Knowledge Representation & Reasoning

Shyamanta M Hazarika
CSE, SoE
Tezpur University

Strictness of FOL

To reason from $P(a)$ to $Q(a)$, need either

- facts about a itself
- universals, e.g. $\forall x(P(x) \supset Q(x))$
  - something that applies to all instances
  - all or nothing!

But most of what we learn about the world is in terms of generics

e.g., encyclopedia entries for ferris wheels, violins, turtles, wildflowers

Properties are not strict for all instances, because

- genetic / manufacturing varieties
  - early ferris wheels
- cases in exceptional circumstances
  - dried wildflowers
- borderline cases
  - toy violins
- imagined cases
  - flying turtles

etc.

Generics vs. universals

✓ Violins have four strings.

vs.

× All violins have four strings.

vs.

? All violins that are not E_1 or E_2 or ... have four strings.

(exceptions usually cannot be enumerated)

Similarly, for general properties of individuals

- Alexander the great: ruthlessness
- Ecuador: exports
- pneumonia: treatment

Goal: be able to say a *P* is a *Q* <u>in general</u>, but not necessarily

It is reasonable to conclude *Q(a)* given *P(a),*
unless there is a good reason not to

Here: qualitative version (no numbers)

Varieties of defaults (I)

General statements

- prototypical: The prototypical P is a Q.

  Owls hunt at night.

- normal: Under typical circumstances, P's are Q's.

  People work close to where they live.

- statistical: Most P's are Q's.

  The people in the waiting room are growing impatient.

Lack of information to the contrary

- group confidence: All known P's are Q's.

  Natural languages are easy for children to learn.

- familiarity: If a P was not a Q, you would know it.
  - an older brother
  - very unusual individual, situation or event

Varieties of defaults (II)

Conventional

- conversational: Unless I tell you otherwise, a P is a Q

  "There is a gas station two blocks east."
  the default: the gas station is open.

- representational: Unless otherwise indicated, a P is a Q

  the speed limit in a city

Persistence

- inertia: A P is a Q if it used to be a Q.
  - colours of objects
  - locations of parked cars (for a while!)

Here: we will use "Birds fly" as a typical default.

Closed-world assumption

Reiter's observation:

There are usually many more -ve facts than +ve facts!

Example: airline flight guide provides

DirectConnect(cleveland,toronto) DirectConnect(toronto,northBay)
DirectConnect(toronto,winnipeg) ...

but not: ¬DirectConnect(cleveland,northBay)

Conversational default, called CWA:

only +ve facts will be given, relative to some vocabulary

But note: KB $\not\models$ -ve facts (would have to answer: "I don't know")

Proposal: a new version of entailment: KB $\models_c \alpha$ iff KB $\cup$ *Negs* $\models \alpha$

where *Negs* = $\{\neg p \mid p$ atomic and KB $\not\models p\}$

Note: relation to negation as failure

a common pattern:
KB´ = KB $\cup$ Δ

Gives: KB $\models_c$ +ve facts and -ve facts

Properties of CWA

For every α (without quantifiers), KB $\models_c \alpha$ or KB $\models_c \neg\alpha$

Why? Inductive argument:

- immediately true for atomic sentences
- push $\neg$ in, e.g. KB $\models \neg\neg\alpha$ iff KB $\models \alpha$
- KB $\models (\alpha \wedge \beta)$ iff KB $\models \alpha$ and KB $\models \beta$
- Say KB $\not\models_c (\alpha \vee \beta)$. Then KB $\not\models_c \alpha$ and KB $\not\models_c \beta$.
  So by induction, KB $\models_c \neg\alpha$ and KB $\models_c \neg\beta$. Thus, KB $\models_c \neg(\alpha \vee \beta)$.

