mathit{hasOrg}(x, y) \wedge \mathit{SitInvHeart}(y) \rightarrow \mathit{SitInvPatient}(x))$$
$$\forall x.\forall y.(\mathit{Human}(x) \wedge \mathit{hasOrg}(x, y) \wedge \mathit{NormalHeart}(y) \rightarrow \mathit{Healthy}(x))$$
$$\forall x.(\forall y.(\mathit{hasOrg}(x, y) \wedge \mathit{Heart}(y) \wedge \neg\mathit{hasLoc}(y, \mathit{right}) \rightarrow \mathit{hasLoc}(y, \mathit{left})))$$
$$\mathit{Human}(\mathit{MaryJones}), \mathit{hasOrg}(\mathit{MaryJones}, \mathit{MJHeart}), \mathit{Heart}(\mathit{MJHeart})$$

Semantics: Stable Models

So far all this is just syntax.

We need to specify the semantics of Datalog$^{\neg}$
$\Rightarrow$ Which are the preferred models?

There was a "war of semantics" in 1980s and 1990s.

Meaning of things like $\neg B \rightarrow A$ and $\neg A \rightarrow B$

Single vs. multiple-models semantics

To date, we have the following:

- Well-founded Semantics
- Stable-model Semantics (aka Answer Set Semantics)

Semantics: Stable Models

We will focus on Stable Model Semantics

Preferred models are called Stable Models (SM)

$$\mathcal{K} \mathrel{|\!\!\approx} \alpha \quad \text{iff} \quad \mathcal{I} \models \alpha \quad \text{for each stable model } \mathcal{I} \text{ of } \mathcal{K}$$

We will see that $\mathcal{K}$ may have

- No stable models, or
- One stable model, or
- Several stable models

Furthermore, if $\mathcal{K}$ contains only Datalog rules (i.e., no negation), then $\mathcal{K}$ has exactly one stable model (the least Herbrand model).

Semantics: Stable Models

We proceed as follows:

1. Define stable models for the propositional case
2. Extend to the case with variables using grounding

A simple propositional example $\mathcal{K}$ with one rule and one fact:

$$\begin{aligned} \mathit{Suspect} \wedge \neg \mathit{Guilty} \rightarrow \mathit{Innocent} \\ \mathit{Suspect} \end{aligned}$$

Intuitively, the rule says the following:

> “A suspect is innocent unless they can be proved guilty”

We only know that Suspect holds, so we intuitively expect that

$$\mathcal{K} \mathrel{|\!\approx} \mathit{Innocent}$$

Semantics: Stable Models

Our example:

$$\begin{aligned} \mathit{Suspect} \wedge \neg \mathit{Guilty} &\rightarrow \mathit{Innocent} \\ &\mathit{Suspect} \end{aligned}$$

Intuitively, the following (Herbrand-style) model should be stable:

$$\mathcal{I}_1 \quad = \quad \{\mathit{Suspect}, \mathit{Innocent}\}$$

To check this, we first compute the reduct $\mathcal{K}^{\mathcal{I}_1}$ of $\mathcal{K}$ by $\mathcal{I}_1$

1. Remove all rules with negative body literal $\neg A$ such that the (positive) literal A is in $\mathcal{I}_1$

   In our case, we don't remove any rule since $\mathit{Guilty} \notin \mathcal{I}_1$

2. Remove all negative literals from remaining rules

$$\begin{aligned} \mathit{Suspect} &\rightarrow \mathit{Innocent} \\ &\mathit{Suspect} \end{aligned}$$

Semantics: Stable Models

Once we have the reduct $\mathcal{K}^{\mathcal{I}_1}$

$$\begin{array}{r} \textit{Suspect} \rightarrow \textit{Innocent} \\ \textit{Suspect} \end{array}$$

We check whether $\mathcal{I}_1$ is the least Herbrand Model of $\mathcal{K}^{\mathcal{I}_1}$, in which case $\mathcal{I}_1$ is a stable model.

Indeed, by using forward chaining we can see that

$$\mathcal{I}_1 \quad = \quad \{\textit{Suspect}, \textit{Innocent}\}$$

is the least Herbrand model of $\mathcal{K}^{\mathcal{I}_1}$ and hence it's a stable model of $\mathcal{K}$

But this is not sufficient to show $\mathcal{K} \mid\!\approx \textit{Innocent}$

($\Rightarrow$) We need to look at all stable models of $\mathcal{K}$

Semantics: Stable Models

Let's check the remaining possibilities:

$$\begin{aligned}\mathcal{I}_2 &= \{\textit{Suspect}, \textit{Guilty}\}\\ \mathcal{I}_3 &= \{\textit{Suspect}, \textit{Innocent}, \textit{Guilty}\}\\ \mathcal{I}_4 &= \{\textit{Suspect}\}\end{aligned}$$

The reducts $\mathcal{K}^{\mathcal{I}_2}$ and $\mathcal{K}^{\mathcal{I}_3}$ are the same and contain just the fact:

$$\textit{Suspect}$$

This is so because *Guilty* $\in \mathcal{I}_2, \mathcal{I}_3$ and hence the reduct doesn't include the only rule we have in $\mathcal{K}$.

