

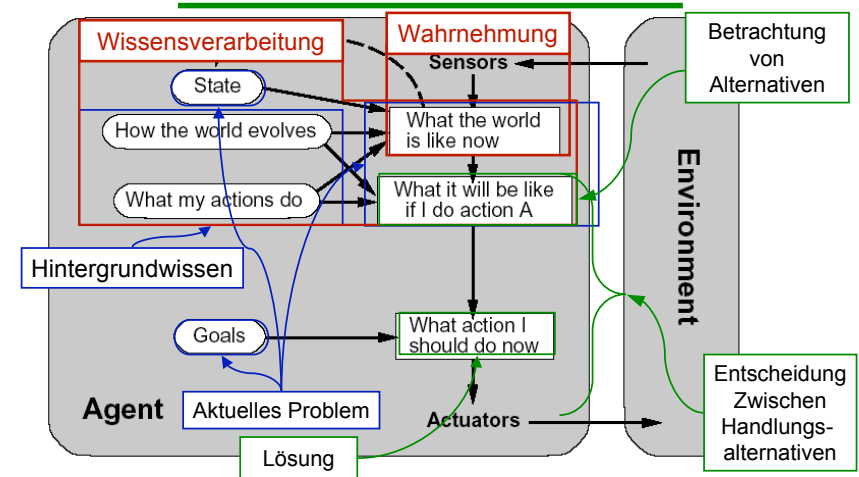
Wissensrepräsentation

Christopher Habel, Özgür Özçep
Sommersemester 2005

Sitzung 13: Schliessen unter Unsicherheit

- Wissensbasierte Agenten: Wissenslücken, Wissensmängel
- Deduktive vs. Nicht-deduktive Wissensverarbeitung
- Vererbung & Vererbungsnetze: Motivation
- Schliessen in Vererbungsnetzen

Goal-based Agent: Sicheres vs. unsicheres Wissen

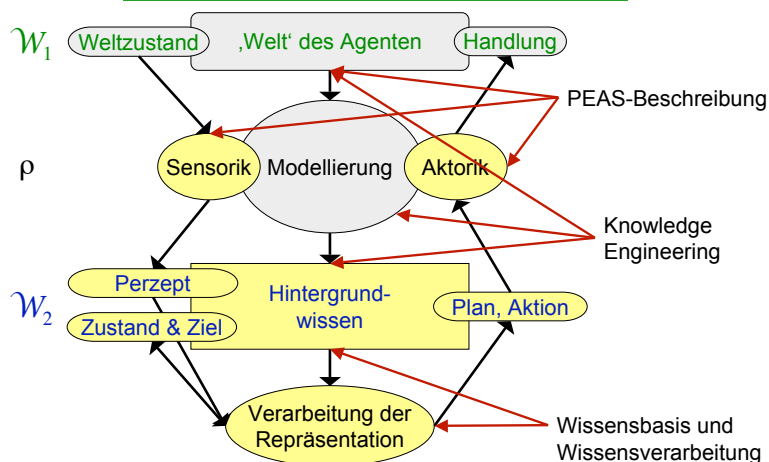


Wissensrepräsentation, SoSe 2005
Ch. Habel / C. Eschenbach

13 – 2

Russell & Norvig - ch.2, f.21

Wissensbasierter Agent



Wissensrepräsentation, SoSe 2005
Ch. Habel / C. Eschenbach

13 – 3

Entscheidung zwischen Handlungsalternativen

- Wenn mehrere Handlungen möglich sind, muss der Agent zwischen den Handlungen entscheiden.

- Die zentralen Fragen:
Was ist zu tun, wenn das Wissen

- unvollständig
- unpräzise / ungenau
- inkorrekt / unsicher

ist?

Aber, wann ist das Wissen des Agenten schon mal vollständig, präzise und sicher?