CWA is an assumption about <u>complete</u> knowledge

never any unknowns, relative to vocabulary

In general, a KB has <u>incomplete</u> knowledge,

e.g. Let KB be $(p \vee q)$. Then KB $\models (p \vee q)$,
but KB $\not\models p$, KB $\not\models \neg p$, KB $\not\models q$, KB $\not\models \neg q$

With CWA, have: If KB $\models_c (\alpha \vee \beta)$, then KB $\models_c \alpha$ or KB $\models_c \beta$.

similar argument to above

Query evaluation

With CWA can reduce queries (without quantifiers) to the atomic case:

$\text{KB} \models_c (\alpha \wedge \beta)$ iff $\text{KB} \models_c \alpha$ and $\text{KB} \models_c \beta$

$\text{KB} \models_c (\alpha \vee \beta)$ iff $\text{KB} \models_c \alpha$ or $\text{KB} \models_c \beta$

$\text{KB} \models_c \neg(\alpha \wedge \beta)$ iff $\text{KB} \models_c \neg\alpha$ or $\text{KB} \models_c \neg\beta$

$\text{KB} \models_c \neg(\alpha \vee \beta)$ iff $\text{KB} \models_c \neg\alpha$ and $\text{KB} \models_c \neg\beta$

$\text{KB} \models_c \neg\neg\alpha$ iff $\text{KB} \models_c \alpha$

reduces to: $\text{KB} \models_c \rho$, where ρ is a literal

If $\text{KB} \cup \textit{Negs}$ is consistent, get $\text{KB} \models_c \neg\alpha$ iff $\text{KB} \not\models_c \alpha$

reduces to: $\text{KB} \models_c p$, where p is atomic

If atoms stored as a table, deciding if $\text{KB} \models_c \alpha$ is like DB-retrieval:

- reduce query to set of atomic queries
- solve atomic queries by table lookup

Different from ordinary logic reasoning (e.g. no reasoning by cases)

Consistency of CWA

If KB is a set of atoms, then KB $\cup$ *Negs* is always consistent

Also works if KB has conjunctions and if KB has only negative disjunctions

If KB contains $(\neg p \vee \neg q)$. Add both $\neg p, \neg q$.

Problem when KB $\models (\alpha \vee \beta)$, but KB $\not\models \alpha$ and KB $\not\models \beta$

e.g. KB = $(p \vee q)$ *Negs* = $\{\neg p, \neg q\}$

KB $\cup$ *Negs* is inconsistent and so for every α, KB $\models_c \alpha$!

Solution: only apply CWA to atoms that are "uncontroversial"

One approach: GCWA

Negs $= \{\neg p$ | If KB $\models (p \vee q_1 \vee ... \vee q_n)$ then KB $\models (q_1 \vee ... \vee q_n)$ $\}$

When KB is consistent, get:

- KB $\cup$ *Negs* consistent
- everything derivable is also derivable by CWA

Quantifiers and equality

So far, results do not extend to wffs with quantifiers

can have KB $\not\models_c \forall x.\alpha$ and KB $\not\models_c \neg\forall x.\alpha$

e.g. just because for every t, we have KB $\models_c \neg\text{DirectConnect}(\text{myHome}, t)$
does not mean that KB $\models_c \forall x[\neg\text{DirectConnect}(\text{myHome}, x)]$

But may want to treat KB as providing complete information about what individuals exist

Define: KB $\models_{cd}$ α iff KB $\cup$ *Negs* $\cup$ *Dc* $\models \alpha$

where the c_i are all the constants appearing in KB (assumed finite)

where *Dc* is <u>domain closure</u>: $\forall x[x{=}c_1 \vee \ldots \vee x{=}c_n]$,

Get: KB $\models_{cd} \exists x.\alpha$ iff KB $\models_{cd}$ $\alpha[x/c]$, for some c appearing in the KB
KB $\models_{cd} \forall x.\alpha$ iff KB $\models_{cd}$ $\alpha[x/c]$, for all c appearing in the KB

Then add: *Un* is <u>unique names</u>: $(c_i \neq c_j)$, for $i \neq j$

Get: KB $\models_{cdu} (c = d)$ iff c and d are the same constant

→ full recursive query evaluation

Non-monotonicity

Ordinary entailment is monotonic

If KB $\models \alpha$, then KB* $\models \alpha$, for any KB $\subseteq$ KB*

But CWA entailment is *not* monotonic

Can have KB $\models_c \alpha$, KB $\subseteq$ KB', but KB' $\not\models_c \alpha$

e.g. $\{p\} \models_c \neg q$, but $\{p, q\} \not\models_c \neg q$

Suggests study of <u>non-monotonic reasoning</u>

- start with explicit beliefs
- generate implicit beliefs non-monotonically, taking *defaults* into account
- implicit beliefs may not be uniquely determined (vs. monotonic case)