The least model of $\mathcal{K}^{\mathcal{I}_2}$ (or $\mathcal{K}^{\mathcal{I}_3}$) is $\mathcal{I}_4 \Rightarrow$ neither $\mathcal{I}_2$, nor $\mathcal{I}_3$ are stable.

Semantics: Stable Models

We finally check whether

$$\mathcal{I}_4 \quad = \quad \{\textit{Suspect}\}$$

is a stable model of

$$\begin{aligned} \textit{Suspect} \land \neg\textit{Guilty} \rightarrow \textit{Innocent} \\ \textit{Suspect} \end{aligned}$$

The reduct $\mathcal{K}^{\mathcal{I}_4}$ is the same as $\mathcal{K}^{\mathcal{I}_1}$, namely

$$\begin{aligned} \textit{Suspect} \rightarrow \textit{Innocent} \\ \textit{Suspect} \end{aligned}$$

But then $\mathcal{I}_4$ is not even a model of $\mathcal{K}^{\mathcal{I}_4}$.

Thus, $\mathcal{I}_1 = \{\textit{Suspect}, \textit{Innocent}\}$ is the only stable model of $\mathcal{K}$ and so $\mathcal{K} \mathrel{|\!\approx} \textit{Innocent}$.

Examples

Consider $\mathcal{K}$ as follows:

$$\begin{array}{rcl} \neg \mathit{Guilty} & \rightarrow & \mathit{Innocent} \\ \neg \mathit{Innocent} & \rightarrow & \mathit{Guilty} \end{array}$$

Recall that we compute a reduct $\mathcal{K}^{\mathcal{I}}$ of $\mathcal{K}$ by $\mathcal{I}$ as follows:

1. Remove all rules with negative body literal $\neg A$ such that the (positive) literal A is in $\mathcal{I}$
2. Remove all negative literals from remaining rules

Examples

Consider $\mathcal{K}$ as follows:

$$\neg \textit{Guilty} \quad \rightarrow \textit{Guilty}$$

Recall that we compute a reduct $\mathcal{K}^{\mathcal{I}}$ of $\mathcal{K}$ by $\mathcal{I}$ as follows:

1. Remove all rules with negative body literal $\neg A$ such that the (positive) literal A is in $\mathcal{I}$
2. Remove all negative literals from remaining rules

Non-monotonic vs Classical Negation

Consider again our propositional example $\mathcal{K}$:

$$\begin{aligned} \mathit{Suspect} \wedge \neg \mathit{Guilty} \rightarrow \mathit{Innocent} \\ \mathit{Suspect} \end{aligned}$$

Lets check whether

$$\mathcal{K} \models \mathit{Innocent}$$

That is, whether entailment holds under monotonic PL semantics.

Clearly, $\mathcal{K}$ is equivalent in standard propositional logic to

$$\begin{aligned} \mathit{Suspect} \rightarrow \mathit{Innocent} \vee \mathit{Guilty} \\ \mathit{Suspect} \end{aligned}$$

Hence $\mathcal{I} = \{\mathit{Suspect}, \mathit{Guilty}\}$ is a model of $\mathcal{K}$ and $\mathcal{I} \not\models \mathit{Innocent}$
$(\Rightarrow)$ $\mathcal{K} \not\models \mathit{Innocent}$

Properties

Let $\mathcal{K}$ be a (propositional) Datalog$^{\neg}$ knowledge base. Then, the following hold:

Theorem

Every stable model of $\mathcal{K}$ is a classical model of $\mathcal{K}$.

Corollary

If $\mathcal{K} \models \alpha$, then $\mathcal{K} \mid\!\approx \alpha$.

Theorem

If a proposition P holds in some stable model of $\mathcal{K}$, then P is a head (or a fact) of some rule in $\mathcal{K}$.

Theorem

If $\mathcal{I}_1$ and $\mathcal{I}_2$ are stable models of $\mathcal{K}$, then neither $\mathcal{I}_1 \subsetneq \mathcal{I}_2$ or $\mathcal{I}_2 \subsetneq \mathcal{I}_1$.

Stable Models: Non-Propositional Case

So far, all this is propositional.

What about...

$$\forall x.(Heart(x) \land hasLoc(x, left) \to NormalHeart(x))$$
$$\forall x.(Heart(x) \land hasLoc(x, right) \to SitInvHeart(x))$$
$$\forall x.\forall y.(Human(x) \land hasOrg(x, y) \land SitInvHeart(y) \to SitInvPatient(x))$$
$$\forall x.\forall y.(Human(x) \land hasOrg(x, y) \land NormalHeart(y) \to Healthy(x))$$
$$\forall x.(\forall y.(hasOrg(x, y) \land Heart(y) \land \neg hasLoc(y, right) \to hasLoc(y, left)))$$
$$Human(MJ), hasOrg(MJ, h), Heart(h)$$

Fortunately, we are still within Bernays-Shönfinkel class

($\Rightarrow$) We can apply grounding and reduce to propositional case.

