Default Reasoning

- When giving information, we don't want to enumerate all of the exceptions, even if we could think of them all.
- In default reasoning, we specify general knowledge and modularly add exceptions. The general knowledge is used for cases we don't know are exceptional.
- Classical logic is **monotonic:** If g logically follows from A, it also follows from any superset of A.
- Default reasoning is **nonmonotonic:** When we add that something is exceptional, we can't conclude what we could before.

Defaults as Assumptions

Default reasoning can be modeled using

- H is normality assumptions
- F states what follows from the assumptions

An explanation of g gives an argument for g.

Default Example

A reader of newsgroups may have a default:
"Articles about AI are generally interesting".

$$H = \{int\_ai\},$$

where int_ai means X is interesting if it is about AI.
With facts:

$$interesting \leftarrow about\_ai \wedge int\_ai.$$

$$about\_ai.$$

$\{int\_ai\}$ is an explanation for $interesting$.

Default Example, Continued

We can have exceptions to defaults:

$$false \leftarrow interesting \wedge uninteresting.$$

Suppose an article is about AI but is uninteresting:

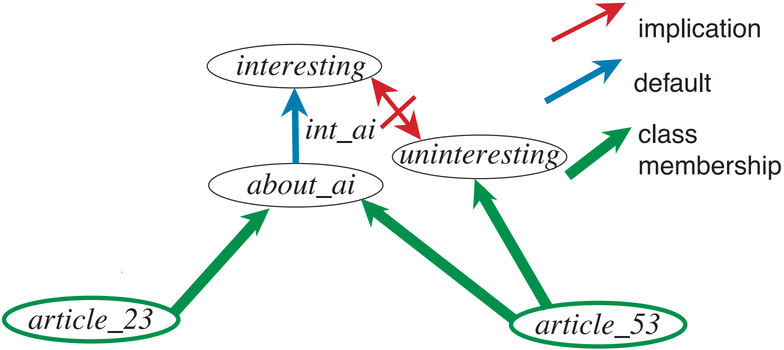
$$interesting \leftarrow about\_ai \wedge int\_ai.$$

$$about\_ai.$$

$$uninteresting.$$

We cannot explain *interesting* even though everything we know about the previous we also know about this case.

Exceptions to defaults



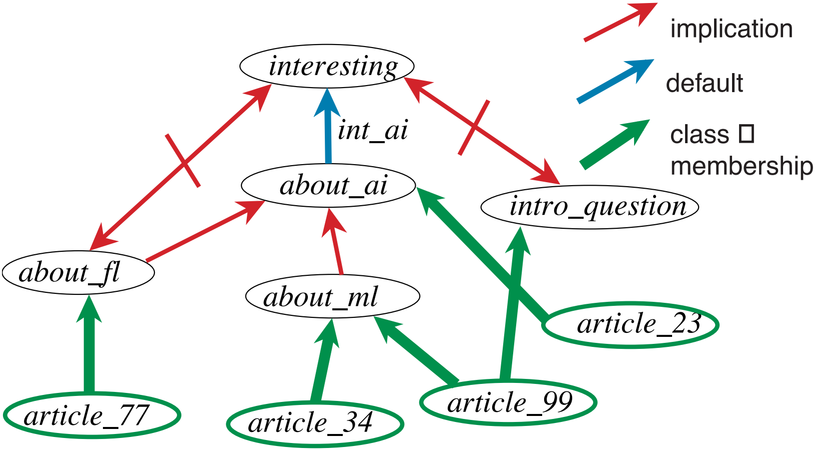
Exceptions to Defaults

"Articles about formal logic are about AI."
"Articles about formal logic are uninteresting."
"Articles about machine learning are about AI."

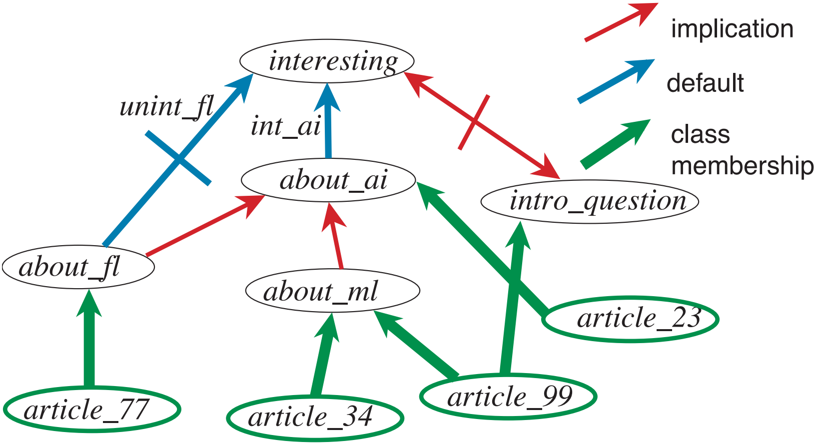
$$about\_ai \leftarrow about\_fl.$$
$$uninteresting \leftarrow about\_fl.$$
$$about\_ai \leftarrow about\_ml.$$
$$interesting \leftarrow about\_ai \wedge int\_ai.$$
$$false \leftarrow interesting \wedge uninteresting.$$
$$false \leftarrow intro\_question \wedge interesting.$$

Given $about\_fl$, is there explanation for $interesting$?
Given $about\_ml$, is there explanation for $interesting$?

