

Beschreibungslogik

Vorlesung im Sommersemester 2017

Organisatorisches

- Form: K4, jeder vierte Termin ist Übung
Fragen und Diskussion jederzeit erwünscht.

- Vorlesungsmaterial:

Folien und Aufgabenblätter auf:

<http://tinyurl.com/ss17-bl>

Beispiele, Beweise, etc. an der Tafel bitte mitschreiben

Organisatorisches

- Zeit und Ort:

Mi 16–18 MZH 6210

Do 12–14 MZH 6210 (Ausnahme: am 20.4. GW2 B2880)

- Vortragender:

Thomas Schneider

Cartesium Raum 2.56

Tel. (218)-64432

ts@cs.uni-bremen.de

- Position im Curriculum: Modulbereich Theorie, Master-Basis
Master-Profilen KIKR und DMI

2

Prüfungsmodalitäten

Übungen + Fachgespräch:

- Übungsaufgaben jede 2. Woche (Kernaufgaben + Zusatzaufgaben)
- Werden in Kleingruppen (2–3 Pers.) bearbeitet, abgegeben und dann von mir korrigiert
- Lösungen für Übungsaufgaben werden von Euch präsentiert.
Ziel: wissenschaftl. Sprache und Präsentieren üben
- Es müssen insgesamt 50% der Punkte erreicht werden, und jede_r muss mindestens 1 Lösung präsentiert haben (Präsentation wird *nicht* gewertet).
- Fachgespräch am Ende des Semesters
- Erste Übung (und Abgabe) am 19.4.2016

Alternativ: Mündliche Prüfung

3

4

Wiederholungsregelungen

Fachgespräch/mündliche Prüfung nicht bestanden?

- 1 Wiederholungsversuch im selben Semester möglich, zu denselben Konditionen wie der 1. Versuch
- Wenn erforderlich, weitere Wiederholungsversuche:
 - als mündliche Prüfung in den folgenden 3 Semestern (je 1 Versuch pro Semester)
 - oder ggf. im SoSe 2018 erneute Teilnahme am Übungsbetrieb (wenn ich die Veranstaltung wieder halte)

Dies ist bereits Deine Wiederholung?

- Erneute Teilnahme am Übungsbetrieb ist möglich, zu den eben genannten Konditionen. Übungspunkte aus vergangenen Semestern werden dann *nicht* anerkannt.
- Bei Nutzung früherer eigener Lösungen: Quellenangabe erforderlich

5

Kapitel 1

Einleitung

7

Literatur

- The Description Logic Handbook, 2nd edition. Baader, Calvanese, McGuinness, Nardi, and Patel-Schneider (eds.), Cambridge University Press, 2007
- An Overview of Tableau Algorithms for Description Logics. Baader und Sattler. *Studia Logica*, 69:5-40, 2001
- Buchkapitel über Komplexitätsresultate in Stud.IP

6

Wissensrepräsentation

Wissensrepräsentation: allgemeines Hintergrundwissen dem Rechner verfügbar machen.

Grundlage für:

- „Intelligente Systeme“, die Daten nicht nur verarbeiten, sondern auch interpretieren können
- Zahlreiche Teildisziplinen der künstlichen Intelligenz, wie z.B. Robotik, automatisches Planen, usw.
- ein „semantisches Web“, in dem Suchmaschinen und andere Agenten Webinhalte interpretieren können (im Gegensatz zu schlüsselwortbasierter Suche)

Wissensrepräsentation – ein Beispiel

Betrachte folgende medizinische Datenbank:

Patient(p_1)	Patient(p_2)
diagnose(p_1, d_1)	diagnose(p_2, d_2)
Tachykardie(d_1)	Hypertonie(d_2)

Und die Anfrage $\text{Patient}(x) \wedge \exists y (\text{diagnose}(x, y) \wedge \text{Herz-Kreislauf-Erkrankung}(y))$

also: "gib mir alle Patienten mit Herz-Kreislauf-Erkrankungen"

Ein klassisches Datenbanksystem liefert keine Antworten, denn es fehlt das Wissen, dass Tachykardie und Hypertonie Herz-Kreislauf-Erkrankungen sind

Wissensrepräsentation

Entwicklung von **Formalimen**, mittels derer Wissen über die Welt in abstrakter Weise beschrieben werden kann und die effektiv verwendet werden können, um intelligente Anwendungen zu realisieren.

