

1.) Beschreiben Sie das Versionsraum-Lernverfahren!

Was passiert jetzt, wenn die Eingabe ein positives Beispiel ist? s. S. 178 !!

(d.h. alle Hypothesen in der Menge S und G müssen dann überprüft werden), ob $h(e)$, $e = \text{positiv}$, zutrifft. Falls ja, ist nichts zu tun, falls nein, ist die Hypothese s zu verallgemeinern, g aus G zu entfernen, weil s bzw. g zu speziell sind und g nicht weiter verallgemeinert werden kann.

2.) Was ist bedingte Wahrscheinlichkeit?

Können Sie das definieren? = Bayes'scher Satz

$$P(B|A) = \frac{P(B) \cdot P(A|B)}{P(A)} \quad \text{So nicht!!} \quad \text{Sondern so: } = \frac{P(A \cap B)}{P(A)}$$

Das genügt nicht! Hier haben Sie ja die bedingte Wahrscheinlichkeit durch sich selbst definiert! Was also ist bedingte Wahrscheinlichkeit? $P(B|A) = \frac{P(A \cap B)}{P(A)}$

286 3.) Können Sie das TMS beschreiben?

Was macht ein TMS?

Es überführt bei Vorliegen neuer Information ein maschinelles System von einem Zustand in einen anderen unter Erhaltung der Konsistenz und verwaltet so Abhängigkeiten zwischen Aussagen oder Daten.

4.) Was ist C_n ? $C_n(F)$ genau aufschreiben!

Wenn $C_n(F) = F$, also ein Fixpunkt ist, dann (C_n : Folgerungsoperator, \sim relation: $F \models G$ bzw. $F \models G$) heißt die Formelmeng F deduktiv abgeschlossen!

1845

S. 19

46ff

5.) Was ist $\text{Cn}(\text{Cn}(F))$?

das ist $= (\text{Cn}(F))$, wegen der deduktiven Abgeschlossenheit! Der Cn -Operator der klassischen Logik ist monoton!

1845

S. 18

43/45

6.) Was heißt $F \models G$?

Aus F folgt logisch G gdw. jede Interpretation, die F erfüllt, auch G erfüllt, (logische Folgerung)

1845

S. 144

7.) Was ist ein Modell eines TMS?

Eine Menge M von Knoten, $M \subseteq N$. Die Knoten in M sind in-Knoten, die anderen, restlichen Knoten in N out-Knoten

1845

S.

8.) Was ist eine Extension? $\Lambda_E = E$

$\Lambda_E = E$, also ein Fixpunkt.
Genauere Definition!

9.) Was ist eine Signatur?

- Signatur in der AL: eine Menge von Bezeichnern, genannt Aussagenvariablen.
- Signatur bei der PL 1?

1- oder n-stellige Fkt.-Symbole werden durch Funktionen auf den Objekten interpretiert

$\Sigma = \{ \text{func}, \text{pred} \}$ = die Menge der Funktions-Symbole und der Prädikaten-Symbole. Dabei hat jedes Symbol $s \in \text{Func} \cup \text{Pred}$ eine arität, Prädikate und Funktionen können 0, 1, 2, ... -stellig sein!

1-stellige Prädikatensymbole: Eigenschaften von Teilmengen des Universums U

Die Funktions- und Prädikaten-Symbole werden interpretiert, d.h., durch die Interpretation werden sie mit einer

2, 3, ..., n-stellige Prädikatensymbole: Relation zwischen den Elementen von U

Belegung versehen: 0-stellige Symbole werden durch ein Objekt = ein Element des Universums U interpretiert
0-stellige Prädikaten-Symbole werden wie Aussagenvariablen ~.

10.) Was ist Skolemisierung?

Verfahren genau angeben in der richtigen Reihenfolge! Pränexform herstellen, dann:
 $\exists x$ kann bei der Skolemisierung ersetzt werden durch eine Konstante/Vorangesetzt, diese Konstante kommt noch nicht vor in der Formel und \exists steht vor \forall oder alleine da: $\exists x P(x) \rightarrow P(a)$ a = neue Konstante

11.) JTMS: Was sind die IN- und OUT-Knoten? (In einem TMN, einem Paar $T = (N, J)$)

IN-Knoten: die zugehörige Aussage wird akzeptiert, als wahr angenommen. OUT-Knoten: nicht akzeptiert
 $\langle I | O \rightarrow n \rangle$ mit $O = \emptyset$ ist eine monotone Begründung, nicht-monoton sonst.
 Was sind die Begründungen?

Eine Begründung $J = \langle I | O \rightarrow n \rangle$ mit $I, O \subseteq N$ und $n \in N$, I = in-Liste, O = out-Liste, n = Konsequenz der Begründung, hat die Funktion einer Regel, zur Beschreibung von Zusammenhängen, z. B. von neuem Wissen, auch aus Nicht-Wissen!

13.) Was ist ein Default?

Eine Regel, die üblicherweise gilt, meistens gilt, Regel, die gilt, solange nicht explizit das Gegenteil bewiesen wird.
 Also: unsichere Regel, aber viel flexibler als sichere Regeln

14.) Was muß ein Modell eines TMS sein?

Es muß 1.) fundiert und 2.) abgeschlossen sein.

15.) Was bedeutet "fundiert"?

Eine Knotenmenge heißt fundiert, wenn es für jeden Knoten eine stützende Begründung gibt. Damit sind zirkuläre Begründungen ausgeschlossen, d. h. Beh., die sich selbst stützen.

16.) Konzeptlernen: Wie läuft das ab?

Was ist die Eingabe?