Exceptions to Defaults



Formal logic is uninteresting by default



Contradictory Explanations

Suppose formal logic articles aren't interesting *by default*:

$$H = \{unint\_fl, int\_ai\}$$

The corresponding facts are:

$$interesting \leftarrow about\_ai \land int\_ai.$$
$$about\_ai \leftarrow about\_fl.$$
$$uninteresting \leftarrow about\_fl \land unint\_fl.$$
$$false \leftarrow interesting \land uninteresting.$$
$$about\_fl.$$

Does *uninteresting* have an explanation?
Does *interesting* have an explanation?

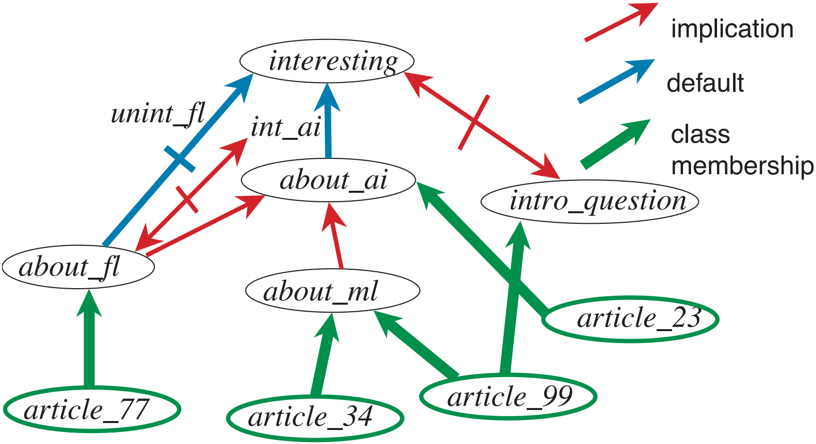
Overriding Assumptions

- For an article about formal logic, the argument "it is interesting because it is about AI" shouldn't be applicable.
- This is an instance of preference for more specific defaults.
- Arguments that articles about formal logic are interesting because they are about AI can be defeated by adding:

  $$false \leftarrow about\_fl \wedge int\_ai.$$

  This is known as a cancellation rule.
- We can no longer explain *interesting*.

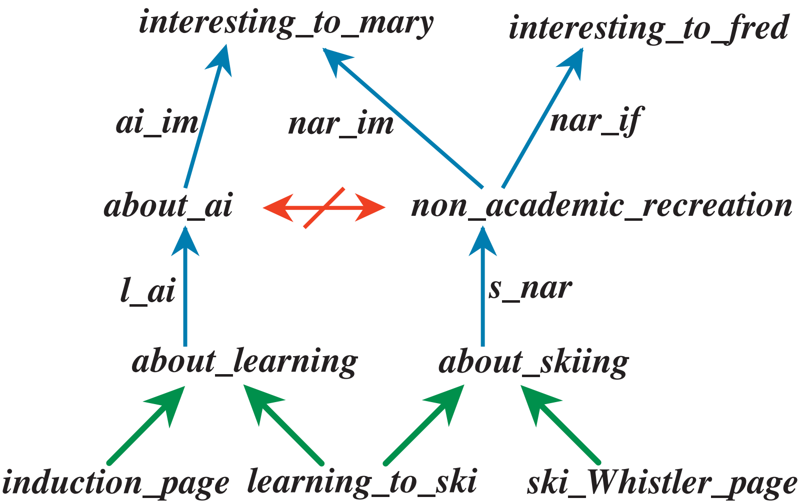
Diagram of the Default Example



Multiple Extension Problem

- What if incompatible goals can be explained and there are no cancellation rules applicable?
  What should we predict?
- For example: what if introductory questions are uninteresting, by default?
- This is the multiple extension problem.
- Recall: an extension of $\langle F, H \rangle$ is the set of logical consequences of F and a maximal scenario of $\langle F, H \rangle$.

Competing Arguments



Skeptical Default Prediction

- We predict g if g is in all extensions of $\langle F, H \rangle$.
- Suppose g isn't in extension E. As far as we are concerned E could be the correct view of the world. So we shouldn't predict g.
- If g is in all extensions, then no matter which extension turns out to be true, we still have g true.
- Thus g is predicted even if an adversary gets to select assumptions, as long as the adversary is forced to select something. You do not predict g if the adversary can pick assumptions from which g can't be explained.

Minimal Models Semantics for Prediction

Recall: logical consequence is defined as truth in all models.
We can define default prediction as truth in all minimal models.

Suppose M_1 and M_2 are models of the facts.
$M_1 <_H M_2$ if the hypotheses violated by M_1 are a strict subset of the hypotheses violated by M_2. That is:

$$\{h \in H' : h \text{ is false in } M_1\} \subset \{h \in H' : h \text{ is false in } M_2\}$$

where H' is the set of ground instances of elements of H.

Minimal Models and Minimal Entailment

- M is a minimal model of F with respect to H if M is a model of F and there is no model M_1 of F such that $M_1 <_H M$.
- g is minimally entailed from $\langle F, H \rangle$ if g is true in all minimal models of F with respect to H.
- Theorem: g is minimally entailed from $\langle F, H \rangle$ if and only if g is in all extensions of $\langle F, H \rangle$.