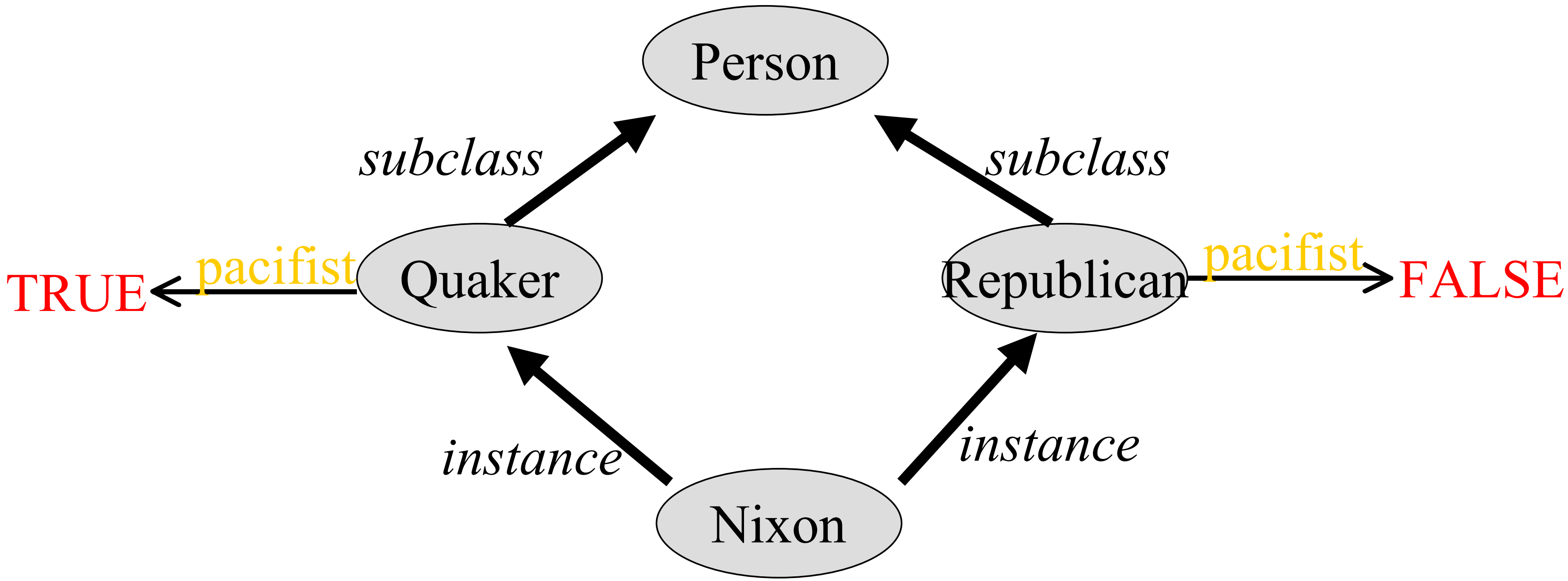
SMART Systems (Vorlesung: KI & XPS)

Ralf Möller, Univ. of Applied Sciences, FH-Wedel

- Beim vorigen Mal:
  - Handlungsplanung
- Inhalt heute:
  - Annahmen-basiertes Schließen
- Lernziele:
  - Default-Schließen und Abduktion
  - Anwendungsbereiche