Wohldefinierte Syntax und Semantik.

Syntax: die Sprache, in der Wissen „aufgeschrieben“ wird
in dieser VL stets symbolisch und logikbasiert

Semantik: fixiert die Bedeutung des repräsentierten Wissens
in exakter, eindeutiger Weise

Wissensrepräsentation

Das Ziel der Wissensrepräsentation etwas konkreter:

Entwicklung von Formalismen, mittels derer Wissen über die Welt in abstrakter Weise beschrieben werden kann und die effektiv verwendet werden können, um intelligente Anwendungen zu realisieren.

[Ronald J. Brachman & Daniele Nardi 2003]

Dies ist anwendungsorientierte Sicht auf WR,
keine philosophische oder kognitionswissenschaftliche.



Wissensrepräsentation

Entwicklung von **Formalimen**, mittels derer Wissen über die Welt in abstrakter Weise beschrieben werden kann und die effektiv verwendet werden können, um intelligente Anwendungen zu realisieren.

Wohldefinierte Syntax und Semantik.

- erlaubt (weitgehende) Anwendungsunabhängigkeit
- wir wollen repräsentieren (=beschreiben),
nicht programmieren wie etwa in Prolog
- basiert auf (logischen) Strukturen

Wissensrepräsentation

Entwicklung von Formalismen, mittels derer **Wissen über die Welt** in abstrakter Weise beschrieben werden kann und die effektiv verwendet werden können, um intelligente Anwendungen zu realisieren.

Die Natur von Wissen ist schwierige (philosophische) Frage.

Wir beschränken uns auf konzeptuelles Wissen:

- Beschreibung des Konzeptes einer Hypertonie, einer Herz-Kreislauf-Erkrankung, einer Vorlesung, eines Studenten etc.
- Eine „was ist ein XYZ“ Beschreibung

Andere Arten von Wissen z.B. zeitliches Wissen, räumliches Wissen, prozedurales Wissen, Wissen über Wissen, etc.

13

Wissensrepräsentation

Entwicklung von Formalismen, mittels derer Wissen über die Welt **in abstrakter Weise** beschrieben werden kann und die effektiv verwendet werden können, um intelligente Anwendungen zu realisieren.

- Wissensrepräsentation bedeutet Abstraktion!
- Alle Aspekte eines Konzeptes im Detail zu identifizieren ist Aufgabe der Philosophie
- Nicht relevante Aspekte eines Konzeptes sollten nicht repräsentiert werden
- Damit ist eine bestimmte WR nicht unbedingt für alle Anwendungen adäquat (trotzdem deklarativ!)

14

Wissensrepräsentation

Entwicklung von Formalismen, mittels derer Wissen über die Welt in abstrakter Weise beschrieben werden kann und die effektiv verwendet werden können, um **intelligente Anwendungen** zu realisieren.

Intelligente Anwendungen können Informationen interpretieren und neue daraus ableiten.

- klassische Datenbank ist in diesem Sinne *nicht* intelligent: eingegebene Daten = ausgegebene Daten, nur Reorganisation
- Schlussfolgern (Reasoning) erforderlich: aus explizit gegebenem Wissen implizites ableiten
→ Fragebogen, Aufgabe 1

15

Wissensrepräsentation

Entwicklung von Formalismen, mittels derer Wissen über die Welt in abstrakter Weise beschrieben werden kann und die **effektiv verwendet werden können**, um intelligente Anwendungen zu realisieren.

