Revision **6.B**

Ontologie für Informationssysteme

Universität Leipzig

Wintersemester 2007/2008

Adrian Immanuel Kieß

13. September 2015

mailto:adrian@immanuelK.net
http://www.immanuelK.net

FRISCH Zum Seminar von Professor Herre. **Teil I.**

Eigene Mitschriften

Seminar vom 22.10.2007

Mitschrift ~~vom~~ zu Herre

- neu Informationen zum Seminar unter:
  - https://wiki.imise.uni-leipzig.de/2007WS/SOntIS

- Flut der Daten aus empirischen Wissenschaften
- Künstliche Intelligenz ⇒ Wissensrepräsentation
  - Wie kann ich das Wissen so repräsentieren, dass damit Lösungen gefunden werden können?
- Programmiersprachen
  - Entwicklungsetappen
    - *Prolog*: Deklarative Sprache. Keine interaktiven Programme, sondern logische Regeln werden aufgestellt.
  - Semantik
    - Viele Fragen
  - natürliche Sprachen formulieren?
  - hier kann Ontologie einen wesentlichen Beitrag leisten
  - Entwicklung eines stabilen Begriffssystemes
    - wie: Raum, Zeit, Objekt, Prozess, usw...

      ↷ daraus entstehen dann Wissensbasen
    - Einsetzen für Planung, Problemlösung...
    - immer auf Grundparadigma: *Wissensbasis*
- 1956
  - Versuche, auf künstliche Weise Problemlösungen zu simulieren

    ⇒ scheiterte
- 1965 ROBINSON

  Resolutionsbeweiser versucht möglichst wenig Schlussregeln zu verwenden, und kommt faktisch mit *einer* Schlussregel aus.

  ⇒ scheiterte auch!
  - ...da die Beweise plötzlich zu lang wurden.
- 1970

  ↷ Ansatz der künstlichen Intelligenz ⟶ falscher Ansatz
- 1960

  Expertensysteme tauchen auf
  - "Im Wissen liegt die Macht."
  - Paradigmenwechsel zu Fokus *Wissen*

        * Heute sagt man *Wissensbasierte Systeme*

- 1970 ⇝ heute
  - Hat sich bis heute bewährt
- 1985/1986

  CYC Größtes Projekt im Rahmen der künstlichen Intelligenz.

  ⟼ D. ~~LENART~~[1] LENAT

  Es sollte das Alltagswissen der Menschen weltweit in unzähligen Schemata über viele Jahre eingegeben und verknüpft werden. Hier geht es um die Frage, wie menschliches Hintergrundwissen, das im Expertensystemen immer ausgeblendet blieb, in Computerprogrammen zugänglich gemacht werden könnte. Das Wissenssystem CYC ist allerdings nicht lernfähig und bleibt auf die Zusammenhänge eingeschränkt, die ihm explizit eingegeben wurden.[2]

  - Begann mit c.a. 50.000.000,00$
    * es sollte
      · richtiger Durchbruch erzielt werden
      · so viel Wissen wie möglich angesammelt werden

      ⟼ Projekt gescheitert!
    * Warum?
      · Schwierig, so große Wissensbasen zu beherrschen

        ⇒ nicht mehr vernünftig zu organisieren
      · LENAT war der Meinung, dass alle logischen Probleme bereits gelöst seien

      ↺ Wichtige Probleme im CYC sind aber ungelöst!
- Benötigt wird zuerst ein einheitliches Begriffssystem — Raum, Zeit, Objekt, etc...
- ⇆ neuronale Netze ⇆ Sozialwissenschaften ⟼ Basisontologie? ⟼ mit fundamentalen Begriffen!

  ⇝ ONTOMED ⋙ GFO[3] (Generic Formal Ontology)

---

[1] Vielen herzlichen Dank an Martin C. für diesen wertvollen Hinweis.
[2] Quotation aus KI–Künstliche Intelligenz, Klaus Mainzer.
[3] *GFO*. Siehe: http://en.wikipedia.org/wiki/General_Formal_Ontology

Vorlesung vom 15.11.2007

- Zeit-Ontologien
  - $\vdash\dashv$ Chronoide[4]
  - Zeitgrenzen
  - Koinzidenz[5] von Grenzen ($\Rightarrow$ F. BRENTANO)
  - Zeitmodell einführen
    - Beispiele
      - ...
    - Eigenschaften sind zu einem Zeitpunkt präsent

$Chron(x)$ Einstelliges Prädikat, X ist ein Chronoid.

$tpart\,(x, y)$, $tcoine\,(x, y)$ x, y sind Grenzen, die koinzidieren,
$lb(x, y) :=$ x ist die linke Grenze von y,
$rb(x, y) :=$ x ist die rechte Grenze von y.

Definitionen

- $tppart(x, y) := tpart(x, y) \wedge x \neq y$

  (proper temporal part-of)
- $tb(x, y) := lb(x, y) \vee rb(x, y)$

  (x is a time boundary)

Definition: *Grenze* v. boundary

- $innerb(x, y) := \exists u(tppart(u, y) \wedge tb(x, u) \wedge \neg tb(x, y))$

  (x is an inner time-boundary of y)
- $Tb(x) := \exists y\, tb(x, y)$

  (x is a time-boundary)
- $t \circ v(x, y) := \exists z(tpart(z, x) \wedge tpart(z, y))$

  (temporal overlap of time-regions)
- $comp(x, y) := Chron(x) \wedge Chron(y) \wedge \exists z(Chron(z) \wedge tpart(x, z) \wedge tpart(y, z))$

  (x and y are compatible chronoids)

[4] https://wiki.imise.uni-leipzig.de/Themen/GOL/Chronoid

[5] Eine *Koinzidenz* ist das zeitliche, jedoch nicht notwendig räumliche Zusammentreffen von zwei oder mehr Ereignissen.

- $l(x) = y =_{df} lb(y, x)$

  (functional lb)

- $r(x) = y =_{df} rb\,(y, x)$

  (functional rb)

- $meets(x, y) =_{df} Chron(x) \wedge Chron(y) \wedge \exists uv\,(rb\,(u, x) \wedge lb\,(v, y) \wedge tcoinc\,(u, v))$

  (Allen's "meets" relation)

- $starts(x, y) =_{df} tppart\,(x, y) \wedge Chron\,(x) \wedge Chron\,(y) \wedge \exists u\,(lb\,(u, x) \wedge lb\,(u, y))$

  (Allen's "starts" relation)

- $ends(x, y) =_{df} Chron\,(x) \wedge Chron\,(y) \wedge tppart\,(x, y) \wedge \exists u\,(rb\,(u, x) \wedge rb\,(u, y))$

  (Allen's "ends" relation)

- $during(x, y) =_{df} Chron\,(x) \wedge Chron\,(y) \wedge tppart\,(x, y) \wedge \neg ends\,(x, y) \wedge \neg starts\,(x, y)$

  (Allen's "during" relation)

Axiome

- $\neg\exists x\,(Tb\,(x) \wedge Chron\,(x))$

  (time-boundaries and chronoid are disjoint)

- $\forall x\,(Tb\,(x) \vee Chron\,(x))$
- $tb\,(y, x) \rightarrow Chron\,(x)$

  (tb applies to chronoids only)

- $tcoinc\,(x, y) \rightarrow Tb\,(x) \wedge Tb\,(y)$

  (x and y are coincident[6])

- $Chron\,(x) \wedge lb\,(u, x) \wedge rb\,(v, x) \rightarrow u \neq v$

  (the boundaries of a chronoid are different)

- $(Chron\,(x) \wedge ((lb\,(u, x) \wedge lb\,(v, x)) \vee (rb\,(u, x) \wedge rb\,(v, x))) \rightarrow u = v)$

  (the left and right boundaries of a chronoid are uniquely determined)

- $Chron\,(x) \rightarrow \exists u\,(lb\,(u, x)) \wedge \exists v\,(rb\,(v, x))$

  (every chronoid has a left boundary and a right boundary)

- $Chron\,(x) \wedge Chron\,(y) \wedge l\,(x) = l\,(y) \wedge r\,(x) = r\,(y) \rightarrow x = y$

  (there are no different chronoids with the same boundaries)

- $Chron\,(x) \wedge lb\,(u, x) \wedge rb\,(v, x) \rightarrow \neg tcoinc\,(u, v)$

  (the boundaries of a chronoid do not coincide)

[6] *coincident*. Occupying the same area in space or happening at the same time: a series of coincident events. See Synonyms for contemporary.