Wissensrepräsentation, SoSe 2005
Ch. Habel / C. Eschenbach

13 – 4

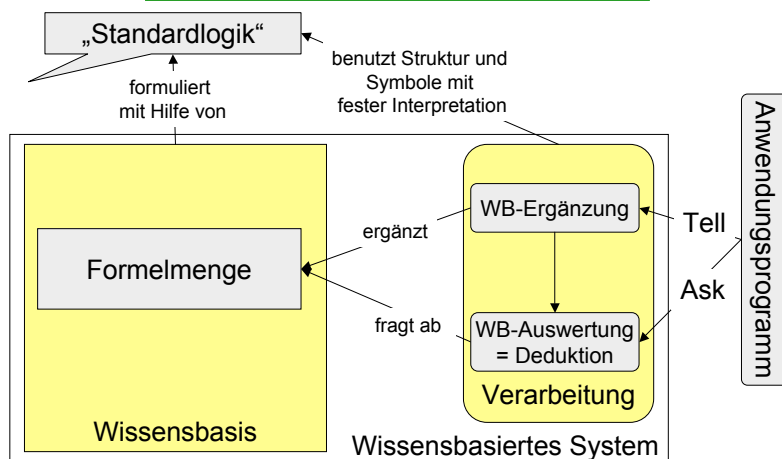
Sie beabsichtigen eine Wochenendreise zu Freunden in Berlin.

Die Aufgaben (Varianten: Mit der Bahn / mit dem Auto.)

- Routen- und Zeitplanung (Bahn / Auto)
- Entscheidung / Wahl des Verkehrsmittels
- Planmodifikation: Was ist zu tun, wenn bei der Planausführung Probleme auftreten? (Verspätung, Stau, Umleitung)
 - Kriterien für die Planung / Entscheidung
 - Wie vollständig, präzise, sicher muss dieses Wissen sein?
 - Wie konstant oder veränderlich ist das einzusetzende Wissen?

- Welche Entscheidungen sind in der Wumpus-Welt zu treffen?
- Wie gut oder schlecht kann die „Wissenslage“ sein?
 - Vollständigkeit, Korrektheit, Genauigkeit
- Welche Handlungsoptionen gibt es, wenn die Wissenslage nicht gut ist? (Risiko?)
- Wie verändern sich die Antworten auf die obigen Fragen, wenn
 - Die Sensorik nicht präzise / „diskret“ ist?
 - Die Welt nicht statisch ist? (Der Wumpus bewegt sich.)
→ Update der Wissensbasis

Wissensbasiertes System
mit deduktiver Wissensbasis



Wissensbasen und Folgerung

Entailment means that one thing follows from another:

$$KB \models \alpha$$

Knowledge base KB entails sentence α if and only if

α is true in all worlds where KB is true

- Wenn KB die Welt korrekt beschreibt
- und $KB \models \alpha$,
- dann beschreiben auch α und $KB \cup \{\alpha\}$ die Welt korrekt.
- Ergänzung von α zerstört bestehende Rep.-Beziehungen nicht

Aus der Perspektive eines Agenten

- Wenn $KB \models \alpha$,
- dann ist der Agent mit KB berechtigt, auf α zu schließen.

2.1 Deductive Reasoning Agents

- How can an agent decide what to do using theorem proving?
- Basic idea is to use logic to encode a theory stating the *best* action to perform in any given situation.
- Let:
 - ρ be this theory (typically a set of rules);
 - Δ be a logical database that describes the current state of the world;
 - Ac be the set of actions the agent can perform;
 - $\Delta \vdash_{\rho} \phi$ mean that ϕ can be proved from Δ using ρ .

<http://www.csc.11v.ac.uk/~mjw/pubs/imas/>

7

Deduktive Wissenbasierte Systeme (PL) Update

Folgerbarkeit – Ableitbarkeit in einem deduktiven wissensbasierten System (PL-basiert)

$$KB \models X \text{ gdw. } KB \vdash X$$

Monotonie-Eigenschaft deduktiver Logiken:

$$\text{Wenn } KB \models X, \text{ dann } KB \cup \{U\} \models X$$

➤ Update / Modifikation

- Wenn ein Agent die Wissensentität X aus der Wissensbasis KB erschliessen kann, so kann er nach der **Erweiterung von KB** um U X auch aus der erweiterten Wissensbasis erschliessen.
- Für wissensbasierte Systeme sind ausser Erweiterungen (Expansionen), andere Update-Operationen notwendig.

```

/* try to find an action explicitly prescribed */
for each  $a \in Ac$  do
  if  $\Delta \vdash_{\rho} Do(a)$  then
    return  $a$ 
  end-if
end-for
/* try to find an action not excluded */
for each  $a \in Ac$  do
  if  $\Delta \not\vdash_{\rho} \neg Do(a)$  then
    return  $a$ 
  end-if
end-for
return null /* no action found */

```

Entscheidbarkeit ?