Nixon Diamond

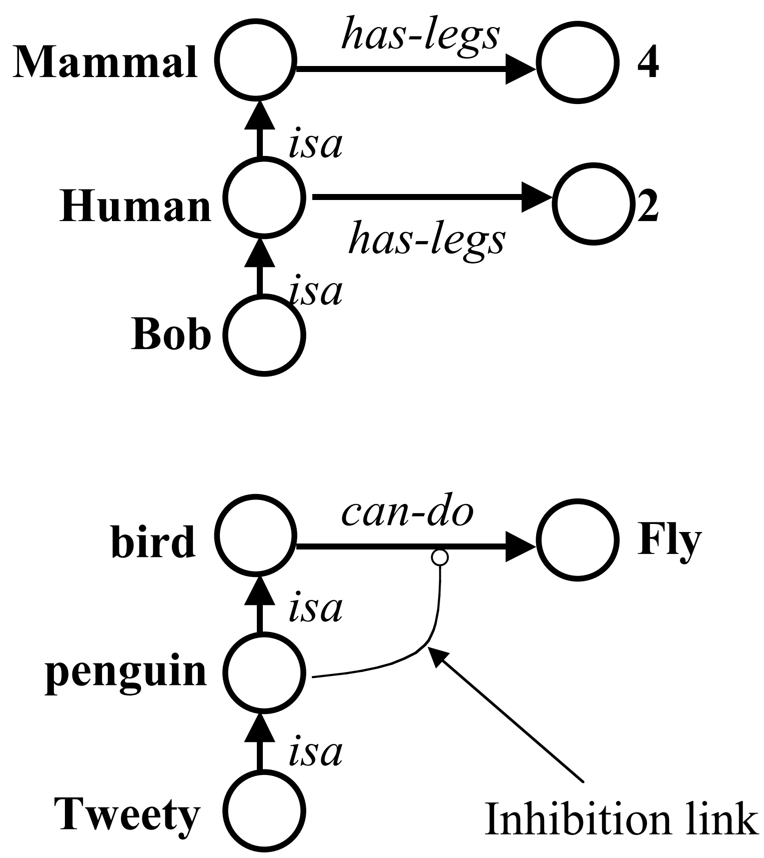
- This was the classic example circa 1980.



http://www.csee.umbc.edu/~ypeng/471-671.html

Exceptions in ISA hierarchy

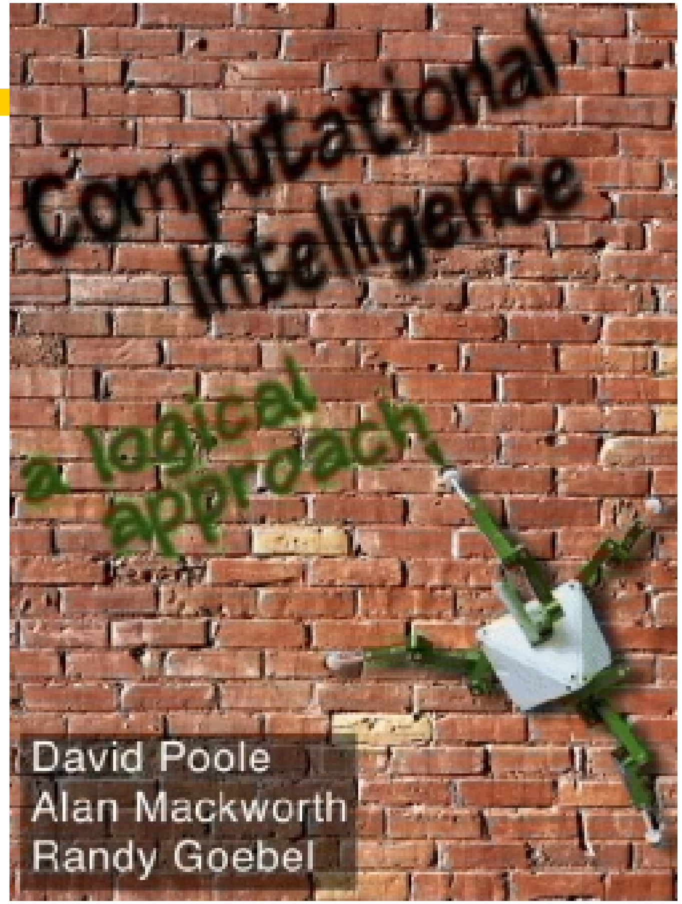
- Properties of a class are often default in nature (there are exceptions to these associations for some subclasses/instances)
  - Closer ancestors (more specific) overriding far way ones (more general)
  - Use explicit inhibition links to prevent inheriting some properties



http://www.csee.umbc.edu/~ypeng/471-671.html

Acknowledgments

- Slides taken from
- Computational Intelligence - A Logical Approach, David Poole, Alan Mackworth, Randy Goebel
- Oxford University Press, New York.
- http://www.cs.ubc.ca/spider/poole/ci.html



Assumption-based Reasoning

Often we want our agents to make assumptions rather than doing deduction from their knowledge. For example:

- In default reasoning the delivery robot may want to assume Mary is in her office, even if it isn't always true.
- In diagnosis you hypothesize what could be wrong with a system to produce the observed symptoms.
- In design you hypothesize components that provably fulfill some design goals and are feasible.