Eingabe: eine Menge von Trainingsbeispielen. Die Hypothesen über das Konzept bestehen aus einer Konjunktion von Constraints bezügl. der Attribut-Werte. Alle Constraints erfüllt durch ein Beispiel $e \rightarrow e$ ist ein positives Beispiel $\rightarrow h(e) = 1$

1845
S. 145 ff

S. 166 ff

S. 152 ff 17.) Was ist ein Entscheidungsbaum?

Er liefert zu Objekten (hier = Attribut/Wert-Paare) eine Entscheidung, welche Klasse es zuwar ordnen ist, z.B. Ja/Nein.
Induktives Lernen. Knoten = Attribute, Kante = deren Werte, Blatt = Klasse
= Abfragen...?

S. 156
S. 158
S. 161-165

18.) Wie findet man das wichtigste Attribut?

= Das Attribut, durch das die meisten Beispiele ^{genau} eig. klassifiziert werden.
Es soll die Beispiele möglichst gleichmäßig aufteilen in positive und negative Beispielmengen anhand H , = Entropie
Top-1. Dann Induction of Decision Trees: ID3 und C4.5 bzw. gain ratio (α)

S. 152 ff

19.) Was sind die Attribute?

Ein mit einem Attribut a markierter innerer Knoten eines DT repräsentiert eine Abfrage, welchen Wert das betrachtete Objekt für das Attribut a hat

S. 154

20.) Was steht an den Kanten?

Die zu a möglichen Attributwerte
Bei 2-wertigen Attributen wie z.B. "Warten" können die Bezeichnungen {ja, nein} auch als Kantenmarkierungen auftreten.

S. 153 ff

21.) Was ist die Eingabe?

Eine Menge von Attribut/Wert-Paaren = die Objekte, die zu klassifizieren sind.

S. 160
Reinhardt
DT

22.) Wie funktioniert der Algorithmus zum Finden des wichtigsten Attributs?

function DT S. 160 und Beschreibung S. 153

- 1.) Wahl des wichtigsten Attributs, d. h. mit dem die Anzahl der Bsp. auf E , die damit endgültig klassifiziert werden können, maximal ist
- 2.) Für jede durch die verschiedenen noch vorhandenen Attributwerte bestimmte Teilmenge von E (= eine neue Instanz des Lernproblems)
4 mögliche Fälle: ① Bsp.-Menge = \emptyset → kein Bsp. war in E → Default-Klassifizierung MajorityVal(E) ② alle Bsp. gleich klassifiziert → Ausgabe!
③ Attributmenge leer, aber noch Bsp. → Fehler in E ? Falsch Attribut? ④ Attributmenge $\neq \emptyset$ und $E \neq \emptyset$ → rekursiv weiter mit 1) Wahl des wichtigsten Attributs

S. 151 ff

23.) Data Mining, was ist das? Wofür dient das?

Ziel ist das Finden relevanter Zusammenhänge in großen Datenmengen. Relevanz: Assoziationsregeln, probabilistische Wenn-dann-Regeln, die Beziehungen zw. Dingen bzw. Merkmalen beschreiben. KDD = Automatisierung der Analyse von Daten, z.B. Marktanalyse, Kundenprofile
Wichtig: Verständlichkeit, Interaktion, große Datenmenge verarbeiten können, Skalierbarkeit,
Ziel: aus Daten interessante Muster, Strukturen, Abhängigkeiten u. dgl. extrahieren

S. 202
- 204

24.) Warenkorb - Analyse: was macht die?

Sie erstellt Assoziationsregeln aus der Datenbasis aller getätigten Transaktionen, wobei Tabelle erstellt wird, in der Spalte sind alle bei jeweils einer Transaktion gekauften Artikel, in den Zeilen steht, in welcher Transaktion der Artikel jeweils gekauft wurde. $\text{minsupp} = ?$ $\text{minconf} = ?$

S. 199

25.) Können Sie mir sagen, was der A priori - Algorithmus macht?

Er bestimmt zunächst die 1-elementigen Itemmengen und sucht dann in den folgenden Durchläufen in den Obermengen häufiger Itemmengen nach weiteren häufigen Itemmengen, bis keine mehr da ist.

- S. 197

26.) Was ist eine Item - Menge?

Eine beliebige Teilmenge $X \subseteq I$, wobei $I = \{i_1, i_2, \dots\}$ eine endliche Menge von Items ist, d.h. von Dingen, deren Beziehung zueinander beschrieben werden soll.

S. 197/198

27.) Wie lautet eine Assoziations - Regel?

$X \rightarrow Y$ mit $X, Y \subseteq I$ und $X \cap Y = \emptyset$, dann erfüllt eine Transaktion t diese Regel gdw. t alle in der Regel vorkommenden Items enthält, also wenn $X \cup Y \subseteq t$

S. 197/198

28.) Was ist ein Support?

Sei $D = \{t_1, t_2, \dots\}$ eine Menge von Transaktionen t , dann ist der Support (X) der relative Anteil von allen Transaktionen aus D , die X enthalten, also $|\{t \in D \mid X \subseteq t\}| / |D|$. D = die Datenbasis
 $\text{support}(X \rightarrow Y) = \text{support}(X \cup Y)$

S. 225/227

29.) Nennen Sie mir quantitative Methoden!

1.) Probabilistische Netzwerke (Markov-Netze für symmetrische, gerichtete - Bayes-Netze für einseitige Einflüsse). Knoten $\hat{=}$ Aussagen, Kanten $\hat{=}$ direkte Zusammenhänge. Kombination Graph + Wahrscheinlichkeitsverteilung \Rightarrow effiziente Methode zur Repräsentation und Verarbeitung unsicherer Abhängigkeiten.