Was, wenn $Do(a)$ und $Do(b)$
($a \neq b$) ableitbar sind ?

Was bedeutet $Do(a)$?

<http://www.csc.11v.ac.uk/~mjw/pubs/imas/>

8

Deduktive Wissenbasierte Systeme (PL) Revision

Die Revisionsituation

t_0 : Wissensbasis: KB_0

aktuell zu berücksichtigende Wissensentitäten: U

t_1 : Wissensbasis: $KB_1 := KB_0 \oplus U$

mit folgenden Forderungen an KB_1 :

- KB_1 soll konsistent sein (wenn KB_0 konsistent ist)
- KB_1 soll sich nur soweit von KB_0 unterscheiden, wie es notwendig ist.
- U soll soweit in KB_1 berücksichtigt werden, wie es möglich ist.

Belief Revision / Truth Maintenance
setzt Entscheidungen voraus.

Deduktive Wissenbasierte Systeme (PL) Unsicherheit

Folgerbarkeit – Ableitbarkeit in einem deduktiven wissensbasierten System (PL-basiert)

$$KB \models X \text{ gdw. } KB \vdash X$$

(Bei Vorliegen von Vollständigkeit und Korrektheit)

Wenn die Welt ein Modell der Wissensbasis ist, dann kann sich der Agent darauf verlassen, dass eine abgeleitete Wissensentität X in der Welt erfüllt ist.

Unsicherheit

(Charakterisierung im Hinblick auf Folgerbarkeit / Ableitbarkeit):

- Y soll auf Erschliessbarkeit aus KB geprüft werden, und es ergibt sich:

$$KB \not\models Y \text{ und } KB \not\models \neg Y$$

Nichtdeduktive Wissenbasierte Systeme (PL) Unsicherheit

Folgerbarkeit – Ableitbarkeit in einem nicht-deduktiven wissensbasierten System

$$KB \vdash_{nd} X \text{ setzt nicht voraus, dass } KB \models X$$

Unsicherheit betrifft jetzt

- Im konkreten Fall, also für X , den Zusammenhang zwischen \models und \vdash_{nd}
- Die „Interpretation“ von \vdash_{nd} , z.B.
 - „wahrscheinlich X “ ist ableitbar“ bzw.
„ X ist mit gewisser Wahrscheinlichkeit ableitbar“
 - „ X ableitbar, falls nichts dagegen spricht“

Inference in logic

derivation of conclusions from given information or premises
by any acceptable form of reasoning.
Inferences are commonly drawn

by **deduction**, which, by analyzing valid argument forms, draws out the conclusions implicit in their premises,

by **induction**, which argues from many instances to a general statement,

by **probability**, which passes from frequencies within a known domain to conclusions of stated likelihood, and

by **statistical reasoning**, which concludes that, on the average, a certain percentage of a set of entities will satisfy the stated conditions.

Entscheidung unter Unsicherheit

Agenten müssen, um handlungsfähig zu sein, auch dann Entscheidungen treffen können, wenn ihr Wissen – im Hinblick auf die Aufgabenstellung – unvollständig, unpräzise / ungenau, (vielleicht) inkorrekt ist.