- $tcoinc\,(x,y) \rightarrow \neg\exists w\,((lb\,(x,w) \wedge rb\,(y,w)) \vee (rb\,(x,w) \wedge lb\,(y,w)))$

  (coinciding boundaries are not boundaries of one chronoid)
- $tcoinc\,(x,y) \rightarrow tcoinc\,(y,x)$
- $Tb\,(x) \rightarrow tcoinc\,(x,x)$
- $Tb\,(x) \rightarrow \exists y\,(x \neq y \wedge tcoinc\,(x,y))$
- $Tb\,(x) \rightarrow \neg\exists yz\,(x \neq y \wedge x \neq z \wedge y \neq z \wedge tcoinc\,(x,y) \wedge tcoinc\,(x,y))$
- $tpart\,(x,y) \rightarrow Chron\,(x) \wedge Chron\,(y)$

  (range restriction)
- $Chron\,(x) \rightarrow tpart\,(x,x)$

  (reflexivity)
- $tpart\,(x,y) \wedge tpart\,(y,x) \rightarrow x = y$

  (anti-symmetry)
- $tpart\,(x,y) \wedge tpart\,(y,z) \rightarrow tpart\,(x,z)$

  (transitivity)
- $\neg tpart\,(x,y) \wedge Chron\,(x) \wedge Chron\,(y) \rightarrow \exists z\,(tpart\,(z,x) \wedge \neg t \circ v\,(z,y))$

  (existence of a non-overlapping part)
- $Chron\,(x) \rightarrow \exists y\,(starts\,(x,y))$

  (every chronoid has a future extenstion)
- $Chron\,(x) \rightarrow \exists y\,(ends\,(x,y))$

  (every chronoid has a past extension)
- $comp\,(x,y)$

  (every two chronoids are compatible)
- $Chron\,(x) \rightarrow \exists y\,(Chron\,(y) \wedge during\,(y,x))$

  (there are no atomic chronoids)
-

$$\begin{gathered} Chron\,(x) \wedge Chron\,(y) \wedge meets\,(x,y) \\ \rightarrow \exists z(Chron\,(z) \wedge l\,(x) = l\,(z) \wedge r\,(y) = r\,(z) \wedge \\ \neg\exists w\,(tpart\,(w,z) \wedge \neg t \circ v\,((w,x) \wedge \neg t \circ v\,(w,y)))) \end{gathered}$$

  (sum of meeting chronoids)
-

$$\begin{gathered} t \circ v\,(x,y)\,\exists w\,(starts\,(w,x) \wedge \neg t \circ v\,(w,y)) \\ \rightarrow \exists z(Chron\,(z) \wedge l\,(x) = l\,(z) \wedge r\,(y) = r\,(z) \wedge \\ \neg\exists w\,(tpart\,(w,z) \wedge \neg t \circ v\,((w,x) \wedge \neg t \circ v\,(w,y)))) \end{gathered}$$

  (sum of meeting chronoids)

–

$$t \circ v\,(x,y) \wedge \exists w\,(starts\,(w,x) \wedge \neg t \circ v\,(w,y))$$
$$\rightarrow \exists z(Chron\,(z) \wedge l\,(x) = l\,(z) \wedge r\,(y) = r\,(z) \wedge$$
$$\neg\exists w\,(tpart\,(w,z) \wedge \neg t \circ v\,(w,x) \wedge \neg t \circ v\,(w,y)))$$

(sum of overlapping chronoids)

–

$$t \circ v\,(x,y)$$
$$\rightarrow \exists z\,(tpart\,(z,x) \wedge tpart\,(z,y))$$
$$\wedge\forall u\,(tpart\,(u,x) \wedge tpart\,(u,y) \rightarrow tpart\,(u,z))$$

(existence of intersection of overlapping chronoids)

- Die Axiome werden zusammengefaßt durch T_B bezeichnet.
- $B := Brentano$
- Fragen über T_B
  1. Ist T_B widerspruchsfrei?
  2. Ist T_B vollständig?
  3. Ist T_B entscheidbar?

Vorlesung vom 22.11.2007

1. Flächen sind existentiell abhängig von *3-dimensionalen Raumregionen*
2. Linien sind existentiell abhängig von Flächen
3. Punkte sind existentiell abhängig von Linien

Behauptung Oberfäche des Tisches ist eine Fläche.

$\vdash\dashv$ Chronoide

$\bigcirc$ Topoid

- "Bretzeln betrachten"...

$$\rightarrow$$
$$\downarrow\sqsubset\sqsupset\uparrow$$
$$\leftarrow$$

  - usw...

- Ein Wurm frisst sich in einen Apfel... und kommt auf der anderen Seite wieder raus... es hat sich ein Tunnel gebildet, also eine Röhre die glatt hindurch geht. Der Wurm kann aber die verrücktesten Wege gehen, und ein wildes Gebilde entsteht; aber: alle Gebilde die entstehen sind untereinander Äquivalent – es entstehen ja Zylinder.
  - Wir kommen daher zur *Knotentheorie*. Wie ist diese "Kurve" in den 3-dimensionalen Raum eingebettet? Ein mathem. ungelöstes Problem. Weiter mit...

$Top(x)$ 3-dimensional, kompakt und zusammenhängend.

- Wir wollen eine "Inselwelt" beschreiben. Die einzelnen Inseln erscheinen als zusammenhängende Gebilde. Daher...

$SReg(x) := spatial\,region$.

$spart(x, y) := x\,ist\,'spatial\,part'\,von\,y$. Ein räumlicher Teil eines dreimius. räuml. Gebildes ist selbst 3-dimensional.

$b(x, y) := x\,ist\,Grenze\,von\,y$.

$scoine(x, y) := die\,Grenzen\,x\,und\,y\,/\,Koinzidenzen$.

- Ist die Menge der Axiome, die entsteht, widerspruchsfrei? Warum Koinzidenzen von Grenzen? $\twoheadrightarrow$ Problem mit Flaggen. Wir bilden eine polnische Flagge. Wie sieht es mit der Begrenzung der roten und der weißen Fläche aus?

$Surf\,Reg(x) := \exists y(Sreg(y) \wedge b(x, y))$,

$LinReg(x) := \exists y(SurfReg(y) \wedge b(x, y))$,

$PtReg(x) := \exists y(LinReg(y) \wedge b(x, y))$ sowie

$SBd(x) := SurfReg(x) \vee LinReg(x) \vee PtReg(x)$.

- Punktgebilde der koinzidierten Grenzen (von Linien) im Raum nennen wir also nun *Brentano-Raum*. Brentano-Punkt im 3-dim. Brentano-Raum. $\longrightarrow$ Bei Beispiel viele Linien auf einen Punkt gezeichnet... es geht darum, was der eigentliche "Schnittpunkt" der Linienenden ist, die sich nun treffen.

$d(x, y) \geq 0, d(x, y)$ **ist positive reele Zahl,**

$d(x, y) = d(y, x)$,

$d(x, y) + d(y, z) \geq d(x, z)$ und

$d(x, y) = 0\, gdw\, x = y$.

- Eine Linie im Raum koinzidiert mit unendlich vielen paarweise verschiedenen Linien.
- Mit wie vielen Fächen kann eine Fläche im $B - Raum$ koinzidieren?
- Materielle Entitäten und räumlich-zeitliche Entitäten. ($\neq$Raum, $\neq$Zeit)

1. Entitäten, die zu Zeitpunkten vollständig existentiell präsent sind;
   $\Rrightarrow$ Presentiale[7]
2. Entitäten, die eine zeitliche Ausdehnung haben;
   $\Rrightarrow$ Prozesse

Materielle Strukturen sind räuml.-zeitlich. Entitäten, welche

1. Presentiale sind– und
2. Raum besetzen

   $at(S, t) := S\, existiert\, zum\, Zeitpunkt\, t$,
   $occ(S, T) := S\, besetzt\, den\, Raum\, T$. (von occupy)

- (Beispiel Tasse) Qualitäten. $\circlearrowleft$ materielle Struktur MS

  $i(x, y) := die\, Qualität\, inhäriert\, in\, y$.
  $i := inhärieren$.
  $Qualität = individuelle\, Eigenschaft$.

Vorlesung vom 6.12.2007

Persistenz

$John\,(t)$ ist Presential
J
$at\,(J, t) :=$ exisitiert zum Zeitpunkt t.
$at\,(J, s)\ s \neq t,\ s < t$
$at\,(J, t_0)\ (t_0 := Zeitp.\, d.\, Geburt)$
$at(J, 60J)$
$\{John\,(t) \mid t \in LT\} = \Delta$
Persistant von John
$Perst(John)$ ist ein *Konzept*,
welches alle diejenigen Eigenschaften/Attribute erfasst, die für alle

[7] $\equiv$Präsential.

$John\,(t)$ gleich sind.
$Perst(John)$ hat als Instanzen genau die Entitäten in Δ.
$< \{Johnt\,(t) : t \in LT\}\,,\ Perst\,(John) >$
$Proc\,(John) :=$ ist ein Prozess, der u.a.
die folgenden Bedingungen erfüllt:
Für alle Zeitpunkte t des Chronoids von $Proc\,(John)$
$(| --\ |\ -\ |\ ---\ |)$ $\quad Proc\,(John) \restriction t - John\,(t)$ ist Einschränkung dieses Prozesses auf t.
$Def. \qquad \models Präsential$
Wir erweitern nun auf den Prozess;
$< \{Johnt\,(t) : t \in LT\}\,,\ Perst\,(John)\,,\ \mathbf{Proc(John)} >$
John ist nun in der Lage auch Handlungen auszuführen.
Dem Persistanten $Perst(John)$ lässt sich ein Individuum zuordnen, notiert $Perp(John)$[8].
$\Rightarrow$ WERTHEIMER. Zur Kognitionstheorie[9].

$Perst(x) := x\,ist\,Persistant$

$Perp(x) := x\,ist\,Perpetuant$

$Proc(x) := x\,ist\,Prozess$

$Pres(x) := x\,ist\,Präsential$

$persist(x, y, t) := Perst(x) \wedge y :: x \wedge at(y, t).$

- Beispiele v. Axiomen:
  - $\forall xy(Perst(x) \wedge y :: x \rightarrow Pres(y))$
  - $Pres(y) \rightarrow \exists t(at(y, t))$

Viele dieser Axiome kann man z.B. in eine Ontologiesprache wie *OWL* übersetzen.

Begriffe. Presentiale. Persistanten. Daraus abgeleitet Perpetuanten.

Prozesse

Temporale Komplexe

Jeder temporale Komplex (t.K.) hat eine zeitliche Ausdehnung.

$prtime(x, y) := y\,ist\,die\,temporale\,Extension\,des\,t.K.\,x.$
y ist z.B. ein Chronoid.
Zusammenhängende t.K.