2.) Fuzzy - Logik für die Modellierung weicher Übergänge, vage Aussagen, graduelle Abstufungen, subjektive Schätzwerte, ungenaue Messwerte

30.) Nennen Sie mir die Schritte, die für die Herstellung einer Klauselmengen nötig sind! Der Reihe nach!

4-stufiger Normalisierungsprozess:

- 1) Pränex - NF herstellen $Q_1 x_1 Q_2 x_2 \dots Q_n x_n [Matrix]$
- 2) konjunktive NF " $Q_1 x_1 Q_2 x_2 \dots Q_n x_n [Matrix \text{ in KNF}]$
- 3) Skolem - NF " \exists ersetzen durch z.B. Skolemkonstant
danach: \forall weglassen
- 4) Klausel - NF " aus der KNF-Matrix Formelmengen erzeugen $\{(\dots), (\dots), (\dots), (\dots)\}$

Herstellung einer Pränex - NF: Alle Negationen hinter, alle Quantoren schieben! Hierbei ändern sich evtl. mehrmals die Quantoren wg. $\neg \forall x (A) \rightarrow \exists x (\neg A)$ und $\neg \exists x (A) \rightarrow \forall x (\neg A)$. Verschiedene Quantoren hintereinander dürfen nie vertauscht werden! Denn $\forall x \exists y \neq \exists y \forall x$

31.) Was sind die Sicherheitsfaktoren bei Mycin? Was bedeutet $CF[-1, 1]$?

Durch die CF werden die Regeln mit einer reellen Zahl zwischen -1 und +1 bzgl. ihrer Gültigkeit quantifiziert. $CF(A \rightarrow B) \in [-1, 1]$

32.) Zählen Sie mir die Komponenten eines wissensbasierten Systems (WBS) auf!

- Wissensbasis
 - W-Verarbeitung-K.
 - W-Erwerb-K.
 - Erklärung-K.
 - Dialog-K.
- Ein logisches System ist durch 4 Komponenten gegeben: Signaturen: Σ , Formeln: $Formel(\Sigma)$, Interpretationen: $Int(\Sigma)$, Erfüllungrelation: \models_{Σ}
Hierbei ist eine Wissensbasis W eine Formelmengen über Σ

33.) Was ist "gain ratio"?

Das ist der normierte Informationsgewinn, also $gain(a)$ geteilt durch die split info(a), die sog. Entropie des Attributs a.
 $gain(a) = I(E) - I(E|a \text{ bekannt})$
Beim Lernalgorithmus TDIDT bzw. beim DT-Lernsystem ID3 wird damit das wichtigste Attribut gefunden.

S. 162/163
165

1845
S. 5144

34.) Was ist Entropie?

$H(P)$ misst die Unsicherheit bzgl. einer zu erwartenden $\omega \in \Omega$ bei beh. P .

Entropie = der mittlere Informationsgehalt H der tatsächlichen Antwort bei k möglichen verschiedenen Antworten, also $H(P(v_1), \dots, P(v_k))$ ist die Entropie von $\{v_1, \dots, v_k\} = -\sum_{i=1}^k P(v_i) \log_2 P(v_i)$. 1 Antwort = 1 bit insgesamt, also bei k möglichen Antworten und alle gleich wahrscheinlich

35.) Was ist eine atomare Formel?

PL1: ein positiver oder negativer Literal, z.B. $p(x)$, $\neg p(x)$
eine Formel der Form $P(t_1, t_2, \dots, t_n)$. Ein Literal ist eine atomare Formel oder die Negation einer atomaren Formel.
AL: eine AL-Formel, die nur aus einer Aussagevariablen besteht

36.) Die Klauselmenge ist nicht äquivalent

sondern nur erfüllbarkeitsäquivalent

zur ursprünglichen Formel! Woran

liegt das? Am Skolemisierungsschritt !! Wg. der Quantorenelimination!

Statt $\forall x \exists y$ jetzt $\forall x (\dots \rightarrow \dots f(x))$. Aus einer

Klausel = Disjunktion aus n Literalen

Formel wird eine \sim Menge

37.) Was ist Monotonie?

$A \subseteq B \Rightarrow Cn(A) \subseteq Cn(B)$ und

aus $F \subseteq H$ folgt $Cn(F) \subseteq Cn(H)$, d.h.

Theoreme von F sind stets Theoreme einer Erweiterung von F

38.) Was heißt "nichtmonotones Schließen"?

Nichtmonotone Logiken ermöglichen revidierbares

Schließen, d.h. Rücknahme nachträglich als falsch

erkannter Schlüsse und Ableitung alternativer Schlüsse.

39.) Formulieren Sie: "Alle Vögel haben

Flügel und Beine" Beim Allquantor: \rightarrow

$\forall x (V(x) \rightarrow F(x) \wedge B(x))$

Ebenso: "Einige Hunde haben Ohren"

$\exists x (H(x) \wedge O(x))$ Beim Existenzquantor: \wedge

Ebenso: "Jeder liebt jemanden"

$\forall x \exists y \text{ liebt}(x, y)$

S. 285

40.) Was ist der Unterschied zwischen monotonem und nicht-monotonem Schließen?

Monoton: Konsequenzoperation $C_n: 2^{\text{Form}} \rightarrow 2^{\text{Form}}$
 Aus $F \subseteq H$ folgt $C_n(F) \subseteq C_n(H)$

Nichtmonoton: Inferenzoperation $C: 2^{\text{Form}} \rightarrow 2^{\text{Form}}$
 Inferenzrelation (statt Folgerungsrelation), also \vdash statt \models , definiert durch $F \vdash G$ gdw. $G \in C(F)$
 zur Darstellung nichtmonotoner Ableitungen.

S. 162 ff

41.) Wie lernt der Entscheidungsbaum anhand der Trainingsmenge und der Testmenge?