Gründe für „unsicheres Wissen“

- Fehlende Information über die Umgebung bzw. über Veränderungen der Umgebung
- Qualität der Information (Wahrnehmung, Kommunikation)
- Wissensverarbeitung (z.B. Schliessen unter beschränkten Ressourcen)

Verarbeitung unsicheren Wissens: Agenda

Definition von

Repräsentations- & Verarbeitungssystemen

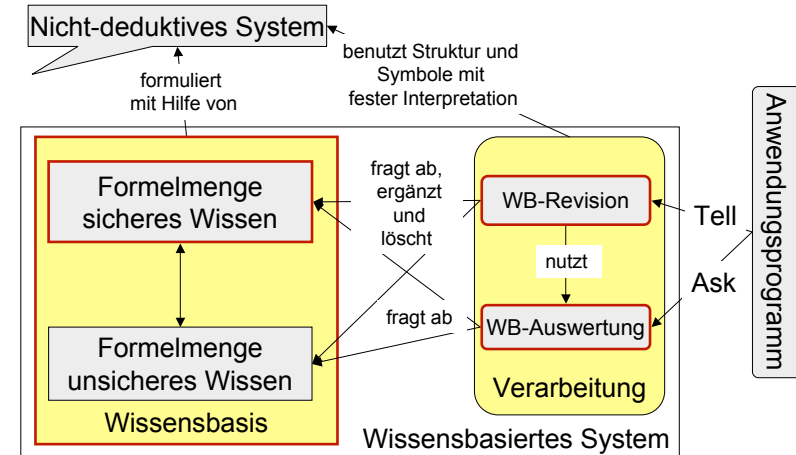
Generelles Schema (s. LOS)

- eine formale Sprache (zur Repräsentation)
- Evaluations- / Interpretationsprinzipien
- semantische Kategorisierungen und Beziehungen
- Ableitungs-, Beweisverfahren

➤ In manchen Systemen der KI sind nicht alle dieser Gesichtspunkte realisiert.

Z.B. existieren zu einigen Ansätzen des „Nicht-monotonen Schliessens“ keine ausgearbeiteten Semantikkonzeptionen (→ semantische Kategorisierungen und Beziehungen)

Wissensbasiertes System mit nicht-deduktiver Wissensbasis



Vererbung (Inheritance)

When thinking object-centered, hierarchy is a natural way to view the world

- importance of abstraction in remembering and reasoning
 - groups of things share properties in the world (e.g., mammals, seafood)
 - don't have to repeat representations (e.g., sufficient to say that "elephants are mammals" to know a lot about them)

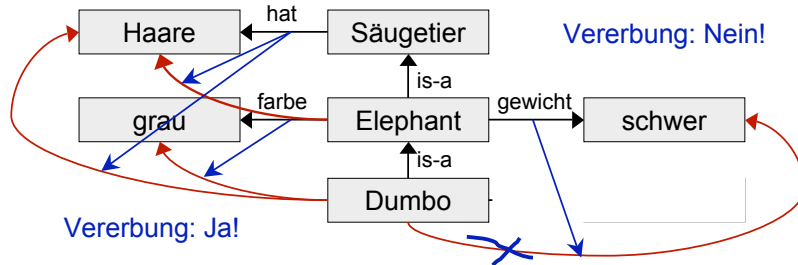
z.B. Frames, Semantische Netze, Beschreibungslogik (description logics)

Literatur

- Brachman, R. J. & Levesque, H. J. (2004). Knowledge Representation and Reasoning. San Francisco, CA: Morgan Kaufmann. (Kapitel 10)
- D. Gabbay, C. Hogger, and J. Robinson (eds.) (1994). *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume 3: Nonmonotonic Reasoning and Uncertain Reasoning*, Oxford University Press.
- Darin insbesondere:
- J. Horty. Some direct theories of nonmonotonic inheritance. In D. Gabbay, C. Hogger, and J. Robinson (eds.), Oxford University Press (1994), pp. 111 - 187.
- D. Touretzky. (1986). The mathematics of inheritance systems. Morgan Kaufmann, Los Altos, Ca., 1986.
- L. A. Stein (1992). Resolving Ambiguity in Nonmonotonic Inheritance Hierarchies, *Artificial Intelligence* 55 (2-3): 259-310.