$prtime(x, y) \wedge prtime(x, z) \rightarrow y = z$

[8] Perpetuant.
[9] Siehe hierzu auch: www.unet.univie.ac.at/~a0401627/Kognitionspsychologie_Allg3vomLeder.doc

$Chron(x) \rightarrow \exists y(tempcomp(y) \wedge prtime(y, x))$.

$\Longrightarrow$ Es wird abgelehnt, dass es Zeitintervalle gibt, in denen nichts geschieht. Es muss immer an irgendeiner Stelle etwas vorgehen.

Prozesse sind spezielle zusammenhängende temporale Komplexe.

Kausaler Zusammenhang

Hier *ohne* kausalen Zusammenhang. Billiardkugel wird vom Eiffelturm in Paris geworfen (P_1), bei Aufprall t_1 wird zeitgleich (ohne Abstand) eine Billiardkugel in Leipzig zu t_1' angestossen (P_2) die bis t_2 rollt.

$tcoine(t_1, t_1')$

$P_1 + P_2 = P \ \ [t_0 \ldots t_2]$

...temporale Komplexe, die kausal zusammenhängen *und kohärent* sind.

Grundklassifikation von Prozessen

stetige Prozesse continuous

Zustände states

Sprünge event

Abgeleitete Begriffe: *changes*, *histories*.

State Gegeben seien $U1$, U_2 zwei Konzepte, deren Instanzen Presentiale sind, $ex(U_1) \cap ext(U_2) = \emptyset$.

P ist ein U_1-State, wenn alle Prozessgrenzen von P Instanzen von U_1 sind.
$U_1 = rot \qquad U_2 = blau$
nur U_1. $| - - - \shortmid - \shortmid - \shortmid - |$
$U_1 + U_2$. $| \, rot \shortmid blau \shortmid rot \cdots |$ bis Einführung auf stetige Prozesse, bei Farbverlauf, beim dem Änderung bei Betrachtung von größeren Abständen ersichtlich ist.

Events sind bestimmte Extremalgrenzen von Prozessen.

Vorlesung vom 13.12.2007

zu den Prozessen

- $prtime\,(x, y) :=$ Projektion eines Prozesses x auf die Zeit y, y ist ein Chronoid.
- $timerestr\,(x, t, y) := x$ ist ein Prozeß, t ist Zeitidentität und y entsteht durch Beschränkung von x auf t.

Wenn t ein Zeitpunkt ist, dann ist y ein Presential.

Axiom $\forall x \forall t \exists y \, (timerestr\,(x,t,y) \wedge \forall xyt \, (timerestr\,(x,t,y) \rightarrow Pres\,(y))$

- $procpart\,(x,y)$
- $layerpart\,(x,y)$

Occurrent ist abhängig von Prozessen.

1. *Events*
2. *Changes*
3. *Histories*

Schiff des Theseus

$Tes\,(t_0) \rightarrow Tes\,(t_1) \rightarrow \cdots Tes\,(t')$ {Komplettes Austauschen der Teile}

$Tes'(s_1) \rightarrow Tes'(s_2) \rightarrow \cdots Tes'(t')$

Weiter mit...

$$R \subseteq M \times M$$
$$a < b \quad \{(a,b)\,|a < b\}$$

Ontologische Auffassung von der Relation

Bsp1 John küsst Mary. $R(K)$

$R(K)$ ist in *Konzept* (Categorie)

$r :: R\,(k)$, r als Instanz von $R(K)$ heisst ***Relator***, r hat Teile, die Rollen genannt werden, für r:

$q_1 =$ Rolle des Küssenden

$q_2 =$ Rolle des Geküssten

K Kuss ist ein Prozeß und der Relator fundiert auf K.

$Obj(I) \wedge (Obj(M) \wedge Cat(R(K)) \wedge \exists q_1 q_2 r(r :: R(K) \wedge Roleof \cdots$ (nicht mehr lesbar)

John ist Vater von Mary, $V := Vaterschaftsbeziehung$

$r :: V$

$Obj(I) \wedge Obj(M) \wedge Role(Vater) \wedge Role(Tochter) \wedge$

$\exists q_1 q_2 r(r :: V \wedge q_1 :: Role(Vater) \wedge q_2 :: Role(Tochter) \wedge role - filler(q_1, I)$

$\wedge role - filler(q_2, M) \wedge role - of(q_1, r) \wedge role - of(q_2, r))$

$$V :: r \quad \begin{array}{ll} \overrightarrow{role-of}\, q_1 & \overrightarrow{filler}\, John \\ \overrightarrow{role-of}\, q_2 & \overrightarrow{filler}\, Mary \end{array}$$

Fakten

$$\underbrace{\overrightarrow{John\, kuesst\, Mary}}_{Prozess}$$

$$\underbrace{< r :: R(K), q_1, q_2, I, M >}_{Fakt}$$

Fakten sind Teile der realen Welt.

$\Rightarrow$ Übergang zu Inferenzen und Propositionen

$\varphi : Es\, regnet,\, W$
$\varphi\, ist\, wahr\, in\, W,\, W \models \psi$
$Gilt\, W \models \psi\, oder\, W \models \neg\varphi$
$R_1(W) \models \psi,\, R_2(W) \models \neg\varphi$

Was sind sinnvolle Teile $PW \leq W$ der Welt, die die Beziehung $PW \models \psi$ sinnvoll prozessieren?

$R(Leipzig)[10.06.05] \models \psi$
$TB \leq W$

Vorlesung vom 24.1.2008

$Student = Person \sqcap$

$\exists hat.Name \sqcap$

$\exists belegt.Lehrveranstaltung \sqcap$

$\exists hat.Adresse$

$\underbrace{Student \sqsubseteq Person}_{is-a\, Relation}$ (Subsumpt. Relation)

OWL Spezifikation vom W3C – Im englischen ausgesprochen als "Eule", Symbol der Weisheit.

OWL Lite

OWL DL

OWL Full

RDF-Modell. Tripel (S, P, O). $S \xrightarrow{P} O$.

Ontologieeditoren mit Beispielen

$\underbrace{\text{SWOOP, Protege}}_{DL}$ $\underbrace{\text{OBOEdit}}_{Aufbau\,von\,Graphen}$

Entwicklung von Ontologien

1. STEP 1. Domänen-Spezifikation & Anforderungsanalyse
   a) Domäne. "Wein"
   b) Welche Objekte sollen betrachtet werden?
   c) (Sicht und Kontext wichtig)
   d) Auflistungen von Fragen:
      i. Eigenschaften des Weins,
      ii. Ist Bordeaux ein Rot- oder Weißwein,
      iii. $\vdots$
2. STEP 2. Konzeptualisierung.
   a) Aufstellung einer Liste von Termen (Konzepte), die in der Domäne verwendet werden,
   b) $T := nat.\,sprachl.\,Definition$
3. STEP 3. Axiomatisierung.
   a) Anordnung der Terme der Liste von STEP 2.
   b) Taxonomie.
      i. $L = \{C_1, C_2, \ldots\}$
      ii. $C_i\,is\text{-}a\,C_j$
      iii. $Weisswein\,is\text{-}a\,Wein$
   c) Weitere Beziehung zwischen Konzepten, $part\text{-}of, \ldots$
   d) Volle Axiomatisierung
      i. $\Longrightarrow$ *Semi Formale Spez. Dokumentation*

- SWOOP

  $\underline{Thing}$

  $- \quad \diagdown Wine - -usw.$

Vorlesung vom 31.1.2007

Domäne "Wein"

1. SCHRITT 1. Domänenspezifikation
   a) Begriffsanalyse, Requirement-Spezifikation, Sichtweise, Eigenschaften, Kontexte und Objekte
   b) Kompetenzfragen
      i. Welche Eigenschaften sollen betrachtet werden?
      ii. Ist *Wein X* für Meeresfrüchte geeignet?
2. SCHRITT 2. Konzeptualisierung
   a) Auflistung der bzgl. Frageliste relevanten Terme, Relationen und Eigenschaften
   b) Bordeaux ist ein Rotwein *part-of* v ist von der Lokalität v
3. SCHRITT 3. Axiomatisierung
   a) $R_1 \ldots R_n$
   b) $R_1(C_1, C_2)$
   c) *Wein 2* ist von der Lokalität Italien (is-a)

- Ansprechpartner
  - Martin Löwe
  - Robert Höhendorf

6. Biomedizinische und medizinische Ontologien

6.1 Medizinische Ontologin

- ICD (International Statistical Classification of Diseases)
  - Aktuelle Ausgabe ICD-10
    * Diese Systeme sind "historisch gewachsen"
    * deutsche Ausgabe
      · ICD-10-GM (GM=german)
    * $F6$ Persönlichkeitsstörung
      · $F60$ —
      · $F60.3$
      · $F60.31\, Besondere\, Typen$

    * Prof. Herre:
        · "Wenn Sie das haben, dann können Sie einfach mal gucken, was Sie da haben."
    * **V**: Verdacht
    * **G**: gesicherte Diagnose
    * **A**: Aussschluss einer Krankheit
    * SNOMED-CT (Systematized Nomenclature of Medicin, Clinical Terms)
    * LOINC (Logical Observation Identifier Names and Codes)
    * Prof. Herre:
        · "Das ist ein sehr dummes System; seit 1994 wird es entwickelt."
    * Ziel
        · Erleichterung des Austausches und Beschreibens von Labordaten und anderen klinischen Daten
    * Galen (generalized Architecture for Language, Encyclopedia and Nomenclature in Medicine)
    * A. Rector (Manchester)
    * basiert von der Repräsentationsform her auf DL
    * MeSH (Medical Subject Headings)
- Kontrolliertes Vokabular
    * Synonym
    * $t_1, t_2 \in$ Term sind synonym, wenn
    * $\{c\,|\}\ldots$
    * Homonome[10] Term
    * $t \in Term$
    * t ist Homonym gdw
    * $\|\{c|sem\,(t,c)\}\| > 1$
    * Beispiel an Tau
        · Tau—Seil (Konzept,
        · Tau—Niederschlag,
        · Tau—19. Buchst. im griech. Alphabet
    * $Term \rightarrow Konzept \begin{cases} Inhalt & (Intension) \\ Umfang & (Extension) \end{cases}$

[10] *Homonym*. Siehe: http://de.wikipedia.org/wiki/Homonym. Beispiel: Arm (Körperteil) und arm ("mittellos").