Design and Recognition

Two different tasks use assumption-based reasoning:

- Design The aim is to design an artifact or plan. The designer can select whichever design they like that satisfies the design criteria.
- Recognition The aim is to find out what is true based on observations. If there are a number of possibilities, the recognizer can't select the one they like best. The underlying reality is fixed; the aim is to find out what it is.

Compare: Recognizing a disease with designing a treatment. Designing a meeting time with determining when it is.

The Assumption-based Framework

The assumption-based framework is defined in terms of two sets of formulae:

- F is a set of closed formulae called the facts.
  These are formulae that are given as true in the world.
  We assume F are Horn clauses.

- H is a set of formulae called the possible hypotheses or assumables. Ground instances of the possible hypotheses can be assumed if consistent.

Making Assumptions

- A scenario of $\langle F, H\rangle$ is a set D of ground instances of elements of H such that $F \cup D$ is satisfiable.

- An explanation of g from $\langle F, H\rangle$ is a scenario that, together with F, implies g.

  D is an explanation of g if $F \cup D \models g$ and $F \cup D \not\models \mathit{false}$.

  A minimal explanation is an explanation such that no strict subset is also an explanation.

- An extension of $\langle F, H\rangle$ is the set of logical consequences of F and a maximal scenario of $\langle F, H\rangle$.

Default Reasoning and Abduction

There are two strategies for using the assumption-based framework:

- Default reasoning Where the truth of g is unknown and is to be determined.
  An explanation for g corresponds to an argument for g.
- Abduction Where g is given, and we are interested in explaining it. g could be an observation in a recognition task or a design goal in a design task.

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 determines 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".

$$H = \{int\_ai(X)\},$$

where $int\_ai(X)$ means X is interesting if it is about AI.

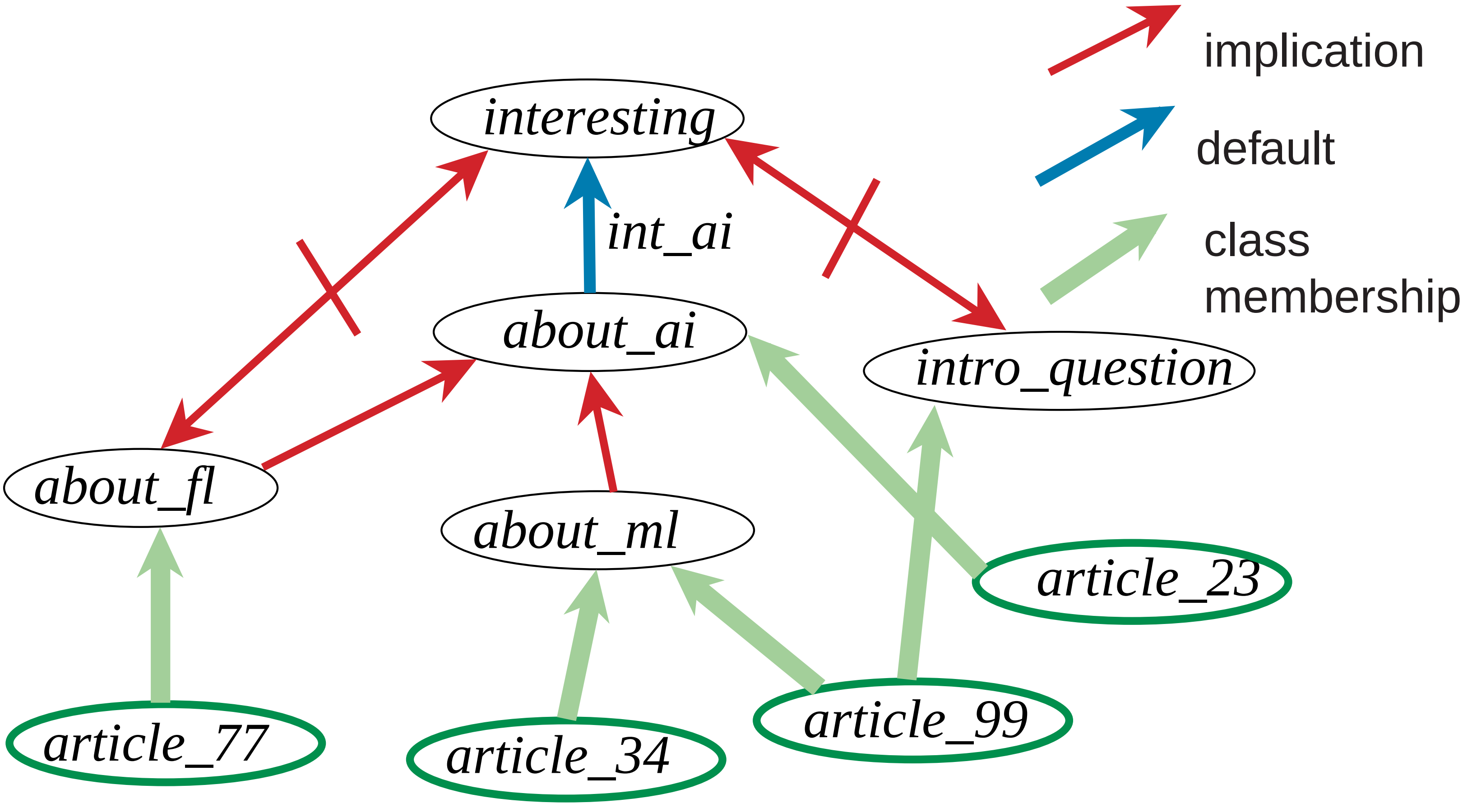
With facts:

$$interesting(X) \leftarrow about\_ai(X) \wedge int\_ai(X).$$

$$about\_ai(art\_23).$$

$\{int\_ai(art\_23)\}$ is an explanation for $interesting(art\_23)$.

Diagram of the Default Example



Exceptions to Defaults

"Articles about formal logic are about AI."

"Articles about formal logic are uninteresting."

"Articles about machine learning are about AI."

$$about\_ai(X) \leftarrow about\_fl(X).$$

$$false \leftarrow about\_fl(X) \wedge interesting(X).$$

$$about\_ai(X) \leftarrow about\_ml(X).$$

$$about\_fl(art\_77).$$

$$about\_ml(art\_34).$$

You can't explain $interesting(art\_77)$.

You can explain $interesting(art\_34)$.

Contradictory Explanations

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

$$H = \{unint\_fl(X), int\_ai(X)\}$$

The corresponding facts are:

$$interesting(X) \leftarrow about\_ai(X) \wedge int\_ai(X).$$

$$about\_ai(X) \leftarrow about\_fl(X).$$

$$false \leftarrow about\_fl(X) \wedge unint\_fl(X) \wedge interesting(X).$$

$$about\_fl(art\_77).$$

$\neg interesting(art\_77)$ has explanation $\{unint\_fl(art\_77)\}$.

$interesting(art\_77)$ has explanation $\{int\_ai(art\_77)\}$.

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 the fact:

$$false \leftarrow about\_fl(X) \wedge int\_ai(X).$$

This is known as a cancellation rule.

With this fact, 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?

- 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 $\langle F, H \rangle$.

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 $\langle F, H \rangle$.

Abduction

Abduction is an assumption-based reasoning strategy where

- H is a set of assumptions about what could be happening in a system
- F axiomatizes how a system works
- g to be explained is an observation or a design goal

Example: in **diagnosis** of a physical system:
H contain possible faults and assumptions of normality,
F contains a model of how faults manifest themselves
g is conjunction of symptoms.

Abduction versus Default Reasoning

Abduction differs from default reasoning in that:

- We don’t care if $\neg g$ can also been explained.
- It is the explanations that are of interest, not just the conclusion.
- H contains abnormality as well as normality assumptions.
- We don’t want to only explain normal outcomes; often we want to explain why some abnormal observation occurred.