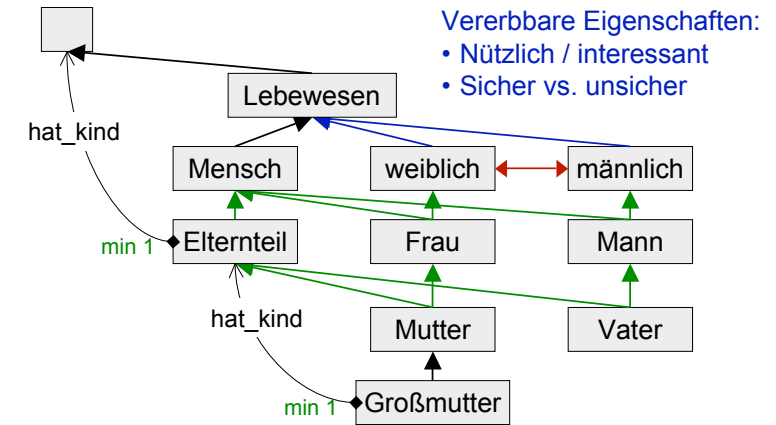
Semantische Netze

- Repräsentationen für Konzept-Wissen
- Graphen mit Knoten für Konzepte und Individuen und beschriftete Pfeilen für binäre Relationen
- Definiert über Bilder, keine klare Semantik, keine spezifizierten Verarbeitungsmechanismen



Beispiel: Konzeptsysteme

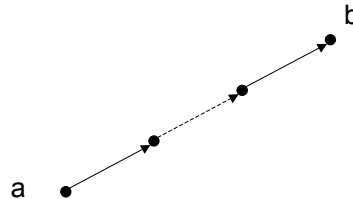
Gruppendiskussion



Inheritance

“Inheritance” is the result of transitivity reasoning over paths

- for strict networks, *modus ponens* in graphical form
- does *a* inherit from *b*?
≡ is *b* in the transitive closure of IS-A (or subsumption) from *a*?

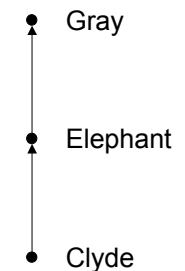


Pfad-basiertes Schliessen

Focus just on

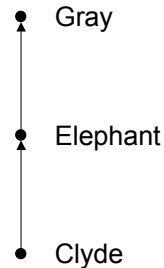
inheritance and transitivity

- many interesting considerations in looking just at **where information comes from** in a network representation
- abstract descriptions, and properties into *nodes* in graphs, and just look at **reasoning with paths** and the conclusions they lead us to



Pfad-basiertes Schliessen (cont'd)

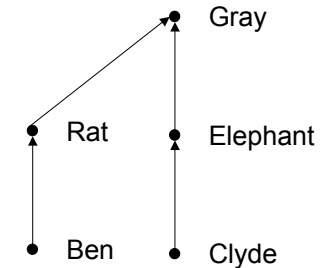
- edges in the network:
Clyde—Elephant, Elephant—Gray
- paths included in this network:
{Clyde—Elephant—Gray},
plus edges
 - in general, a path is a sequence of 1 or more edges
- conclusions supported by the paths:
Clyde → Elephant; Elephant → Gray;
Clyde → Gray



Vererbungsnetze (1)

Strict inheritance in trees

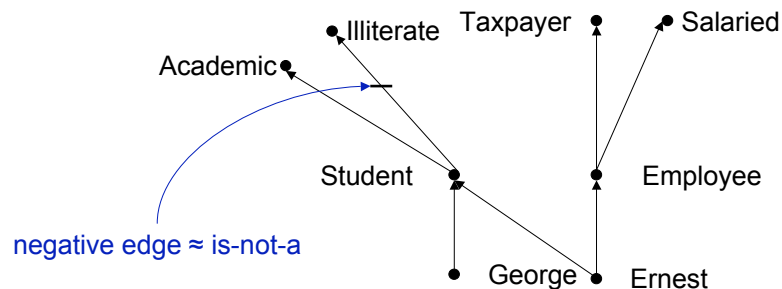
- as in description logics
- conclusions produced by complete transitive closure on all paths (any traversal procedure will do);
all reachable nodes are implied



Vererbungsnetze (2)