    * N. will wissen, ob **T. Mann** oder **Autor** *des Romans(T_2)* **"Der Zauberberg"**(T_1) ist
    * $T_1 = T_2$

Biomedizinische Ontologien

- OBO-Ontologien
  - Open Biomedical Ontologies (OBO) und
  - GO (Gene Ontologie)
  - Siehe: http://obo.sourceforge.net/
  - Motivation für OBO
    * Creating ontologies takes a lot of work $\Rightarrow$ reuse
  - GO
    * GO ist a member vocabulary of OBO
- GFO $\longmapsto$ Analyse von Domäne (unser Fokus) $\longmapsto$ formal. Repr. Ontologie.
- multivalent anwendbar

Teil II.
Zeugs

KI und Computer

Turing-Maschine und Berechenbarkeit

Wenn Intelligenz mit Computern realisiert werden soll, dann muss sie auf Rechenleistungen zurückführbar sein.

> Nicht jedes Programm einer Turing-Maschine ist so einfach wie die Addition. Im Prinzip kann aber das Rechnen mit natürlichen Zahlen auf das Manipulieren von 0 und 1 mit den Elementarbefehlen einer Turing-Maschine zurückgeführt werden. Allgemein unersucht die Arithmentik mehrstellige Funktionen f mit den Argumenten $x_1, \ldots, x_n$, also z.B. $f(x_1, x_2) = x_1 + x_2$. Jedes Argument ist eine Zahl, die auf einem Turing-Band durch eine Einserkette dargestellt wird. Die restlichen Felder sind leer, d.h. mit 0 bedruckt. Am Anfang einer Rechnung stehen also auf einem Turing-Band nur Einserketten, die durch Nullen getrennt sind, also: $\ldots 0x_1 0x_2 0 \ldots 0x_n 0 \ldots$. Eine n-stellige Funktion f mit Argumenten $x_1, \ldots, x_n$ heißt Turing-brechenbar, wenn es eine Turing-Maschine gibt,

> die mit einer Bandbeschriftung $\ldots 0x_1 0x_2 0 \ldots 0x_n 0 \ldots$ beginnt und nach endlich vielen Schritten mit der Banbeschriftung $\ldots f(x_1, \ldots, x_n) 0 \ldots$ stoppt. Dabei wird der Funktionswert $f(x_1, \ldots, x_n)$ durch eine entsprechende Einserkette dargestellt.

Computer sollen Probleme maschinell lösen. Bei schwierigen Problemen wird sich ein Benutzer kaum noch zusätzlich um die Zuordnung von Werten und Speicherstellen kümmern können. Daher wurden problemorientierte Programmiersprachen eingeführt, die es einem Nutzer erlauben, Probleme und ihre Lösungsverfahren in einer problemnahen Sprache auszudrücken. Ein Beispiel war die frühere Programmiersprache FORTRAN, die mit ihren Formelausdrücken zur Lösung mathematischer Aufgaben geschaffen wurde.

Problemorientierte Sprachen werden als höhere Programmiersprachen bezeichnet, da sie nicht maschinenorientiert sind und auch nicht in die Maschinensprache übersetzt werden müssen. Andere Beispiele von höheren am Nutzer orientierten Programmierpsrachen legen auf einen übersichtlichen strukturierten Programmierstil Wert, wie z.B. PASCAL. Die erwähnten KI-Sprachen LISP und PROLOG eignen sich als höhere Programmiersprachen besonders für Wissensrepräsentationen.

KI, Unendlichkeit, Unentscheidbarkeit und Unvollständigkeit

Die Begriffe universelle Brechenbarkeit und Entscheidbarkeit sind zwar sehr mächtig und umfassen einen großen Bereich wissenschaftlicher, technischer und praktischer Probleme. Wie die Mathematik zeigt, ist menschliches Denken aber mächtiger als Algorithmen, die nach endlicher Zeit stoppen: Mathematik ist mehr als Rechnen! Wir können nicht nur unendlich große Mengen bilden. Wir müssen sie sogar voraussetzen, um maschinelle Algorithmen vollständig definieren zu können. Sind damit Grenzen künstlicher Intelligenz verbunden?

> Das Unendliche kommt bereits durch das elementare Zählen in die Mathematik: $0, 1, 2, 3, \ldots$. Wir beginnen mit 0 und erhalten dann jede weitere Zahl, indem wir die vorausgehende Zahl n um 1 erhöhen, also $n + 1$. Dieser Prozess bricht nie ab. Zu jeder noch so großen Zahl n gibt es immer noch eine Nachfolgerzahl $n + 1$. Bei der Bildung einer neuen Zahl wird auf die bereits gebildete Vorgängerzahl zurückgegriffen. Man spricht daher auch vom *Prinzip der Rekursion*. Nach diesem diesem Prinzip lassen sich Funktionen für alle unendlichen natürlichen Zahlen definieren. So wird unsere beim Problem des Handlungsreisenden erwähnte Fakultätsfunktion f, die jeder Zahl n die Fakultät $n!$ zuordnet, rekursiv definiert, indem man ihren Wert in 0 gleich 1 setzt und vom Wert in n zum Wert in $n + 1$ durch Multiplikation mit $n + 1$ übergeht, also $f(0) = 1, f(1) = 1, f(2) = 2, f(3) = 6, f(4) = 24, \ldots$. Auch z.B. Addieren, Multiplizieren und Potenzieren werden nach dem Prinzip der Rekursion für alle natürlichen Zahlen definiert.

Bei der Berechnung primitiv-rekursiver Funktionen sind also unendliche Programmschleifen, die nicht abbrechen, ausgeschlossen. Wir kommn für ihre Definition noch ohne das Unendliche aus. Damit sind allerdings nicht alle Algortihmen erfasst. Historisch konnte der Münsteraner Logiker W. Ackermann 1928 erstmals eine berechenbare Funktion angeben, die nicht primitiv-rekursiv ist. Bei Turing-Programmen sind Schleifen möglich, die so lange (engl. *while*) durchlaufen werden müssen, bis eine Zahl gefunden ist, die eine besimmte Prüfbedingung erfüllt. Bei Computerprogrammen mit *while*-Schleifen kann es vorkommen, dass keine Zahl mit der Prüfbedingung gefunden wird und das Programm (wenigstens theoretisch) unbegrenzt weiterläuft. Jedenfalls lässt sich bei *while*-Schleifen die Anzahl der Durchläufe nicht voraussagen.

KI und Information

Die Anzahl der Binärentscheidungen (bit), die zu einem Zeichen führt, heißt Informationsgehalt I des Zeichens, also in unserem Beispiel $I = 2$ bit. Allgemein sind bei einem Vorrat von N Zeichen $N = 2^I$ Auswahlverfahren mit I Binärentscheidungen möglich. Der Informationsgehalt I des Zeichens ist daher der Logarithmus von N zur Basis 2 (logarithmus dualis), d.h. $I = ld\,N$ bit.

> Wenn alle Zeichen $z_i (1 \leq i \leq N)$ eines Zeichenvorrates mit gleicher Wahrscheinlichkeit p_i auftreten, dann ist $p_i = 1/N$, also $N = 1/p_i = p_i^{-1} = -ld\,p_i$. Treten alle Zeichen mit verschienen Wahrscheinlichkeiten auf, dann beträgt der Infomationsgehalt des Zeichens z_i mit der Wahrscheinlichkeit p_i des Auftretens $i(z_i) = -ld\,p_i$. Ein wahrscheinliches Zeichen hat also geringeren Informationsgehalt als ein unwahrscheinliches. Der Informationsgehalt eines Zeichens lässt sich daher als Neuigkeitswert oder Überraschungsgrad für den Empfänger auffassen.

Information und Computertechnik

Bei der Anwendung des wahrscheinlichkeitstheoretischen Informationsbegriffs auf Computer wird die Erwartungswahrscheinlichkeit von Zeichenfolgen untersucht, die von einem Computer erzeugt werden. Computer setzen Algorithmen der Informationsverarbeitung voraus. Wie komplex sind die informationsverabeitenden Algorithmen von KI-Systemen?

Informationen von Computern werden durch Zeichenreihen (Daten) dargestellt. Wir betrachten z.B. die binären Sequenzen

$s_1 = 00000000000000000$

$s_2 = 0010001000100010$

$s_3 = 0110110011011110.$

Die algorithmische Informationstheorie untersucht die Fragen: Welcher Algorithmus bzw. welches Computerprogramm kann diese Sequenzen erzeugen? Wie komplex sind diese Programme? Die Sequenzen s_1 und s_2 sind von großer Regelmäßigkeit und können daher durch ein kurzes Programm erzeugt werden. In Sequenz s_1 ist 16-mal der Druckbefehl von 0 nacheinander wiederholt. s_1 entspricht daher dem Programm "16 mal 0". s_2 entspricht dem Programm "4 mal 0010". Demgegenüber weist die Sequenz s_2 keine Regelmäßigkeit auf. Intuitiv ist der "Informationsgehalt" von regulären Mustern wie s_1 und s_2 gering, da durch die Wiederholung von Teilmustern nichts Neues erzeugt wird. Nach dem Vorschlag von Kolmogorov und Chaitin nimmt der Informationsgehalt und die Komplexität eines Musters mit wachsender Irregularität bzw. abnehmender Regularität zu.