Will consider three approaches:

- minimal entailment: interpretations that minimize abnormality
- default logic: KB as facts + default rules of inference
- autoepistemic logic: facts that refer to what is/is not believed

Minimizing abnormality

CWA makes the extension of all predicates as small as possible

by adding negated literals

Generalize: do this only for selected predicates

Ab predicates used to talk about typical cases

Example KB: Bird(chilly), ¬Flies(chilly),
Bird(tweety), (chilly ≠ tweety),
$\forall x[\mathrm{Bird}(x) \wedge \neg\mathrm{Ab}(x) \supset \mathrm{Flies}(x)]$ ← All birds that are normal fly

Would like to conclude by default Flies(tweety), but KB |≠ Flies(tweety)

because there is an interpretation $\mathfrak{I}$ where $I[\mathrm{tweety}] \in I[\mathrm{Ab}]$

Solution: consider only interpretations where $I[\mathrm{Ab}]$ is as small as possible, relative to KB

for example: KB requires that $I[\mathrm{chilly}] \in I[\mathrm{Ab}]$

this is sometimes called "circumscription" since we circumscribe the Ab predicate

Generalizes to many Ab_i predicates

Minimal entailment

Given two interps over the same domain, $\mathfrak{I}_1$ and $\mathfrak{I}_2$

$\mathfrak{I}_1 \leq \mathfrak{I}_2$ iff $I_1[\mathrm{Ab}] \subseteq I_2[\mathrm{Ab}]$ for every Ab predicate

$\mathfrak{I}_1 < \mathfrak{I}_2$ iff $\mathfrak{I}_1 \leq \mathfrak{I}_2$ but not $\mathfrak{I}_2 \leq \mathfrak{I}_1$ read: $\mathfrak{I}_1$ is more normal than $\mathfrak{I}_2$

Define a new version of entailment, $\models_{\leq}$ by

KB $\models_{\leq} \alpha$ iff for every $\mathfrak{I}$, if $\mathfrak{I} \models$ KB and no $\mathfrak{I}^* < \mathfrak{I}$ s.t. $\mathfrak{I}^* \models$ KB
then $\mathfrak{I} \models \alpha$.

So α must be true in all interps satisfying KB that are *minimal* in abnormalities

Get: KB $\models_{\leq}$ Flies(tweety)

because if interp satisfies KB and is minimal, only *I[chilly]* will be in *I[Ab]*

Minimization need not produce a *unique* interpretation:

Bird(a), Bird(b), $[\neg\mathrm{Flies}(a) \vee \neg\mathrm{Flies}(b)]$ yields two minimal interpretations

KB $\not\models_{\leq}$ Flies(a), KB $\not\models_{\leq}$ Flies(b), KB $\models_{\leq}$ $\mathrm{Flies}(a) \vee \mathrm{Flies}(b)$

Different from the CWA: no inconsistency!

But stronger than GCWA: conclude a or b flies

Fixed and variable predicates

Imagine KB as before + $\forall x[\text{Penguin}(x) \supset \text{Bird}(x) \wedge \neg\text{Flies}(x)]$

Get: KB $\models \forall x[\text{Penguin}(x) \supset \text{Ab}(x)]$

So minimizing Ab also minimizes penguins: KB $\models_{\leq} \forall x \neg\text{Penguin}(x)$

McCarthy's definition: Let **P** and **Q** be sets of predicates

$\mathfrak{I}_1 \leq \mathfrak{I}_2$ iff same domain and

1. $I_1[P] \subseteq I_2[P]$, for every $P \in$ **P** — Ab predicates
2. $I_1[Q] = I_2[Q]$, for every $Q \notin$ **Q** — fixed predicates

so only predicates in **Q** are allowed to vary

Get definition of $\models_{\leq}$ that is parameterized by what is minimized *and* what is allowed to vary

Previous example: minimize Ab, but allow only Flies to vary.

Problems:
- need to decide what to allow to vary
- cannot conclude $\neg$Penguin(tweety) by default!

only get default $(\neg\text{Penguin}(\text{tweety}) \supset \text{Flies}(\text{tweety}))$