Nach dem Verfahren DT zur Klassifikation der Beispiele. Bei leerer Bsp.-Menge wird anhand der Klassifikation des Elternknoten Majority Val(E) als Defaultwert zur Klassifikation verwendet. Bei gleicher {Ja, Nein}-Klassifikation aller Bsp. wird diese ausgegeben. Bei leerer Attributmengen, aber noch +/- Bsp. vorhanden ist evtl. falsches Bsp. in der Trainingsmenge oder man muß zusätzliche Attribute einführen. Ansonsten, also +/- Bsp. vorhanden und $A \neq \emptyset$, wählen wir das beste Attribut $\alpha \in A_{\text{gen}}$. Wichtigkeit: Bewertung des Lernerfolgs anhand einer Testmenge von separaten Bsp., ob diese korrekt klassifiziert werden. Verbesserung durch mehr Bsp. in die Trainingsmenge!

S. 297

42.) Was ist ein Bayesisches Netz?

Es modelliert kausale Zusammenhänge
 Knoten = betrachtete Aussagen, Kanten = direkte Zusammenhänge

1845
S. 143

43.) Was sind In-Mengen, Out-Mengen bei einem TMS? $\langle I, O, \rightarrow \rangle$??

Die in-Knoten $\in I = \text{in-Liste}$ sind die akzeptierten, geglaubten, gewußten Aussagen, die out-Knoten $\in O = \text{out-Liste}$ die nicht-geglaubten, nicht-akzeptierten Aussagen. Damit ein Knoten eines TMN durch eine Begründung $\langle I, O \rightarrow n \rangle$ als akzeptiert gilt, müssen alle Knoten der in-Liste I akzeptiert werden, es darf aber kein out-Knoten akzeptiert werden. Durch die Angabe von out-Knoten können also Schlüsse gezielt verhindert werden und damit nicht monotone Ableitungen ermöglicht werden. Im TMN sind die Knoten Rechtecke und die Begründungen Kreise.

Kanten-
 markierung
 = + für in-K.
 = - für out-K.

S. 286/287!!

44) Wie generiert man einen Entscheidungsbaum?

Ans. eine sog. Trainingsmenge (= Bsp.-Menge = Attribut/Wert-Paare mit Klassifikation als positiv/negativ)

Möglichst kleine Entscheidungsbaum finden:
das wichtigste Attribut zuerst testen, d.h. das, das am meisten zur Differenzierung beiträgt. Rekursiv, mit Lernalgorithmus DT, s. S. 159/160

Das wichtigste Attribut ist das mit dem größten Informationsgewinn (Entropie)

45) Was ist ein Kalkül? Eigenschaften eines Kalküls?

Kalkül = eine Menge von ^{Log.}Axiomen und Inferenzregeln

Axiome = Menge elementarer Tautologien oder Menge elementarer Widersprüche

Regeln = Menge von Vorschriften, wie aus Formeln weitere Formeln abgeleitet werden können. Bsp: MP, MT, \wedge -Einf., \wedge -Einf., \vdash

46) Was ist maschinelles Lernen?

47) Nennen Sie die Komponenten eines wissensbasierten Systems!

- Wissensbasis
 - regelmäßiges Wissen $\xrightarrow{\text{Experten}}$ permanentes W.
 - fallspezifisches Wissen $\xrightarrow{\text{Benutzer}}$ temporäres W.
- Wissensverarbeitungskomponente
 - ~ erwerbs ~ (Experten-Schnittstelle)
- Erklärungskomponente (Benutzer-Schnittstelle)
- Dialogkomponente (Experten- u. Benutzer-Interface)

48) Wie charakterisiert man Inferenz? (Peirce)

1.) Deduktion, stets wahr!

Schließen von (allgemein) gültigem regelhaftem Wissen auf einen speziellen Fall, was für alle Autos gilt, gilt auch für konkretes Auto.
 B: Batterie geladen nötig - Batterie = leer \rightarrow Auto starten unmöglich

2.) Induktion, nicht stets wahr!

Man erschließt aus einzelnen Tatsachen regelhaftes Wissen

B: Regel, aufgrund von Beobachtung: if Batterie leer then Auto stellt

3.) Abduktion, nicht stets wahr!

Man erschließt einen Sachverhalt aus einer Beobachtung, aufgrund regelhaften Wissens, cf. Diagnoseverfahren, so wird zwar eine mögliche Erklärung gefunden. Sie muß aber nicht sein.

B: W = "Auto mit leerer Batterie startet nicht" W = "Auto stellt" \rightarrow B: Batterie leer

49) Was ist Syntax? Was ist Semantik?

AL: Signatur, Junktoren, Formeln

AL: Belegung, $f \rightarrow \text{BOOL}$

PL: Signatur, \sim , \wedge , \vee , \exists

$\text{Int}(\Sigma)$,
 Wahrheitswert $\llbracket - \rrbracket$,
 $\llbracket - \rrbracket : \text{Formel } \Sigma \rightarrow \text{BOOL}$

PL-Interpretation: \mathcal{U} , Func, Pred,

$\llbracket A \rrbracket = 1(A)$ falls $A = \text{Atom}$

50) Was sind Regeln?

Formalisierte Konditionalität: "if A wahr, dann schließe, daß B." ^{wahr}
 Wenn immer gilt, daß $A \rightarrow B$, dann ist es eine deterministische Regel. 2 Bedingungen: Keine Disjunktion in der Prämisse und nur 1 Literal in der Konklusion.
 Regeln repräsentieren Zusammenhänge zw. Objekten (Mengen von ~)
 Regel = abstraktes Wissen, Faktum = konkretes Wissen
 Regeln sind gerichtet, d. h., die Prämisse muß wahr sein.
 Ausweg, falls nicht: MT zusätzlich implementieren.