Eine beliebige Sequenz von Zeichen kann durch verschiedene Programme ausgedruckt weren. Für die Sequenz s_2 könnten wir auch ein längeres Programm angeben, dass z.B. 18-mal nacheinander die Druckbefehle für die angebenen Symbole auflistet, also "Drucke 0, Drucke 0, Drucke 1, Drucke 0, ...". Das kürzeste dieser Programme heißt Minimalprogramm. Die Länge eines Programms lässt sich in der Einheit *bit* messen. Für ein kurzes Programm z.B. s_1^* zur Erzeugung der Binärsequenz s_1 aus 16 Nullen reicht eine Programmlänge von etwa $ld\,16 = 4bit$. Für viele der 2^{16} möglichen Binärsequenzen aus 16 Stellen wird wegen der hohen Regellosigkeit erforderlich, 16 Programmbefehle für die einzelnen Symbole aufzulisten. Solche Programme benötigen eine Länge von mindestens 15 oder 16 bit.

KI und Internet

Zur Übermittlung der Daten bringt der Sender die Übertragungsleitung abwechselnd in verschiedene Spannungszustände. Da eine Nachricht in Maschinensprache binär codiert ist, kann eine Leitung mit zwei Zustandswechseln ein Bit pro Zustandsänderung übertragen. Es kann aber auch mit mehr als zwei Übertragungswerten gearbeitet werden. Bei z.B. acht Spannungswerten mit 0, 1, 2, 3, 4, 5, 6 und 7 Volt können drei Bits pro Zustandsänderung übertragen werden. Die Übertragungskapazität einer Leitung für Nachrichten hängt daher von der Menge der unterschiedlichen Werte und der Häufigkeit ab, mit der ein Sender den Wert pro Sekunde übertragen kann. Die Maßeinheit für den Wechsel pro Sekunde wird nach dem französischen Ingenieur Emile Baudot mit *baud* bezeichnet.

Teil III.
Abschriften

Andres Mitschriften

CYC $A \to B$ Als Ersatz bei der materiellen Implikation benutzt.

- darauf bauen immer komplexere Ontologien auf

Struktur und Repräsentation von Kategorien

- Begriffe von Zeit/Raum[11] unabhängig; als Instanzen dieser Begriffe
- primitive Kategorie
  - Begriff Katze $\Rightarrow$ Individuum (Katze) bzw. Instanz
- man braucht einen Rahmen um domainspezifische Ontologien zu bauen

Semantische Funktion von Ontologien

- Integration von Ontologien
- gerne universelles Framework bei Ontologien
- GO (Gene Ontology) $\Rightarrow$ BP und CC und BF (Biological Function)
- Framework soll die spez. Ontologie steuern und Verbindungen zwischen diesen zulassen

  Leben besteht aus Schnappschüssen, das verstrickende Element sind kontinuierliche Prozesse

Begriff des Wissens $a\,priori\,Wissen$

Wissensaquisition Wissenserhebung $\Rightarrow$ Strukturierung

- Rahmensystem dient dazu, Wissen zu formalisieren

Problem Semantik von rein syntaktischen Objekten bilden

Erfüllbare Form der Formalisierung Mathematik als Grundlage der Axiomatisierung von Mengentheorie.

- Für alle Individuen i: $i :: Affe$, so $i :: Saeugetier$.

[11] http://rosw.cs.tu-berlin.de/voelz/PDF/RaumZeit.pdf

1. 1. Schritt
   - Grundlegende Terme erkennen und in Beziehung setzen durch die *is-a* Beziehung .

★**automatisieren** Benutzt Wissen aus Toplevel bzw. Ontologien, als Vorr. bei der Automatisierung.

1. Problemspezifikation
   a) Angabe einer Domäne
2. Konzeptualisierung
   a) Grundkonzepte zwischen den bestehenden Relationen

   Folgende Fragen sind zu beantworten:
   a) Welche Objekte bei D sind relevant?
   b) Welche Konzepte und Relationen sind bei der Beschreibung des Wissens relevant?

   $\Sigma = (C_1 \dots C_n, R_1 \dots R_n, Obj\,(N)) + Taxonomien$

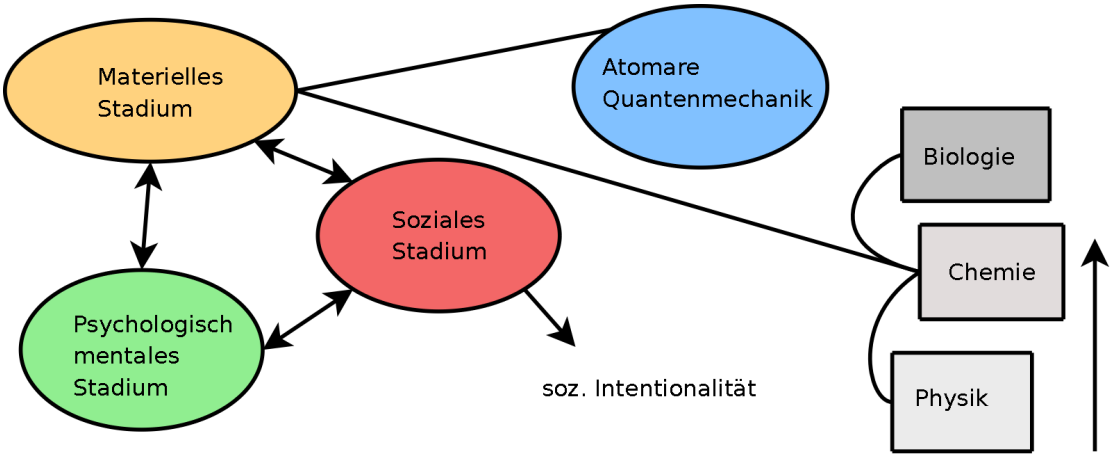
   $\curvearrowright$ Konzeptualisierung d. Domäne

   Frage Was ist ein gutes System?

   $\Longrightarrow$ Möglichst wenige Begriffe, die anderen daraus ableiten.
3. Axiomatisierung und Formalisierung
   - $Ax(C_1, \dots, C_K, R_1 \dots R_L) \rightarrow \forall xyz\,(R\,(X,Y) \wedge R\,(Y,Z) \rightarrow R\,(X,Y))$
   - Aus inkonsistenten Wissensbasen können Reasoner falsches Wissen schließen.

Toplevel Ontology

Abbildung 1: Toplevel Ontology



Materielles Stratum Objekte usw. Theorie u. Experimente

Bsp. Quantenmechanik

- GFO — ausgebildet (relativ gut) für die materielle Welt
- auftauchende Probleme
  - Bsp. Architekt (mit Wissen über Hausbau) ⇒ Gestaltung ⇒ Plan (Vorr. Idee), . . . ⇒ alles folgende baut dann auf ⇒ Haus.
  - PLAN/IDEE $\Longrightarrow$ PLAN $\Longrightarrow$ HAUS

Soziales Stratum Mehrere Individuen treten in Wechselbeziehungen

- soz. Intensionalität
- GFO — beschränkt sich im Moment auf das materielle Stratum

  Universal, Konzept, Symbol = Klassifikation der Welt
- Unterscheidung zw. abstraktem Symbol un den Token
- individuelles Token
- Symbole / Instanzen sind Tokens
- Symbolsysteme
- Symbole müssen verarbeitet werden
- Turingtest

Vier Arten von Hauptkategorien

- konzeptionalistische
- Nominalismus
- ...

Wissensaquisition bei Ontologien

Individuen Können nicht instanziiert werden

Konzepte, Kategorien müssen instanziierbar sein

Frage: Welches Zeitmodell, und für welchen Zweck?

- Ontologie d. Zeit von FRANZ BRENTANO $\Rightarrow$ HUSSERL[12] $\Rightarrow$ $Heidegger$

Chronoid mit a, b Zeitgrenzen. Zeitgrenzen haben keine vom Intervall unabhängige Existenz.

(Punktausdehnung O, Intervall größere Ausdehnung als O)

$$lb\,(x,y) := x\,ist\,linke\,Grenze\,des\,Intervalls\,y\,(left, \cdots$$
$$rb\,(x,y) := y\,ist\,rechte\,Grenze\,\cdots"\cdots"\cdots \quad (\cdots right\,boundary)$$

$$tpart\,(x,y) := x,y \quad Chronoide,\,x\,ist\,Teil\,von\,y$$

$tpart(x,y)$

$$|-----|$$
$$y$$

$$|--|$$
$$x$$

- BRENTANOS Koinzidenzrelation: $coinc(x,y) := x\,und\,y\,sind\,Zeitgrenzen\,die\,koinzidieren$
- Probleme bei der Abbildung von Prozessen mit Hilfe der rellen Zahlen

[12]Siehe hierzu auch: http://www.fritzkraemer.de/philosophie/zwischenpruefung.pdf

Bsp. *Wurf nach oben*

$$\begin{array}{ccc} & \circledast & \\ & \uparrow & \\ \Uparrow & \uparrow & \Downarrow \\ auf & \uparrow & ab \\ \frown & \frown & \frown \end{array}$$

$$\begin{array}{ccc} \Uparrow auf & & ab \Downarrow \\ & |--\dagger--| & \\ & \uparrow & \\ & (was\, passiert\, hier?) & \end{array}$$

Annahme: Letzer Zeitpunkt d. Aufwärtsprozesses und des Abwärtsprozesses kein Punkt oder umgekehrt.

