Default Reasoning

- When giving information, you don’t want to enumerate all of the exceptions, even if you could think of them all.
- In default reasoning, you specify general knowledge and modularly add exceptions. The general knowledge is used for cases you 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 you add that something is exceptional, you can’t conclude what you 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 $g$. 
A reader of newsgroups may have a default: “Articles about AI are generally interesting”.

\[ H = \{ \text{int}_\text{ai}(X) \}, \]

where \( \text{int}_\text{ai}(X) \) means \( X \) is interesting if it is about AI. With facts:

\[ \text{interesting}(X) \leftarrow \text{about}_\text{ai}(X) \land \text{int}_\text{ai}(X). \]

\[ \text{about}_\text{ai}(\text{art}_{23}). \]

\( \{ \text{int}_\text{ai}(\text{art}_{23}) \} \) is an explanation for \( \text{interesting}(\text{art}_{23}) \).
We can have exceptions to defaults:

\[ \text{false} \leftarrow \text{interesting}(X) \land \text{uninteresting}(X). \]

Suppose article 53 is about AI but is uninteresting:

\[ \text{about}_\text{ai}(\text{art}_53). \]
\[ \text{uninteresting}(\text{art}_53). \]

We cannot explain \( \text{interesting}(\text{art}_53) \) even though everything we know about \( \text{art}_23 \) you also know about \( \text{art}_53 \).
Exceptions to defaults

Interesting article about AI implication default class membership

Uninteresting article

Implication default class membership

Article 23

Article 53
Exceptions to Defaults

“Articles about formal logic are about AI.”
“Articles about formal logic are uninteresting.”
“Articles about machine learning are about AI.”

\[
\begin{align*}
\text{about}\_\text{ai}(X) & \leftarrow \text{about}\_\text{fl}(X). \\
\text{uninteresting}(X) & \leftarrow \text{about}\_\text{fl}(X). \\
\text{about}\_\text{ai}(X) & \leftarrow \text{about}\_\text{ml}(X). \\
\text{about}\_\text{fl}(\text{art}\_77). \\
\text{about}\_\text{ml}(\text{art}\_34). \\
\end{align*}
\]

You can’t explain \text{interesting}(\text{art}\_77).
You can explain \text{interesting}(\text{art}\_34).
Exceptions to Defaults

about_fl → about_ai → int_ai → interesting

intro_question → article_23

about_ml → article_34

about_ai → article_99

article_77

implication
default
class
membership
Formal logic is uninteresting by default

- `about_fl`
- `int_ai`
- `about_ai`
- `intro_question`
- `article_77`
- `article_34`
- `article_99`
- `implication`
- `default`
- `class membership`
- `unint_fl`
Contradictory Explanations

Suppose formal logic articles aren’t interesting by default:

\[ H = \{ \text{unint}_\text{fl}(X), \text{int}_\text{ai}(X) \} \]

The corresponding facts are:

\[
\begin{align*}
\text{interesting}(X) & \leftarrow \text{about}_\text{ai}(X) \land \text{int}_\text{ai}(X). \\
\text{about}_\text{ai}(X) & \leftarrow \text{about}_\text{fl}(X). \\
\text{uninteresting}(X) & \leftarrow \text{about}_\text{fl}(X) \land \text{unint}_\text{fl}(X). \\
\text{about}_\text{fl}(\text{art}_77). \\
\text{uninteresting} (\text{art}_77) & \text{ has explanation } \{ \text{unint}_\text{fl}(\text{art}_77) \}. \\
\text{interesting} (\text{art}_77) & \text{ has explanation } \{ \text{int}_\text{ai}(\text{art}_77) \}.
\end{align*}
\]
Overriding Assumptions

Because `art_77` is about formal logic, the argument "`art_77` 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:
\[
\text{false} \leftarrow \text{about}_{-}\text{fl}(X) \land \text{int}_{-}\text{ai}(X).
\]
This is known as a cancellation rule.

You can no longer explain `interesting(art_77)`.
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 \( \langle F, H \rangle \).
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.
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 $\langle F, H \rangle$. 