51) Was geschieht bei der Inferenz mittels Regeln?

Eingabe: Vorwärtsverkettung: datengetriebene Inferenz d. h.,
 WB (Objekte, Pred) bekanntes fallspezifisches Wissen als Startpunkt, erfüllte Prä-
 Faktummengen \rightarrow Konklusion muß wahr sein \rightarrow abgeleitete Fakten als
 faktisches Wissen in den Inferenzprozess eingehend solange,
 bis keine neuen Fakten mehr abgeleitet werden können.
 Eingabe: WB + Faktummengen Ausgabe: Menge gefolgter Fakten (Kyklo-folgen)
 Rückwärtsverkettung: zielorientiert. Ausgehend vom Zielobjekt
 wird RB durchsucht nach Regeln mit Zielobjekt in der Konklusion
 BACKCHAIN, RULE EXE. Eingabe: RB, Faktummengen Ziel Ausgabe: Wert
 B: Prolog-Programmierung des Zielobjekts, falls ableitbar

51) Was ist ein Entscheidungsbaum?

52) Wie lautet Occam's Razor?

Bevorzuge die einfachste mit allen Beobachtungen konsistente Hypothese! d.h. hier: generiere den kleinsten möglichen Entscheidungsbaum!

53) Was sind Konzepte?

Def: Ein Konzept c ist eine 1-stellige Funktion über einer Grundmenge M von Beispielen/Instanzen, nämlich $c: M \rightarrow \{0, 1\}$. Für $x \in M$ gilt: x gehört zum Konzept gdw. $c(x) = 1$, x gehört nicht zum Konzept gdw. $c(x) = 0$.

Strukturierung des Konzeptraumes anhand der "Spezieller-als"-Relation. Suchraum = Menge aller möglichen Hypothesen. Wenn die Extension von h_1 eine Teilmenge der Extension von h_2 ist, dann gilt $h_1 \leq h_2$. Also ist \leq eine partielle, d.h. reflexive, transitive, antisymmetrische Ordnungsrelation auf der Menge der Konzepte, d.h. auf dem Suchraum.

54) Bei Terminierung 3 mögliche Situationen:

- 1.) S leer und/oder G leer $\Rightarrow V_B$ ist zu $\{\}$ kollabiert, d.h. es gibt keine konsistente Hypothese für die Trainingsbeispiele im gegebenen Hypothesenraum L_C
- 2.) S und G sind identisch, also $S = G = \{h\}$, d.h. h ist das einzige Konzept aus L_C das konsistent bzgl. Trainingsbeispiele ist.
- 3.) Normalfall, S und G enthalten verschiedene h . Diese sind alle konsistent.

55) Was kann es bedeuten, wenn der Versionsraum kollabiert?

- 1) Fehler in der Beispielmenge, d. h. ein ^(negativ)positives Beispiel wird dem Lernverfahren als ^(positiv)negatives Beispiel präsentiert oder
- 2) nicht ausreichend mächtige Konzeptsprache L_C , um das Zielkonzept auszudrücken

222 56) Was ist fallbasiertes Schließen? (Lernen aus Erfahrung)

Statt Regelbasis hier eine Fallbasis, in der frühere Erfahrungen gespeichert sind. Lösung neuer Probleme durch Herausuchen des relevantesten Falls aus der CB und Adaption der Lösung. Schließen ist hier kein regelbasierter, sondern ein erinnerungsbasierter Prozess. Formaler Ansatz, Erfahrungen zu verorten. Motto "Ähnliche Probleme - ähnliche Lösungen", Bei neuem Problem gleiche Typ.

233 57) Was ist das Indexvokabular?

242

58) Wie sieht ein Default aus?

Prämisse, Konklusion, Begründungen. Die Prämisse muss erfüllt sein, die Begründung muss nur mit der momentanen Wissensbasis sein.

$$\delta = \frac{\psi: \psi_1, \dots, \psi_n}{\chi}$$
 mit $\psi = \text{prerequisite}$, $\{\psi_1, \dots, \psi_n\} = \text{justifications}$
 $\chi = \text{consequent}$. Extension $E = \text{Erweiterung der \{Fakten\} um akzeptable Thesen.}$

59) Wofür ist der Default gut?

Er ersetzt die explizite Aufführung / Aufzählung von Ausnahmen, was ja auch oft schwierig ist, da die Menge der Ausnahmen meistens unendlich ist. Default-Theorie $T = (W, \Delta)$ mit $W = \text{Fakten, Axiome}$, $\Delta = \text{Defaults}$. Die Fakten sind als PL1-Formeln gesetzt.

60) Was heißt nicht-monoton?

Das heißt, daß neue Information nicht zu einer Erweiterung der Menge der Folgerungen, sondern zu einer Revidierung schon gezogener Schlüsse führen kann.

61) Was ist nichtmonotones Schließen?

Das ist eine Inferenzoperation $C: 2^{\text{Form}} \rightarrow 2^{\text{Form}}$, die auch die Konflikte zwischen altem und neuem Wissen berücksichtigt und verarbeiten kann, z.B. mittels Defaults oder als TMS und so vermeidet, daß die Wissensbasis inkonsistent wird. Nichtmonotone Regeln = unsichere Regeln im TMS u. Petri

4895
S. 204

62) Erklären Sie die Operation Δ

4895
S. 154

63) Welche justifications sind bzw. heißen monoton?

Die mit leave out - Liste, also $O = \emptyset$

64) Was sind Fuzzy-Mengen?