$auf\ e]\,[a\ ab \rightarrow coinc \mid\ e, a\, Abstand\, O\ --\, und\, sind\, trotzdem\, verschieden$

(normale mehrdimensionale Räume: zwei Punkte haben den Abstand O, dann sind sie gleich)

- Punkt im Brentano-Raum hat unendlich viele überabzählbare Teile

4D-Ontologie[13]

...haben keine ★, sondern Prozesse.

$\forall x\,(occ\,(X,T) \wedge spart\,(S,T) \rightarrow \exists y\,(matpart\,(Y,X) \wedge occ\,(y,s)))$ (gilt nicht, bspw. auf atomarer Ebene)

$matpart \neq spart$

$occ\,(G_1,S_1) \wedge occ\,(G_2,S_2) \wedge coinc\,(S_1,S_2)$

4D versus **3D**: In principle, there are infinitely many ways in which we can model the world, so it is perhaps surprising that there are two main approaches, with on the whole minor variations, that dominate the literature. We will call these the 3D paradigm and the 4D paradigm, though they are also known as endurantism, and perdurantism.

Aus einer ppt-Folie. (Adrian)

[13] *4D Ontologien*. Siehe auch: http://suo.ieee.org/SUO/SUO-4D/index.html und www.blackwell-synergy.com/doi/pdf/10.1111/j.1468-0068.2006.00624.x.

3D Ontologie Enduratism

4D Ontologie Perdurantism

- A 4D ontology treats all individuals — things that exist in space-time — as spatio-temporal extents, i.e. as 4D objects.
- The principles of the 4D paradigm are:
  - Individuals exist in a manifold of **four dimensions**, **three space** and **one time**. So things in the past and future exist as well as things in the present.
  - The four dimensional extent is viewed from outside time rather than in the present.
  - Individuals (including physical objects) extend in time as well as space and have both temporal parts and spatial parts.
  - When two individuals have the same spatio-temporal extent they are the same thing. (However not all version of 4D insist on this principle).
  - Thus a 4D object is not (usually) wholly present at a point in time, but its whole is extended in space as well as time. The object at a point in time is a temporal part of the whole. Change is naturally expressed through a four dimensional classical mereology, which Simons: "Parts: A Study in Ontology" in his seminal work, describes in one page.

Conclusion

- A 4D ontology sees physical objects as extended in time as well as space
- ISO 15926 is a data model that is also a 4D ontology
- It uses a possible worlds approach rather than modal logic

Morphologie

$[HAUS] \leftarrow H \;\; P$ befindet sich in H

$occ\,(P,T) \wedge occ\,(H,S) \rightarrow spart\,(T,S)$

- Presentiale reichen nicht aus, für die Beschreibung der Welt

Persistanten

Paul trifft heute John, mit dem er vor zehn Jahren die Schule besuchte.

$R,\ Paul\,(R)\,,\ John\,(R)\,,\ h-30=t_1$
$Paul(t_1),\ John(t_1)$

Persistant ist ein —, $Perst\,(John)$

Instanzen des Persistanten $Perst\,(John)$ sind die Initialen in $\{John\,(t) \mid t \in Perst\,(C)\}$.

Perpetuanten Individuum

$John \quad Perst\,(John)$

$$\underset{1h}{|---\,|\sim\sim\sim|}$$

$exhib\,(n, q, t)$

Endurant Ein Endurant ist ein Individuum, das zu jedem Zeitpunkt *seiner Existenz* vollständig präsent ist *und* durch die Zeit persistiert (gleich bleibt).

$\Sigma(to)\,E\,(t_1)$ Entweder sie haben zwei Entitäten, oder die Zeitpunkte sind gleich.

Schiff von T(h)eseus

Siehe: http://www.thomas-buchheim.de/Theseus_Referat.pdf

> Eine Situation ist ein *kohärentes System* von Fakten, dass als ganzes angefasst werden kann. Situationen sind zu einem Zeitpunkt vollständig präsent.

Veranschaulichung 100m-Lauf $\curvearrowright$ $Schnappschuss$ $\twoheadrightarrow$ Situationstheorie.

$\Longrightarrow$ Situation, $AtomS\,(S)$

Situide sind Prozesse, die die Koherenz Kohaehersibilität erfordern — Situation um zeitliche Dimension erweitert.

$Vater\,(John, Max)$

$\mathbf{W}\ldots \subseteq \mathbf{R}_{ealität}\,W \models \Phi,\ gdw.\,es\,ex.\,Fakt\,f \in W:\ "(f\,entspricht\,dem\,\Phi)"$

$\longrightarrow\ die\,genaue\,Analyse\,dessen\,nicht\,so\,einfach$

> "Es regnet"

$\Phi := Es\,regnet.$
$\omega \models \Phi \quad \omega \models \Phi, \neg\Phi\ldots$ das es gleichtzeitig *regnet* und *nicht regnet* in der Welt.
$Reg\,(Leipzig) \models R\,(Leipzig)\,[10.5.06] \models \Phi$
Situide erfassen genug um Aussagen zu treffen, und als Wahr zu gelten.

Struktur und Repräsentation von Kategorien

drei Arten Konzepte, Universalien, Symbole

- Konzepte-Verbindung "zu fest"

Universalien Noch eine Verbindung zu den realen Dingen

- $Affe := Term(Affe)\, oder\, Real(Affe)$

1. erste Auffassung
   - Affe ist ein Universal und es gibt eine "direkte" Korrelation bez. $Term\,(Affe)$ und $Affe\,(real)$.
   - $[John] \rightarrow John\,(real)$
2. zweite Auffassung

$$\begin{aligned} Term &\rightarrow Affe(real) \updownarrow \\ &\rightarrow Konzept \updownarrow \end{aligned}$$

   - $\curvearrowright$ Existente aber schwummige Beziehung
   - *Konzept* hat als Begriff soziale Komponente

- Wissenserhebung, Konzeptualisierung, Axiomatisierung (Prädikatenhalt., 2. Stufe)
- Was ist eine Domäne? Ein Teil der Welt.

  Die Spezifikation der Domäne ist von erheblichen Komplexitätsgrad

Strukturen und Repräsentationen von Kategorien

Konzept Universal, Symbol

Domäne und die Spezifikation

$D \subseteq \omega$ (Domäne ist ein Teil der Welt)

HB: Objekte der Domäne (HB: Herman B. . . ys)

- $D = (Objekte, V \ldots, CP)$
- V: Menge von V. . .
- CP: Menge von Klassifikationsprinzipien

Konzeptualisierung

$concept\,(D) = (Ind\,(D)\,, Cat\,(D)\,, Rel\,(D))$

Bsp. $D = (Ind, CP, V)$

- X ist Affe, oder X ist Brücke
- Jeder Ausdruck der natürlichen Sprache mit einer Variablen definiert eine Kategorie.

Axiomatisierung

- Zusammenhänge zwischen den Kategorien

$r \in Rel\,(D)$

$r\,(C_1, C_2) \quad r := is - a \quad Affe(C_1)\, is - a\, Säugetier(C_2)$

$is - a\,(Affe, Säugetier)$

$\forall x\,(x :: C_1 \rightarrow x :: C_2)$

- Formalismus in dem die $Cat\,(D)$ erfasst werden
- Klassifikationsprinzipien

Struktur der Konzepte

- Unterscheidung Extension/Intension von Konzepten

Struktur von Typ

$0 - Typ$ d. Individuen (diese haben keine Insanzen)

X sei eine Menge von Typen, dann is $< X >$ ein Typ.
Weitere Typen gibt es nicht. (in der Regel Symbole)

$Bsp.$
Eine Kategorie vom Typ $< \{X\} >$ hat die Eigenschaft, dass alle Instanzen Typen haben, die Element von $\{0\}$ sind.

Diese Kategorien heissen *primitiv*. (sehr verbreitet)

- $< \{< \{0\} >\} >$ keine Kategorie deren Instanzen Individuen sind

Instanziierungsbezug ist im allgemeinen nicht transitiv.

$is - a$ Relation immer transitiv.

Mathmatik hat Begriffe die Instanzen haben...

Konzeptspezifikation

Konzepte als Mengen von Eigenschaften

$Catp\,(C, 0) := C\, ist\, kategorialer\, Teil\, von\, D$

Ein Konzept heisst wohldefiniert, wenn es keine abgeleiteten endlichen ::-Ketten gibt

$Ext\,(C) = \{a \mid a :: C\}$ (Extension C ist eine Menge)

$Int\,(C) = \{D \mid cat\,(D, C)\}$ erlaubter Fall, Menge von Prädikaten

$C\, is - a\, D$ (arbeitet mit Extensionen)

$C\, is - sub\, D\, gdw. \forall x(catp\,(X, C) \rightarrow catp\,(X, D)$

$C\, is - sub\, D$, so $D\, is - a\, C$ (unscharfe Regel)

$Catp$:

Ontologie der Morphologie

- ★Kasati, Warzi
- Medizinische Daten
- Unterscheidung Term (als String) / Konzept

Ontologie in der Medizin und Biologie

1. ICD-10
   a) 22 Hauptkategorien
2. SNOMED-CD
   a) alles in medizinischen Termen zu erfassen
   b) 18 Kategorien
3. COINC
   a) —
4. GALEN
   a) Toplevel Ontologie
   b) A. Rector (Univ. Manchester)

   c) "Why Clinical Terminology is hard"

5. MESH
   a) Sind alles taxonomische "Ontologien",
6. UMLS (Unified Medical Language System)
   a) Integration von medizinischen Taxonomien in eine allgemeine Basisstruktur
   b) semantisches Netz dient der Integration der einzelnen Systeme
   c) Kagegorisierung aller Terme sowie Basisrelationen und die Einordnung der semantischen Beziehungen zwischen diesen Kategorien