Schlussfolgerungsprobleme sollten entscheidbar sein und von möglichst geringer (Berechnungs-)Komplexität

Wechselwirkung:

- geringe Komplexität und Entscheidbarkeit vs.
- benötigte Ausdruckstärke

Ideale Balance abhängig von Anwendung und Abstraktionsgrad, also nicht ein WR-Formalismus, sondern viele!

16

Wissensrepräsentation

- Prädikatenlogik erster Stufe (FO)
 - Formale Syntax, deklarative Semantik
 - Hohe, dennoch beschränkte Ausdrucksstärke
 - Sehr hohe Berechnungskomplexität: unentscheidbar
- Aussagenlogik (AL)
 - Formale Syntax, deklarative Semantik
 - sehr beschränkte Ausdrucksstärke
 - relativ geringe Berechnungskomplexität
 - effektive Schlussfolgerungssysteme („SAT-Solver“) verfügbar

BL ist Kompromiss bzgl. Ausdrucksstärke und Komplexität

17

Beschreibungslogik

Beschreibungslogik ist wichtige Familie von Formalismen zur Wissensrepräsentation

Andere Formalismen: siehe KI-Vorlesung

BLen sind eine Familie von Logiken

- zur Repräsentation konzeptuellen Wissens
- mit entscheidbaren Schlussfolgerungsproblemen
- für die viele effiziente Reasoner zur Verfügung stehen

18

Beschreibungslogik

Historisch hervorgegangen sind BLen aus

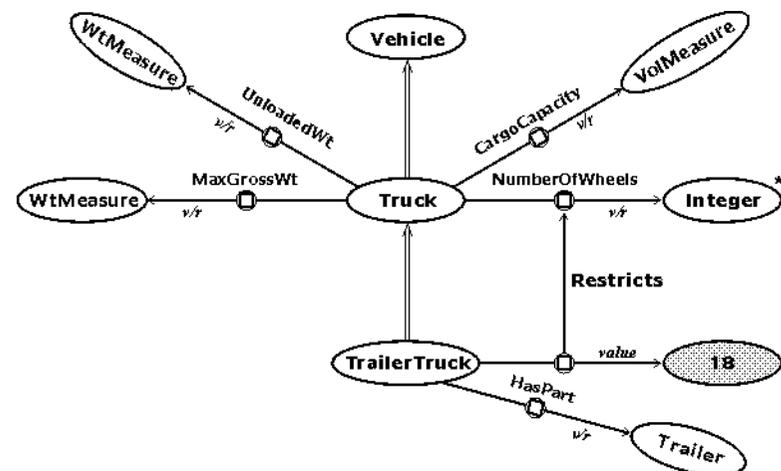
- structured inheritance networks [Ronald J. Brachman 1978]
- einem bekannten historischen WR-System namens KL-ONE [Ronald J. Brachman & James G. Schmolze 1985]

Diese frühen Formalismen

- waren nicht logikbasiert
- hatten keine formale Syntax und Semantik
- haben dennoch fast alle grundlegenden Ideen heutiger BLen eingeführt

19

KL-ONE Beispiel



20

Beschreibungslogik

Zentrale Elemente der WR mit Beschreibungslogik:

- Konzepte

Konzept beschreibt Klasse von Objekten, z.B. Bereich Universität:

Vorlesung, Student, Universität, Fachbereich, etc.

Können durch logische Ausdrücke (Formeln) beschrieben werden:

$\text{Mensch} \sqcap \exists \text{hört.Vorlesung}$

- Rollen

Sind binäre Relationen zwischen Objekten

hört, lehrt, istTeilVon, etc.

können (meist) nicht durch komplexe Ausdrücke beschrieben werden

21

Beschreibungslogik

Zentrale Elemente der WR mit Beschreibungslogik:

- TBoxen (terminologische Boxen)

Definieren Konzepte und setzen diese zueinander in Beziehung

Konzeptdefinition z.B.

$\text{Student} \equiv \text{Mensch} \sqcap \exists \text{hört.Vorlesung}$

Allgemeines Hintergrundwissen / Constraint z.B.