Strict inheritance in DAGs

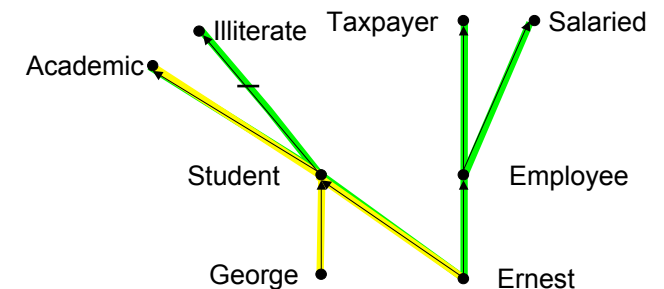
- as in DL's with multiple AND parents
- same as above: all conclusions you can reach by any paths are supported



Schliessen in Vererbungsnetzen: Upward vs. Downward Reasoning

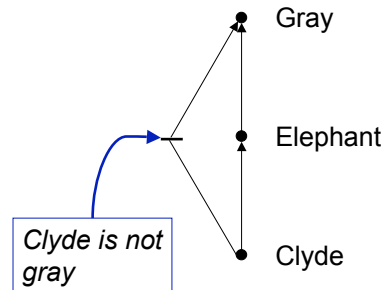
Was ist erschliessbar über:

- Academics ?
- Ernest



Defeasible inheritance (1)

- inherited properties do not always hold, and can be **overridden** (defeated)
- conclusions determined by searching upward from “**focus node**” and selecting first version of property you want



• defeasible

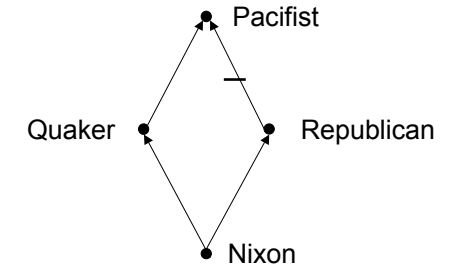
- **anfechtbar**
- **annullierbar**

• „Polarität“ von Kanten

- positiv is / is-a
- negative is-not / is-not-a

Defeasible inheritance (2): Nixon Diamond

- Mehrere Vererbungen möglich, die sich aber widersprechen können. (ambiguous net)
- Welche Eigenschaft soll hier zugesprochen werden?
- Welche Strategien für die Entscheidung zwischen den Vererbungsmöglichkeiten?



Defeasible Inheritance

Aufgabe zur Nachbereitung

Konzeptsystem der Wochentage etc.

- Werkstage, Samstag, Sonntag
- Sonn- und Feiertage
- Spezifische Feiertage:
 - Neujahr
 - Karfreitag, Ostern
 - Himmelfahrt
 - Pfingsten
 - 3. Oktober
 - Weihnachten, Sylvester

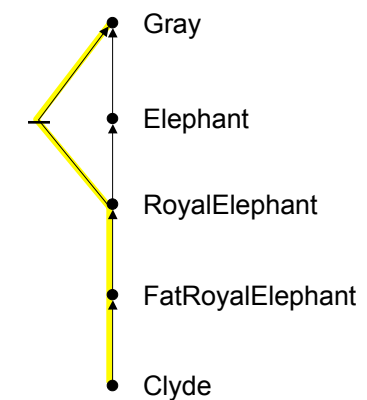
- Welche Vererbungs-Netze?
- Welche Strategien für Konflikte?

Shortest Path Heuristic

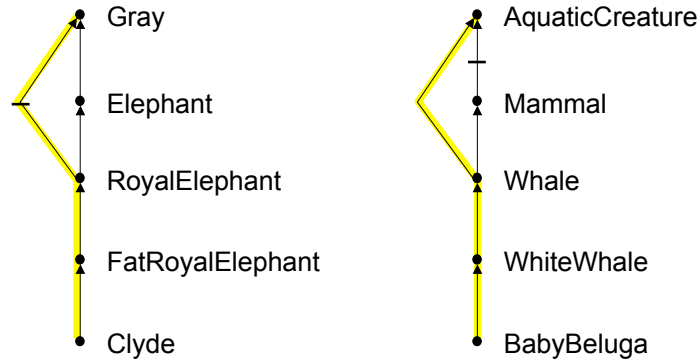
Defeasible inheritance in DAGs

- links have polarity (positive or negative)
- use **shortest path heuristic** to determine which polarity counts

- Zugrundeliegende Idee: Berücksichtige das spezifischste subsumierende Konzept.