Basisrelation

- Axiomatisierung nicht-hierarchischer Relationen
  1. OBO-Ontologien (Open Biological Ontology)
  2. GO (Gene Ontology)

$D \leftrightsquigarrow Ontologische\,Analyse$
$D = (Ind, V, PC)$
$Concept(D)$

Vortrag

Inhalt

- Semantik von wissensrepräsent. Syst.
- linguistische System
- wissensbasierte Systeme

  Herkunft Ontologie: "Lehre des Seins als Seiendes"
- Beschreibung der allgemeinen Zusammenhänge von Strukturen, der Welt —
- Hauptgedanke: best. Prinzipien entwickeln als Grundlage bei einer präz. Theorie der Modellbildung

  Ziel Ingeneurswissenschaft
- kategoriales System der allgemeinsten Begriffe und Axiomatisierung[14] dieser

[14] Unter der *Axiomatisierung* einer Theorie versteht man ihre Darstellung in der Weise, dass gewisse Sätze dieser Theorie, die Axiome, an den Anfang gestellt werden und weitere Sätze durch logische Deduktion aus ihnen abgeleitete werden. (Wikipedia)

- allgemeinste Kategorien (Raum, Zeit, Objekt...)
- $\Longrightarrow$ Ausdehnen des Abstraktionsgrades; Näherung an "Objekt", bspw. Abel
- Kategorien haben Instanzen. Bspw. Haus $\Rightarrow$ Haus in der Isastraße 29, Leipzig.
- seltene Begriffe haben Strukturen
- Top-Level-Ontologie zur Beurteilung der allgemeinsten Kategorien

  Ziel Mehr Transparenz

Vorlesung Betrachten ex. Ontologien,

Schwerpunktvorlesung Bioinformatik

1. Grundbegriffe der auf Logik basierenden Wissensaquisition
   a) Geschichte der KI, insbesondere Semantik der Wissensrepräsentation
   b) Normale Sprachen und Semantik
   c) Prädikatenlogik der ersten Stufe
2. Prinzipien der Wissensverarbeitung
   a) Stufen der Wissensverarbeitung
      i. (Domänen aufbauen, welche Begriffe und Konzepte...)
   b) schwierigste Teil Formalisierung und Axiomatisierung
3. Grundlagen der "Normalen Ontologie"
   a) Vorstellung einer Ontologie
   b) Raum, Zeit
   c) Repräsentationssprachen bei Ontologien, bspw. FOL
   d) Verwendung von Top Level Ontologien

Informationssytem Fkt. Repräsentation von Information, insbesondere semantischer Informationen ("Der Inhalt gehörter Wörter")

- System von Symbolen mit festgelegter Semantik (oft nicht Normal realisiert)
- natürlichsprachliche Semantik
- $FOL\,PK1$ (Prädikatenlogik der ersten Stufe)
- $T \subseteq FOL\ \alpha = (\alpha, (R^\sigma),\ R \in \mathbb{R}$
- $\alpha \models T$
- Vollständigkeitstheorem

Prinzipien der Wissensverarbeitung

- Normalerweise sehr stark "Logiklastig"

Allgemeines

- "Was ist Wissen?"
  - Erkenntnislehre
  - Ursprung des Wissens
  - Beziehung des Wissens zur Welt
  - Verifizierung von Wahrheit

Vorstellung Wissen kann man haben, weitergeben, usw.

- Teile des Wissens sind *a-priori* (Platon, Welt der Ideen)
- eine andere Art des Wissens kann man ohne jede praktische Anwendung und Erfahrung sofort anwenden (Geometrie)
- *a-posteriori* Wissen. Bspw. medizinisches Erfahrungswissen
- *Dialog Menon*[15] von Platon (Reclam Ausgabe)
- meistes Wissen ist abgeleitetes Wissen, das durch Schließen gewonnen wird

$\models \gamma_1$ benötigt Prämissen, oft "Weltwissen", das stillschweigend vorrausgesetzt wird

$\{\gamma_1, \gamma_n\} \models \varphi_1$

x_1 benötigt ebenfalls Ableitung

- dann unendlicher Regress, benötigt Abbruch; an dieser Stelle treten Axiome
- diese Anfangsbedingung (Axiome) setze ich als wahr, ob diese wahr sind, wird nicht gezeigt — es wird vorrausgesetzt
- abschließende Klärung nicht möglich
- $\mathfrak{D} \models Ax \qquad \forall x\, P(x) \qquad n \models P(1), \models P(2), \neg P(---)$ ↯
- best. Sätze können in endlich vielen Schritten widerlegt werden
- Bsp. auch durch Experimente
- damit sind Axiome auch eine "Glaubenssache"

[15] *http://ourworld.compuserve.com/homepages/KrausePlonka/literatur/TKP_sokr_Menon.htm*

Theoriebildung

- Verallgemeinerung aus empirischen Daten
- Bspw. Beobachten von Vögeln. Raabe ist Vogel.
- $S(r_1), S(r_2) \ldots, S(r_k)$ k-schwarze Raben

Generalisierung $\forall x\,(Raabe\,(x) \rightarrow S\,(x))$ aber; "mit Vorsicht zu genießen"

- $\forall x\,(Vogel\,(x) \rightarrow fliegt\,(x))$
- (Substitution u. Abtrennungsregel) $Vogel\,(c) \quad Subst.$
- $Vogel\,(c) \rightarrow fliegt\,(c) \;+\; fliegt\,(c)$
- $X \leq y$ und $X \models \mathfrak{e},$ so $Y \models \mathfrak{e}$ (Klassische Bedingung der Monotonie)
- wenn man die Wissensbasis erweitert, mit: $\{Pinguin\,(c)\}$ – dann Bedingung der Monotonie verletzt
- Theorie soll erklären und zum Verständnis der Welt beitragen
- soll Vorhersagen treffen
- Theorie benutzen zur Konstruktion
  - erklären, vorherzusagen, zu konstruieren
- euklidische Ebene (*a-priori* Wissen?)
- Euklid. Ebene:
  - $g \mid o_p$ Parallelenaxiom als Bsp. / eine gerade parallel
- Individualisierung – Schaffung individuelle Objekte und Situationen
- wie können solche Entitäten so wichtig sein?
- Bsp. postulieren[16] einer unendlichen Menge##
- X ist unendl., wenn eine $1-1-deutige$ Funktion
  - $j:\, x \rightarrow y$ und $y \lneqq x$
- Bsp. Newton / Inertialprinzip[17]
  - Ein Körper K ist entweder in Ruhe, oder bewegt sich geradlinig gleichförmig, wenn keine Kraft auf ihn wirkt

[16] *postulieren*. Etwas fordern oder zur Bedingung machen.
[17] Trägheitsprinzip

  - (diese banale "Bed." sind nicht herstellbar) *Rätsel*: Wie kommt man auf sowas?

- Bemerkungen zum Wahrheitsbegriff des Wissens
  - Wenn Axiom *wahr*, dann alles abgeleitete *wahr*
  - $\mathfrak{a} \models \vartheta \quad \forall (x)\ \vartheta (c)$ Beweis nicht möglich, aber kann widerlegt werden, durch Gegenbeispiel
  - Wenn $c \in \mathfrak{a}$, so d. $\mathfrak{a} \models \neg\vartheta (\mathfrak{c})$
  - $\circlearrowleft\ \mathfrak{a} \not\models \forall x \vartheta (x)$
  - $A \vdash \vartheta, \neg\vartheta$, dann A inkonsistent

Stufen der Wissensverarbeitung

1. Wissenserhebung
   a) Wissen liegt meist in *natürlichsprachlichen Beschreibungen* vor, aber auch Bilder sind Wissensquellen
   b) Messergebnisse die in Datenbanken vorliegen, usw.
2. Konzeptualisierung
   a) Wissen ausgehend von der Analyse des Merkmales; welche Objekte sind Relevant
      i. Welche *Objekte* sind relevant für die Domäne D?
      ii. Welche *Konzepte* und *Relationen* sollten eingeführt werden?
         A. $C_1 \dots C_n$ (Konzepte)
         B. Relationen C_i, C_j
         C. Grundobjekte von D
   b) Konzeptualisierungen sind nicht eindeutig

Axiomatische Methode

$C_1 =_{def} \varphi (G_1 \dots G_K)$

$G_1 =_{def} \varphi_1 \left(D_1^1, \dots, D_{K1}^1\right)$

$\circlearrowleft$ führt ebenfalls wieder zu unendlichem Regress

$\rightarrowtail$ `Abbruch`. Erklären "*primitiver Begriffe*"

Diese werden nicht durch explizite Def. erklärt, charakterisiert. Seien $P_1 \dots P_n$ primitive Begriffe, über die keine Festlegungen getroffen werden. Dann entspricht dies der Anm. von Axiomen über $P_1 \dots P_m$, die logisch gültig sind.

$\forall x (P (x) \vee \neg P (X))$ (Inhaltsleer)

$\rightarrowtail$ `Problem`. Einführung von Axiomen für $\{P_1 \dots P_m\}$. Keine tautologischen, sondern sollten inhaltliche Substanz haben.

1. Domäne
   a) Strecken der euklidischen Ebene
   b) Kongruenzrelation. $s \sim t$ gdw. s und t die gleiche Länge besitzen
   c) $\equiv$ Symbol für die Kongruenz, $=$ Identität
   d) $\alpha(Str., =^{\delta}, \equiv^{\delta'}) \mid (a,b) \in =^{\delta} \; gdw. \, a = b$
   e) $\sum = (\equiv, =)$, andere Möglichkeit $\sum = (\equiv, =, U)$
   f) $\forall x \, (U(x) \leftrightarrow x = x)$
   g) U einstelliges Prädikat, liefert nichts wesentliches, deswegen lässt man es in der Regel weg
2. Axiomatisierung
   a)

$$\begin{array}{rcl} Ax\,(\equiv, =) & \mid & \alpha \models \varphi \\ \varphi\,(\equiv, =) & \mid & \end{array}$$

   b) Wie findet man Axiome?