$\text{Student} \sqcap \text{Vorlesungssaal} \sqsubseteq \perp$

22

Beschreibungslogik

Zentrale Elemente der WR mit Beschreibungslogik:

- ABoxen (assertionale Boxen)

Beschreiben Individuen (= Objekte) und deren Eigenschaften

z.B.

$\text{Student}(\text{hans})$

$\text{Vorlesung} \sqcap \exists \text{hatThema.InformatikThema}(\text{blv})$

$\text{hört}(\text{hans}, \text{blv})$

→ Fragebogen, Aufgabe 2

23

Ontologien

Ein sehr populärer Ansatz zur Wissensrepräsentation ist die Verwendung von *Ontologien*

Ontologien dienen der Repräsentation von konzeptuellem Wissen, sehr ähnlich den TBoxen der Beschreibungslogik

Im Prinzip kann eine Ontologie in beliebiger (meist logischer) Sprache verfasst werden

Besonders populär ist aber OWL: the Web Ontology Language:

- Standardisiert vom W3C (World Wide Web Consortium)
- Zugeschnitten aufs Web: XML-Syntax etc.
- Im Wesentlichen eine Beschreibungslogik

24

Ontologien – Beispiel SNOMED

Im Gesundheitswesen ist standardisierte Terminologie wichtig.

Beispiel: Austausch medizinischer Daten zwischen Ärzten, Krankenhäusern, Krankenkassen, etc.

Es sollten

- dieselben Begriffe (Konzepte) verwendet werden
- verwendete Begriffe in derselben Weise verstanden werden

Es gibt verschiedene Standards:

- ICD-10: Int. Classification of Diseases and Related Health Problems
Publiziert von WHO, verwendet z.B. in Deutschland
- SNOMED CT: Standardized Nomenclature of Medicine, Clinical Terms
Verwendet z.B. in USA, GB, Schweden, Australien,...

25

Ontologien – Beispiel SNOMED

SNOMED

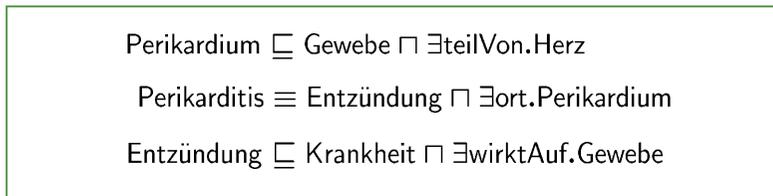
- von einem internationalen Nonprofit-Konsortium entwickelt (IHTSDO)
- standardisiert etwa 500.000 medizinische Begriffe aus allen Bereichen (Krankheiten, Behandlungen, Medikamente, etc.)
- wird regelmäßig aktualisiert, hat >40 Autoren gleichzeitig
- basiert auf einer Ontologie, in der diese Begriffe formal mittels einer Beschreibungslogik definiert werden

Die Mediziner erhalten natürlich keine logischen Ausdrücke sondern einen Katalog von Begriffen.

26

Ontologien – Beispiel SNOMED

Ontologiefragment



Katalog



27

The screenshot shows the SNOMED CT Browser interface. At the top, there are navigation buttons for 'Release: International Edition 20160131', 'Perspective: Full', 'Feedback', and 'About'. The main search area shows 'pericarditis' with 246 matches found in 0.214 seconds. A list of results includes 'Pericarditis', 'Acute pericarditis', 'Viral pericarditis', 'Pericarditis sicca', 'Fungal pericarditis', 'Uremic pericarditis', 'Q fever pericarditis', 'Chronic pericarditis', 'Uraemic pericarditis', 'Mycotic pericarditis', 'Familial pericarditis', 'Subacute pericarditis', 'Pyogenic pericarditis', and 'Adhesive pericarditis'. On the right, the 'Concept Details' panel for 'Pericarditis (disorder)' (SCTID: 3238004) is shown, including its parents (Disorder of pericardium, Inflammatory disorder of mediastinum, Inflammatory disorder of the cardiovascular system) and children (34 Children). A note indicates 'Associated morphology -> Inflammation' and 'Finding site -> Pericardial structure'.