Shortest Path Heuristic



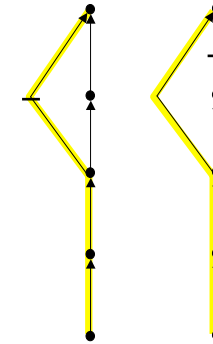
→ Clyde ist nicht grau.

→ BabyBeluga ist ein im Wasser lebendes Tier.

Shortest Path Heuristic

Pfade im Vererbungsnetzwerk

- fungieren als Argumente
- stützen Schlussfolgerungen
- Einige Pfade werden von anderen verhindert (are "preempted")
- Andere sind zulässig (are "admissible")



➤ Das Vererbungsproblem ist das Problem, die zulässigen Pfade zu bestimmen.

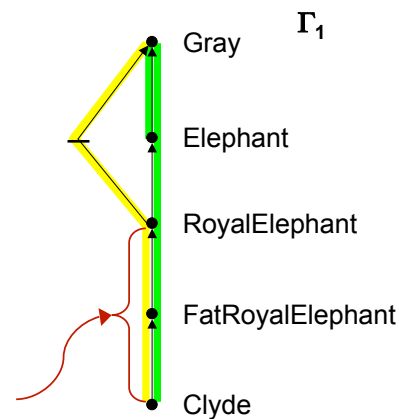
Pfade und Argumentationen

Pfade

- werden von einem Knoten ausgehend konstruiert
- sie repräsentieren Argumentationen

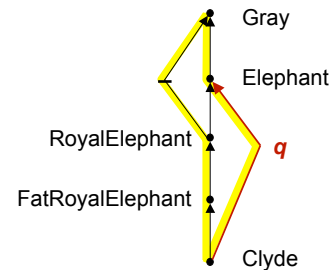
Die letzte Kante eines Pfades korrespondiert zu einem Grund, eine bestehende Argumentation fortzusetzen.

Verschiedene Argumentationen können gemeinsame Anteile haben.

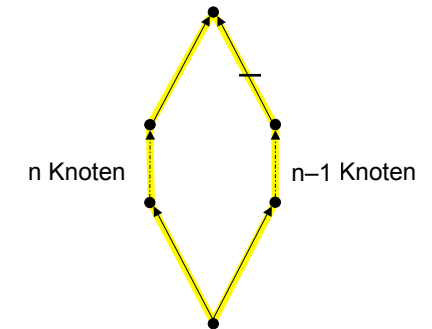


Probleme mit der Kürzeste-Wege-Heuristik

Redundante Kanten beeinflussen das Ergebnis, z.B. eine zusätzliche direkte Verbindung, etwa durch einen „gespeicherten Schluss“



Irrelevante Kanten können das Ergebnis beeinflussen, z.B. bei grosser Pfadlänge

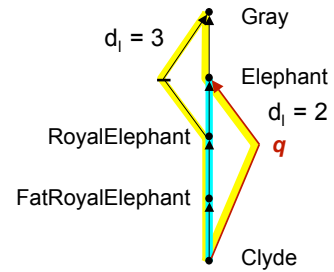


Inferential distance

Consider “inferential distance”: not linear distance, but

- topologically based – a node **a** is nearer to node **b** than to node **c** if there is a path from **a** to **c** through **b**

➤ idea: conclusions from **b** preempt those from **c**



Schliessen unter Unsicherheit – Vererbung und Verwandtes

Die nächsten Vorlesungen:

- Eine Formalisierung des Schliessens in Vererbungsnetzen (Lynn Stein)
 - Ansätze einer semantischen Fundierung
- Typen von Vererbung: strikte vs. nicht-strikte Vererbung
- Defaults (Raymond Reiter, David Poole)
 - generelle Konzeption
 - closed world assumption
 - default Logik(en)
 - Semantik
 - Beweistheorie