Example of User Modeling

Suppose a n infobot wants to determine what a user is interested in. We can hypothesize the interests of users:

$$H = \{interested\_in(Ag, Topic)\}.$$

Suppose the corresponding facts are:

$$selects(Ag, Art) \leftarrow$$
$$about(Art, Topic) \wedge$$
$$interested\_in(Ag, Topic).$$

$$about(art\_94, ai).$$

$$about(art\_94, info\_highway).$$

$$about(art\_34, ai). \qquad about(art\_34, skiing).$$

User Modeling Example: explanations

There are two minimal explanations of $selects(fred, art\_94)$:

$\{interested\_in(fred, ai)\}$.

$\{interested\_in(fred, information\_highway)\}$.

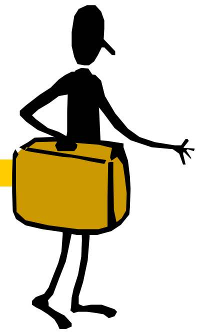
If you observe $selects(fred, art\_94) \wedge selects(fred, art\_34)$, there are two minimal explanations:

$\{interested\_in(fred, ai)\}$.

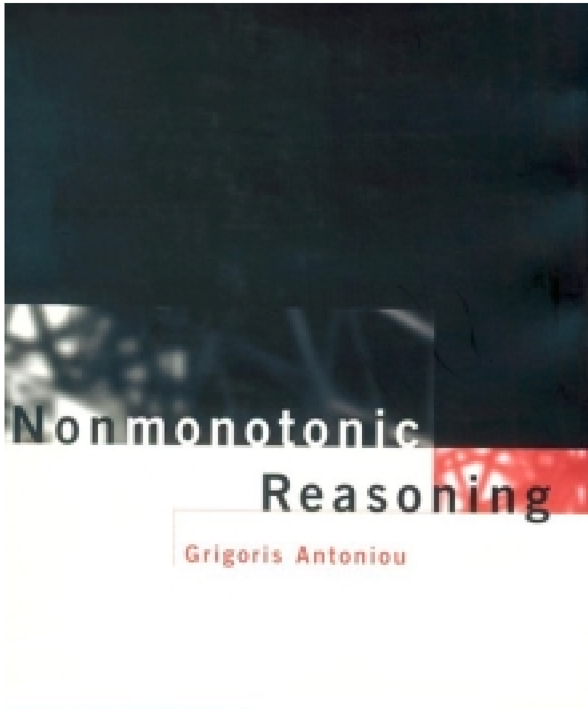
$\{interested\_in(fred, information\_highway),$

$interested\_in(fred, skiing)\}$.