Stable Models: Non-Propositional Case

So, to compute all the stable models of $\mathcal{K}$:

1. Compute the grounding of $\mathcal{K}$ over the Herbrand universe
2. Compute all the stable models of the resulting propositional KB

Obviously, the grounding could be of exponential size

But this is a computational hazard, not a conceptual one.

Intuitively, the following Herbrand model should be stable:

$$\begin{aligned}\mathcal{I}_1 \quad = \quad & \{\mathit{Human}(\mathit{MJ}), \mathit{hasOrg}(\mathit{MJ}, h), \mathit{Heart}(h), \mathit{hasLoc}(h, \mathit{left}) \\ & \mathit{NormalHeart}(h), \mathit{Healthy}(\mathit{MJ})\}\end{aligned}$$

Whereas the following one should not be

$$\begin{aligned}\mathcal{I}_2 \quad = \quad & \{\mathit{Human}(\mathit{MJ}), \mathit{hasOrg}(\mathit{MJ}, h), \mathit{Heart}(h), \mathit{hasLoc}(h, \mathit{right}) \\ & \mathit{SitInvHeart}(h), \mathit{SitInvPatient}(\mathit{MJ})\}\end{aligned}$$

Stable Models: Non-Propositional Case

To check whether

$$\mathcal{I}_1 = \{Human(MJ), hasOrg(MJ, h), Heart(h), hasLoc(h, left)$$
$$NormalHeart(h), Healthy(MJ)\}$$

is stable, notice that even though the grounding is huge, the only PL formulas that matter are the following:

$$Heart(h) \wedge hasLoc(h, left) \rightarrow NormalHeart(h)$$
$$Human(MJ) \wedge hasOrg(MJ, h) \wedge NormalHeart(h) \rightarrow Healthy(MJ)$$
$$hasOrg(MJ, h) \wedge Heart(h) \wedge \neg hasLoc(h, right) \rightarrow hasLoc(h, left)))$$
$$Human(MJ), hasOrg(MJ, h), Heart(h)$$

The reduct of $\mathcal{I}_1$ over those formulas is

$$Heart(h) \wedge hasLoc(h, left) \rightarrow NormalHeart(h)$$
$$Human(MJ) \wedge hasOrg(MJ, h) \wedge NormalHeart(h) \rightarrow Healthy(MJ)$$
$$hasOrg(MJ, h) \wedge Heart(h) \rightarrow hasLoc(h, left)))$$
$$Human(MJ), hasOrg(MJ, h), Heart(h)$$

And clearly $\mathcal{I}_1$ is the least model

Stable Models: Non-Propositional Case

To check whether

$$\mathcal{I}_2 = \{Human(MJ), hasOrg(MJ, h), Heart(h), hasLoc(h, right) \\ SitInvHeart(h), SitInvPatient(MJ)\}$$

is stable, the relevant PL formulas are the following:

$$Heart(h) \land hasLoc(h, right) \rightarrow SitInvHeart(h)$$
$$Human(MJ) \land hasOrg(MJ, h) \land SitInvHeart(h) \rightarrow SitInvPatient(MJ)$$
$$hasOrg(MJ, h) \land Heart(h) \land \neg hasLoc(h, right) \rightarrow hasLoc(h, left)))$$
$$Human(MJ), hasOrg(MJ, h), Heart(h)$$

The reduct of $\mathcal{I}_2$ over those formulas is

$$Heart(h) \land hasLoc(h, right) \rightarrow SitInvHeart(h)$$
$$Human(MJ) \land hasOrg(MJ, h) \land SitInvHeart(h) \rightarrow SitInvPatient(MJ)$$
$$Human(MJ), hasOrg(MJ, h), Heart(h)$$

And clearly, $\mathcal{I}_2$ is not the least model.

Quick Recap

We have seen that by using Datalog with non-monotonic negation

1. We can formalise the closed-world assumption
2. We can express default statements

The key notion is that of a Stable Model as a "preferred" model.

Checking whether a propositional model is stable involves

1. Eliminating negation by computing the reduct
2. Checking if the candidate model is the least model of the reduct

Checking whether a FOL Herbrand interpretation is a stable model involves

1. Computing the propositional grounding of the KB
2. Checking whether the candidate model is stable for the grounding

Note that stable models in the FOL case are always Herbrand models.

What have we left out?

Much more than we have covered!!

The field of NMR is huge and we have just seen the tip of the iceberg

Extensions related to what we have seen:

- Stable models and disjunctive rules (disjunction in the head)
- Stable models and general propositional formulas
- Combinations of classical and non-mon negation

Relationships with other areas

What we have seen is not only relevant to KR.

There are strong connections with other fields:

- Answer Set Programming (ASP)

  Using negation we can encode search problems

- Deductive databases,

  Database systems which can conclude new data using rules

- Logic programming (Prolog)