Bsp. *Tautologien suchen*

$x = x \vee \neg x = x \,|\, x \equiv x \rightarrow x \equiv x$

nächster Trick; finden sie Aussagen, die nicht aus dieser logischen Bed. folgen, aber in dem Standardmodell *wahr* sind.

$T_0 = \{x \equiv x \rightarrow x \equiv x\} \cup \forall x \, (u)$

$\exists X_1 X_2 \, (X_1 \neq X_2) = \varphi_0$

$T_0 \neq \varphi_0$, $\alpha \models \varphi_0$, $T_1 = T_0 \cup \{\varphi_0\}$

genau dann der Fall, wenn $Mod\,(T_0) \nsubseteq Mod\,(\varphi_0)$

$\alpha_0 = (\{\alpha\}, =^{\varphi}, \equiv^{\varphi})$, so gilt T_0 nicht in α_0 (hat nur ein Element)

$T_1 \, \exists X_1 X_2 X_2 \, (\bigwedge X_i \neq X_j) \quad T_1 \nvDash \varphi_1 \, \alpha \models \varphi_1$

(Modell, Modellklasse, usw.)

$\exists n :=$ es gibt wenigstens n viele Strecken

$T_{0D} = T_0 \cup \{\exists n / n < \omega\}$

Ex. φ: $T_\infty \nvDash \varphi$ und $\alpha \models \varphi$

$\psi_2 := \exists x_1 \dots x_n \, (\bigwedge \neg x_i \equiv x_j)$ unendl. viele Kongruenzklassen ex.

$S = T_\infty \cup \{\psi_n / n < \omega\}$

S_∞ lässt sich nicht erweitern.

Bsp. $A = \{a, b\} \quad A^* = \{a_1 \dots a_n / \dots\}$

$w \leq v := w$ ist Präfix von v.

Ist dies eine partielle Ordnung?

$$v \leq vv \leq w \wedge w \leq v \rightarrow v = w$$
$$v \leq w \wedge w \leq z \rightarrow v \leq z$$

- zu jedem Knoten gibt es zwei Nachfolger, diese sind unterschiedlich, sie unterscheiden sich an einer Stelle
  - Folie: Verschiedene Arten von einem Expertenwissen zu erwerben
  - Phasenorientiertes Vorgehen bei der Wissensmodellierung
  - $Szenario(1)$
  - 1. Wissenserhebung
    - (verbirgt Konzept)
    - $\underbrace{Ein\,Igel}_{C}\ \underbrace{ist\,ein}_{is-a}\ \underbrace{Säugetier}_{D=}$
  - 1. Wissenserhebung (2)

    berifft Taxonomie / Konzepte erscheinen als Klassen
    - Objektklassen/Instanzen der Schnupfendomäne
    - Konzepte brechen auf Konzepte ★berunkter am Ende ein der Texonomie stehen Konzepte, Schnitt zu den Individuen sollte möglich sein.

    Ausblick Ontologie.

Grundlagen der Formalen Ontologie

- Informatik-Ontologien spielt phil. nicht die zentrale Rolle
- Wissensbasen entwickeln
- Top-Level-Ontologie (Foundational Ontologien)
- befasst sich mit den allgemeinsten Begriffen der Welt
- Was ist das Ding und welchen Existenzmodus es hat, reales vs. intentionales Objekt
- Raum, Zeit, Objekt, Preozess, Individuum... sind allgemeine Begriffe

- Bsp. Onto-Med (IMISE)
- http://www.onto-med.de/
- General Formal Ontology (GFO)
- A Foundational Ontology Inegrating Objects and Processes (nicht abgeschlossen)
- andere Top-Level-Ontologien: DOLCE, SUMO, ★BWW.
- GFO einzige, die die Integration von Objekten und Prozessen leistet
- DOLCE/SUMO interpretierbar in GFO, umgekehrt funktioniert nicht
- SWOOP beruht auf Description Logic DL
- DL ziemlich ausdrucksschwacher Teil der FOL
- je ausdrucksfähiger, desto schwieriger Inferenzmethoden zur Verfügung zu stellen
- CL — Common Logic, ähnlich FOL, erinnern an $\lambda - Kalkuehl$ (sehr ausdrucksstark)
- möglichst große Ausdrucksfähigkeit vs. effiziente Inferenzsysteme
- GFO: 450 Axiome deren ★WSP-Freiheit zu zueigen ist, oder auch beseitigt
- Axiomatisierung der Grundbegriffe; Vollständigkeit, Entscheidbarkeit usw.
- $D - Domäne$ (=Situid) besser ist Klasse von Situiden
- ★Tirck-Presentiale durch Zeit ex. Persentuale? oder so. Perpetuanten
- (Bsp. für Situoide, Operation $Ont\,(D) = (L, \Sigma, \forall x\,(\Sigma))$ ontologische Signatur
- Begriffsvollständigkeit (möglich), wenn aber vollständig in Begriffen
- dann noch Vollständigkeit der Axiome
- Bsp. $R\,(X, Y) \quad \forall xy\,(R\,(X, Y) \wedge \neg R\,(X, Y))$ Tautologie die schon aus den Grundaxiomen d. Logik folgt, deshalb etwas inhaltsleer
- acht Kriterien für Wissensbasen
- Bsp. Unschärfe von Begriffen und nichtmonotones Schließen
- OBOedit (gerichtete Graphen anderer Kanten-Konzepte stehen)
- viele Aussagen, die hier enthalten sind mit ★— einfach zu lesen (★— nichtmonotones Schließen)

- der Experte kennt immer seine Ausnahmen
- Bsp. Alle Vögel können fliegen (es gibt aber Pinguine)
- Normale Inferenzsysteme erkennen dies aber nicht, musste mit aufgenommen weden, leider ungelöst

Semantik der Wissenrepräsentation

- Begriffe der natürlichen Sprache auf Normale Ebene zu heben
- Bedeutung für
  - Linguistik
  - $\begin{cases} \text{Kognitive Psychologie} \\ \text{Cognitive Science} \end{cases}$
- Arten der Fehlschlüsse, Bsp. manipulatorische Aussagen
- Definitionstheorem

Ontologie – gr. Ontos – Sein – Seindem

- Metaphysik von Aristoteles

1. Platon Parmenicedes (Aristoteles: Kategorienlehre, Metaphysik)
2. Spinoza (Ethik)
3. Leibniz (Monadologie[18])
4. Wittgenstein (Log. Tractatus)
5. Hegel (Phänomenologie des Geistes)

- Heidegger, Sartre

[18] Die *Monadologie* (von griechisch monas: Eins, Einheit) ist die von Gottfried Wilhelm Leibniz begründete Monadenlehre und Titel des Werkes von 1714, in dem er diese in 90 Paragraphen darlegt. Die dort erläuterte Monadologie ist die Lehre von den Monaden oder einfachen Substanzen bzw. letzten Elementen der Wirklichkeit und ist das zur Lösung metaphysischer Probleme dienende Kernstück der Philosophie Leibniz':

Die Urmonade ist Gott; alle anderen Monaden sind ihre Erzeugnisse; sie können nur von Gott vernichtet oder erschaffen werden und nicht von selbst entstehen oder vergehen. Die Welt besteht aus Aggregaten von vielen Monaden, die alle voneinander verschieden und jedoch insofern gleichsam als Entelechien autonom tätig sind, als sie Appetit auf und die Fähigkeit zur Perzeption aufweisen (Prinzip der Vielheit in der Einheit). Perzeptionen sind prinzipiell nicht durch bloß mechanische Gründe erklärbar: Selbst wenn man eine Maschine bauen könnte, die zu Perzeptionen befähigt wäre und die man betreten könnte, würde man im Inneren nur sich stoßende Teile vorfinden, niemals aber eine Erklärung für eine Perzeption. (aus Wikipedia)

RK Stufe:

1. Die Rose ist rot
2. Diese Rose ist rot
3. Die Rose ist eine Pflanze
4. Sokrates ist Mensch. Sokrates ist Weise.
5. $\begin{cases} \text{Das Rot dieser Rose ist dunkel} \\ \text{Das Rot dieses Pullovers geht ins violett} \end{cases}$
6. Die Weisheit d. Sokrates wurde uns durch die Dialoge des Platon überliefert
7. Diese Rose war vor zwei Wochen rot, ist jetzt aber braun
8. Das Rot dieser Rose ist dunkler als das Rot dieses Kugelschreibers
9. John ist größer als Mary
10. Dieses Chamäleon wurde in einer Stunde I-mal grün

-

1. $\forall x \left(R\left(x\right) \rightarrow R_0\left(x\right)\right)$ Rose $R,\ R_0$ rot sein

-

1. Sokrates ist Mensch.
   a) $M\left(x\right),\ S - Sokrates$
   b) $M\left(s\right)$ $[S :: M]$ Mensch und Sokrates ist Instanz von Mensch
   c) $\Longrightarrow s \in ext\left(M\right)$

-

1. $Rose\left(x\right) \rightarrow Rot\left(x\right)$ x ist rotes Ding, wir reden aber von einer Eigenschaft des Roten
2. $Rose\left(x\right) \rightarrow Rot\left(x\right)$
3. $Rose\left(x\right) \rightarrow Pflanze\left(x\right)$
4. $Rose(\textit{individuelle Eigenschaften werden benötigt})$
5. $Weisheit\left(S\right)$ — falsch um die Wahrheit des Sokrates zu beschreiben

...