Zusammenfassung, Kernpunkte

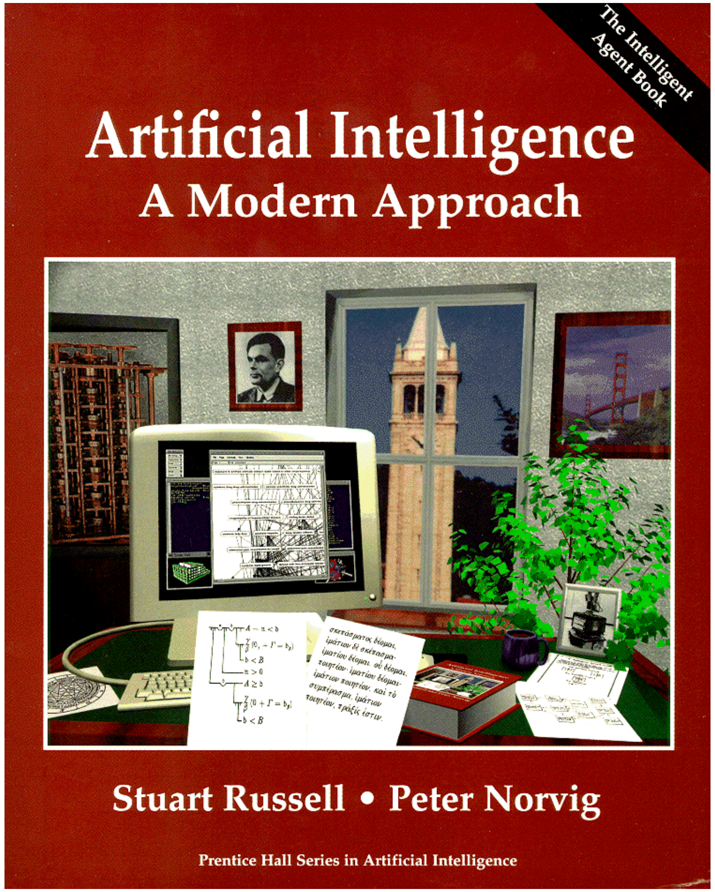


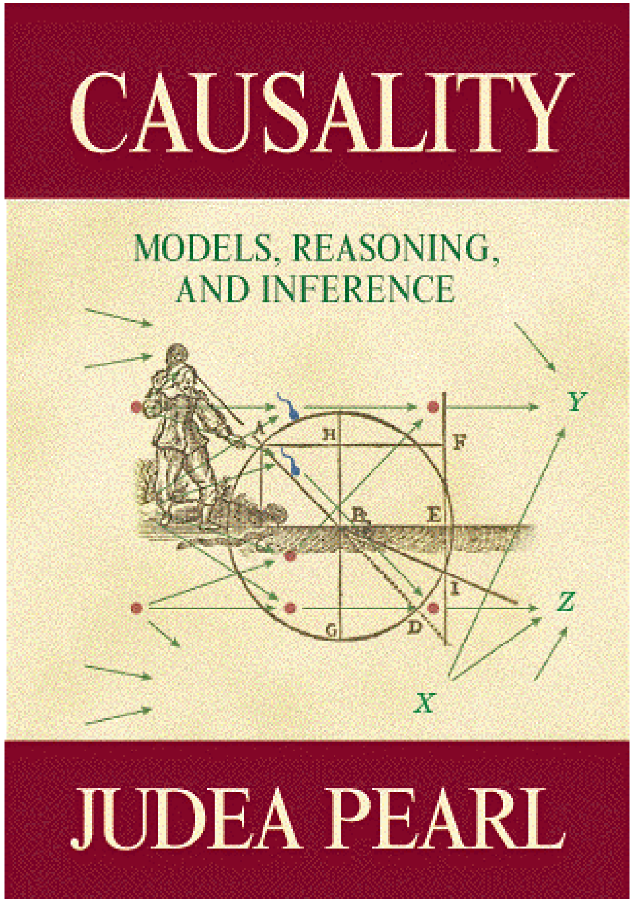
- Default-Schließen und Abduktion
- Anwendungsbereiche



Weitere Literatur zum Thema dieser Vorlesung

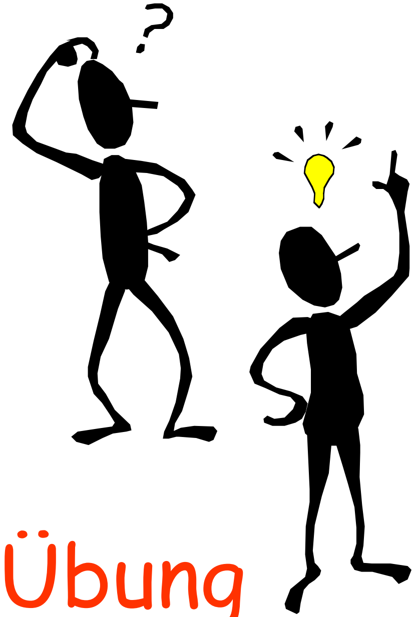
Vertiefungen und Ergänzungen für die gesamte Veranstaltung



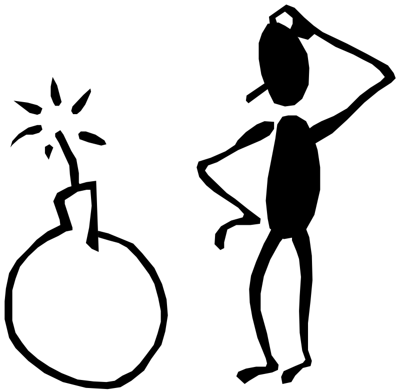


Das war's ... (fast)

Vorlesung



Übung



Klausur