Eine Fuzzy-Menge μ ist eine Abbildung $\mu: X \rightarrow [0, 1]$ der Menge X in das Einheitsintervall. $X = \text{Grundmenge/Universum}$. μ heißt normiert, wenn $\exists x \in X. \mu(x) = 1$. Menge aller Fuzzy-Mengen von $X = \mathcal{F}(X)$. Wenn μ normiert ist, heißt das, es gibt im Universum X mindestens ein Element mit maximalem Zugehörigkeitsgrad. Die Fkt $\mu_{\text{groß}}$ z.B. ist eine Fuzzy-Menge. Mit Fuzzy-Mengen kann man wohl Aussagen, Graduelle, ungenaue Maße,

65) Wie sind Extensions bei einer Fuzzy-Menge definiert?

Sei $\Phi: X^n \rightarrow Y$ eine Abbildung und X, Y Mengen. Dann ist $\hat{\Phi}: (\mathcal{F}(X))^n \rightarrow \mathcal{F}(Y)$ die Extension von Φ , definiert durch:

$$\hat{\Phi}(\mu_1, \dots, \mu_n)(y) := \sup(\min\{\mu_1(x_1), \dots, \mu_n(x_n)\} \mid \text{Abb. fuzifizieren möglich} \mid (x_1, \dots, x_n) \in X^n \text{ und } y = \Phi(x_1, \dots, x_n)\}$$

64) Was ist Modus ponens? Modus tollens?

14

236 65) Erläutern Sie die 4 Arbeitsschritte eines CBR!

RETRIEVE: wird immer vom Rechner durchgeführt, anhand der Indexierung der Fälle. Probleme bei der Selektion: 1. Suchproblem, 2. Ähnlichkeitsprobleme. Grobruche, dann Feinsuche

REUSE: Nach Abgleich mit dem neuen Fall wird eine Lösung generiert für das aktuelle Problem.

REVISE: provisorische Lösung evaluieren durch Vgl. mit ähnl. Fällen aus der CB, ggf. Korrektur, Reparatur, Tests, Simulation

RETAIN: neuen, gelösten Fall in die CB aufnehmen, ggf. Indexvokabular modifizieren, CB umorganisieren

66) Welche besonderen Eigenschaften sollte ein Modell haben?

Anpassung seiner Parameter an die gegebenen Daten und die Zielsetzung des KDD-Prozesses. Quantifizierung der Güte eines Modells mit der maximum likelihood als Wahrscheinlichkeit eines Modells bei gegebener Datenbasis.

258 67) Was ist ein Truth Maintenance System?

Ein JTMS oder ATMS arbeitet mit nichtmonotonen Regeln bzw. verwaltet Annahmen, unter denen eine Aussage ableitbar ist. Immer gekoppelt mit einer Inferenzkomponente, die klassisch-logische Schlüsse vollzieht. Mit einem TMS sind nichtmonotone Ableitungen möglich.

68) Wie funktioniert Konzeptlernen?

Aus einer Menge von Beispielen, die als positive oder negative Bsp./Instanzen markiert sind, für ein zu erlernendes Konzept, soll automatisch eine allgemeine Definition dieses Konzepts c generiert werden. Dieses Lernen kann aufgefasst werden als Suchvorgang im Raum aller möglichen Hypothesen. Diese sind als Konjunktionen von Constraints repräsentiert. 3 Fälle sind für jedes mittels dieser Constraints eingeschränkte Attribut möglich: "?", " \emptyset ", "Fußball o. dgl." für "jedes, kein, nur dieses Wert akzeptabel".

184 69) Was ist ein Merkmalsbaum?

252 Ein Hilfsmittel, um Attributwerte hierarchisch zu strukturieren, um nicht immer gleich zum allgemeinsten "?" Constraint verallgemeinern zu müssen, wenn 2 verschiedene Attributwerte zusammentreffen. Die Konzeptbeschreibung besteht jetzt aus Paaren $[A, B]$, wobei A und B Merkmale aus 2 M.-Bäumen sind. Die Abb. $[A, B]: \text{Beispielsprache} \rightarrow \{0, 1\}$ ist definiert durch: $[A, B]$ ergibt für ein Bsp. (a, b) $1 \Leftrightarrow A \text{ im Baum } \geq a, B \text{ im Baum } \geq b$

70) Wie läuft die Verkettung bei MYCIN?

Sicherheitsfaktoren für die Propagation durch das Regelnetzwerk. Propagationsregeln für \wedge , \vee , serielle, parallele Kombination, bei Konjunktionsbedingung für die 4. Regel, daß $CF[A, \Gamma]$ für keine Aussage A und keine Hypothese-Teilmenge $\Gamma = 1$ sein darf, ebenso wenig darf $CF[A, \Delta] = -1$ sein. $CF =$ Grad des Glaubens an eine Hypothese. $CF=1$, definitiv wahr, $CF=-1$ = falsch. Prinzipiell arbeitet die MYCIN-Inferenzkomponente wie ein rückwärtsverkettendes Regel-Interpreter. $CF=0$: indifferente Haltung. 4 Propagationsregeln. $CF > 0$, wenn A B unterstützt, sonst $CF < 0$.

71) Quantitative Ansätze zur Verarbeitung bzw. WR?

Wegweisendes Beispiel hierfür war MYCIN. Den Aussagen bzw. den Formeln werden numerische Größen zugeordnet, die den Grad ihrer Gewissheit, ihres Einflusses, ihrer Zugehörigkeit zu einer bestimmten Menge u.dgl. ausdrücken. Dazu noch entsprechende Verfahren, um aus diesen Größen neues - quantifiziertes - Wissen abzuleiten.

72) Was ist ein Sicherheitsfaktor?