Ausschnitt aus dem IHTSDO SNOMED CT Browser:

<http://browser.ihtsdotools.org>

28

Struktur Vorlesung

- Kapitel 1: Einleitung
- Kapitel 2: Grundlagen
- Kapitel 3: Ausdrucksstärke und Modellkonstruktionen
- Kapitel 4: Tableau-Algorithmen
- Kapitel 5: Komplexität
- Kapitel 6: Effiziente Beschreibungslogiken
- Kapitel 7: ABoxen und Anfragebeantwortung

29

Kapitelüberblick

Kapitel 3: Ausdrucksstärke und Modellkonstruktionen

- Nicht-Ausdrückbarkeit von Eigenschaften in \mathcal{ALC} und Erweiterungen
- Baummodelleigenschaft
- Endliche Modelleigenschaft

31

Kapitelüberblick

Kapitel 2: Grundlagen

- Basis-Beschreibungslogik \mathcal{ALC}
- TBoxen
- Schlussfolgerungsprobleme
- Erweiterungen von \mathcal{ALC}

30

Kapitelüberblick

Kapitel 4: Tableau-Algorithmen

Wollen zeigen, dass relevante Schlussfolgerungsprobleme entscheidbar sind und vertretbare Komplexität haben. Außerdem: *praxistaugliche* Algorithmen!

- Tableau-Verfahren und dessen Korrektheit
- Komplexitätsanalyse
- Optimierungen
- In manchen Fällen: Sicherstellen von Terminierung

PRAXIS-AUFGABE GEPLANT

32

Kapitelüberblick

Kapitel 5: Komplexität

Ziel: exakte Komplexitätsanalyse, Kennenlernen von Standardtechniken

- Obere Schranken mittels Verfahren aus Modallogik
- Untere Schranken mittels Reduktion von spieltheoret. Problemen
- Unentscheidbare Erweiterungen

33

Kapitelüberblick

Kapitel 7: ABoxen und Anfragebeantwortung

Ziel: Schlussfolgerungsprobleme mit Instanzdaten (ABoxen) und Hintergrundwissen (TBoxen) untersuchen

- Grundbegriffe
- Typen von Anfragen
- Konjunktive Anfragen mit ABoxen+TBoxen
- *Query Rewriting*: Nutzung von Datenbanktechnologie

35

Kapitelüberblick

Kapitel 6: Effiziente Beschreibungslogiken

Betrachten \mathcal{EL} :
geringere Ausdrucksstärke, aber effizientes Schlussfolgern
(SNOMED basiert im Wesentlichen auf \mathcal{EL} !)

- Grundbegriffe
- Polynomieller Algorithmus für Schlussfolgerungsprobleme
- Erweiterungen von \mathcal{EL} mit höherer Komplexität

KLEINES PROGRAMMIERPROJEKT GEPLANT

34

Beim nächsten Mal ...

- Kapitel 1: Einleitung
- Kapitel 2: Grundlagen
- Kapitel 3: Ausdrucksstärke und Modellkonstruktionen
- Kapitel 4: Tableau-Algorithmen
- Kapitel 5: Komplexität
- Kapitel 6: Effiziente Beschreibungslogiken
- Kapitel 7: ABoxen und Anfragebeantwortung

36

EMPFOHLEN

Daniele Nardi, Ronald J. Brachman:
An Introduction to Description Logics.
Description Logic Handbook 2007: 1-40

WETTERFÜHREND

Ronald J. Brachman:
What's in a Concept: Structural Foundations for Semantic Networks.
International Journal of Man-Machine Studies 9(2): 127-152 (1977)

Ronald J. Brachman, James G. Schmolze:
An Overview of the KL-ONE Knowledge Representation System.
Cognitive Science 9(2): 171-216 (1985)