CF quantifiziert die Gültigkeit einer Regel mit einer reellen Zahl $\in [-1, 1]$, die also angibt, wie sicher B ist, wenn A wahr ist. $CF(A \rightarrow B) \in [-1, 1]$

$CF(A \rightarrow B) < 0 \hat{=} A$ liefert Evidenz gegen B

$CF(A \rightarrow B) > 0 \hat{=} \text{" " " für } B$

$CF(A \rightarrow B) = 1 \hat{=} B$ ist definitiv wahr

$CF(A \rightarrow B) = -1 \hat{=} B$ ist definitiv falsch

Gefragt wird hier nicht mehr, ob ein Zielobjekt wahr ist, sondern mit welcher Gewissheit. Es werden von MYCIN also Evidenzen gesammelt, um das Ziel zu beweisen.

CF s sind nur eine heuristische Methode, also nicht theoretisch fundiert, aber sehr erfolgreich.

Propagationsregeln: Konjunktion $CF[A \wedge B] = \min\{CF[A], CF[B]\}$
Disjunktion $CF[A \vee B] = \max\{CF[A], CF[B]\}$

73) Was ist die Erklärungskomponente von MYCIN?

- 1.) QA-Modul, beantwortet einfache Fragen bzgl.
- 2.) Reasoning status checker, aufgerufen mit HOW? WHY?

74) Was ist eine Instanz, beim Konzeptlernen?

Für ein Konzept c , also eine 1-stellige Funktion $c: M \rightarrow \{0, 1\}$ über einer Grundmenge M ist $x \in M$ eine Instanz. Die Elemente von M heißen also Beispiele oder Instanzen. Für $c(x) = 1$ sagen wir " x erfüllt c , x ist ein positives Beispiel von c ", für $c(x) = 0$ sagen wir " x gehört nicht zum Konzept c , x ist ein negatives Beispiel".

75) Was ist eine Extension?

Die Menge aller positiven Instanzen für c , also alle x mit $c(x) = 1$, das ist die Extension von c .

76) Was ist die Eingabe beim Konzeptlernen?

L_E und L_C und das Zielkonzept $c \in L_C$ und $P \subseteq L_E$ und $N \subseteq L_E$, also $N \cup P$ als Trainingsmenge mit $c(p) = 1$ mit $c(n) = 0$, die in der Beispielsprache L_E beschriebenen Beispiele

Hierbei dürfen P und N keine Fehler enthalten und L_C , also die Konzeptsprache, bestehend aus den Konzeptbeschreibungen $k: L_E \rightarrow \{0, 1\}$ muß mächtig genug sein, um c ausdrücken zu können. Induktives Lernen! Das anhand der Beispielmengen vom Konzeptlernverfahren gelieferte Konzept h muß vollständig und korrekt bzgl. der gegebenen Beispielmengen sein. Übereinstimmung des gelernten Konzepts h mit dem zu erlernenden Konzept c kann nicht garantiert werden!

77) Was ist das Ziel des Konzeptlernens?

17

Bestimmung eines Konzeptes c aus der Konzeptsprache L_C so daß $h(e) = c(e)$ für alle Beispiele $e \in L_E$ ist.

78) Was ist Kandidaten - Elimination?

Als initialer Hypothesenraum H wird L_C genommen.

$L_C = \text{alle Konzepte } \langle \dots, \dots, \dots, \dots, \dots \rangle$

Bei jedem neuen Beispiel werden aus H alle Hypothesen entfernt, die nicht mit der vorgegebenen Klassifikation von diesem Beispiel übereinstimmen.

79) Wie kann der Hypothesenraum nach durchsucht werden?

Unter Ausnutzung der "spezieller-als" - Relation auf der Menge L_C , mit der speziellsten Hypothese als initialem h , h wird dann bei jedem neuen positiven Beispiel e , das noch nicht von h abgedeckt wird, soweit verallgemeinert, daß auch e mit abgedeckt wird. Resultat ist schließlich eine speziellste Hypothese, die mit den gegebenen Beispielen konsistent ist.

Oder umgekehrt: Beginn mit dem allgemeinsten h , h wird dann bei jedem neuen negativen Beispiel e , das fälschlicherweise von h abgedeckt wird, soweit spezialisiert, daß e nicht mehr mit abgedeckt wird. Resultat ist eine allgemeinste Hypothese, die konsistent ist mit den gegebenen Beispielen.

80) Was versteht man unter spezieller bzw. allgemeiner Generalisierung?

Hypothese h ist speziellste Generalisierung einer Beispielmenge B , wenn 1.) h ist vollständig und korrekt bzgl. B .

2.) $\neg \exists h' < h$, das vollständig und korrekt bzgl. B ist.

Hypothese h ist allgemeinste Generalisierung, wenn 1.) gilt und $\neg \exists h' . h < h' , \dots$

81) Was ist ein Versionsraum?

$V_B = \{h \in L_C \mid h \text{ ist korrekt und vollständig bzgl. } B\}$
 ist der Versionsraum der Beispielmenge B (Trainingsmenge)

82) Wie funktioniert das Versionsraumlernverfahren?

Grundidee: statt Auswahl einzelner Hypothesen und Rücknahme dieser Auswahl bei Widersprüchen sind ständig alle noch möglichen Hypothesen als Menge dargestellt. Prinzip \approx Kandidaten-Elimination, aber statt direkter Repräsentation der div. Hypothesen hier die Repräsentation durch 2 Begrenzungsmengen S und G , also kompakter.

83) Repräsentationstheorem für Versionsräume?

$$V_B = \{h \in L_C \mid \exists s \in S \exists g \in G (s \leq h \leq g)\}$$

84) Versionsraum - Lernverfahren konkret beschreiben?

Für jedes neue Beispiel $e \notin B$ (B = die Menge der schon verarbeiteten Beispiele!) müssen S und G als Repräsentation des V_B überprüft und ggf. angepasst werden. Falls eine Hypothese $h \in S \cup G$ mit e übereinstimmt, d.h. falls $h(e) = 1$ für $e = +\text{Bsp.}$ und $h(e) = 0$ für $e = -\text{Bsp.}$, ist nichts zu tun. Falls nicht, 2 Fälle möglich: $h(e) = 0$ für $e = +\text{Bsp.}$ oder $h(e) = 1$ für $e = -\text{Bsp.}$

85) Was heißt, ein Konzept ist konsistent?

Konzept $c: B \rightarrow \{0, 1\}$ ist korrekt und vollständig = konsistent bzgl. B , wenn alle positiven Beispiele abgedeckt werden und kein negatives, also wenn $\forall b \in B. c(b) = 1$ für b pos. (vollständig) und $\forall b \in B. c(b) = 0$ für b neg. (korrekt)

86) Welche Lernalgorithmen basieren auf dem Verfahren DT?

Die TDIDT-Verfahren. Hierbei zentral: Attributauswahl so, daß die Beispielmenge möglichst genau in die Teilmengen des positiven bzw. negativen Beispiels aufgeteilt wird. ID3 und C4.5 anhand von H , von $I(E)$, von $\text{gain}(a)$, von $\text{gain ratio}(a)$ und $\text{split info}(a)$.

Das wichtigste Attribut ist das, das am meisten zur Differenzierung beiträgt. "Wichtig" ist hierbei relativ, das hängt von der jeweiligen (Rest-)Beispielmenge ab!

87) Was ist eine deduktiv abgeschlossene Theorie?

Ein Fixpunkt des Operators C_n , denn deduktiv abgeschlossen heißt, daß $C_n(F) = F$, daß also die C_n -Anwendung keine neuen Erkenntnisse mehr bringt, daß also jede C_n -Anwendung auf diese Formelmengen bzw. Theorie wieder zur selben Formelmengen / Theorie führt, daß also die Folgerung aus einer Formelmengen die Formelmengen selbst ist.

88) Wie kommt man zu einer Klauselmengen?

- ① Formel in bereinigte Form bringen, d.h. ggf. gebundenes Umbenennung von Variablen so daß hinter jedem \forall, \exists eine andere Variable steht
- ② Beseitigung der Implikatoren \Rightarrow und \Leftarrow durch Äquivalenzen $\neg \dots \vee$
- ③ Negationszeichen mittels de Morgan und doppelte Negation ganz nach innen schieben
- ④ Quantoren ganz nach außen schieben, ①-④ \rightarrow vorwiegend technische NF
- ⑤ Skolemisierung, d.h. alle \exists raus \Rightarrow Formel erfüllbarkeitsäquivalent
- ⑥ Alle Allquantoren weglassen, übrig bleibt die Matrixformel
- ⑦ Matrixformel in KNF = Disjunktion von Literalen bringen
- ⑧ daraus die Mengendarstellung $\{\{L_{1,1}, L_{1,2}, \dots\}, \dots, \{L_{m,1}, L_{m,2}, \dots\}\}$ als Menge von Klauseln

89) Was ist ein Unifikator? Unifikation? mgu?

Eine Substitution σ heißt Unifikator der Terme s und t , wenn $\sigma(s) = \sigma(t)$ gilt. In diesem Fall sind s und t unifizierbar. Nicht unifizierbar sind die Terme $f(x, b)$ und $f(a, c)$, aber folgende Terme sind unifizierbar:
 $t_1 = f(x, g(a, y))$ $t_2 = f(b, z)$

Unifikatoren: $\sigma = \{x/b \quad y/a \quad z/g(a, a)\}$ $\sigma' = \{y, a\}$
 $\mu = \{x/b \quad z/g(a, y)\} = mgu$

Für unifizierbare Terme gibt es immer auch einen allgemeinsten Unifikator (mgu). Ein Unifikator μ heißt mgu, wenn es zu jedem Unifikator σ von s und t eine Substitution σ' mit $\sigma = \sigma' \circ \mu$ gibt. $\sigma(t)$ heißt Instanz des Terms. Instantiation von Variablen. Bei der Unifikation werden Terme gleich gemacht, wichtig für Resolution.

90) Was ist der Resolutionskalkül?

Die einzige Regel ist die Resolutionsregel, wobei die Klauseln $\{L, K_1, \dots, K_m\}$ und $\{\neg L, M_1, \dots, M_n\}$ Elternklauseln heißen und die abgeleitete Klausel $\{K_1, \dots, K_m, M_1, \dots, M_n\}$ Resolvente. Die Literale $L, \neg L$ heißen Resolutionslitterale.

Ziel der Resolution: Widerspruch ableiten. Dieser elementare Widerspruch wird durch die leere Klausel \square repräsentiert.

Ziel: Die Erfüllbarkeit einer Formel F oder einer Formelmenge F zu beweisen. Wie? Durch Hinzufügen von $\neg F$ bzw. $\neg F$ zur Formel A bzw. Formelmenge A . Gelingt der Beweis, dass $A/\neg F/\neg F$ bzw. $A/A \wedge \neg F/\neg F$ unerfüllbar sind, dann ist $A \models F$, $A \models F$, $A \models F$ u.dgl. bewiesen.

91) Resolutionalkül korrekt und widerlegungsvollständig?

$F \models G$ gdw. $F \wedge \neg G \vdash_R \square$ $\square =$ unerfüllbare Formel
d.h.:

Um zu zeigen, dass die Formel G aus einer gegebenen Formelmeng F logisch folgt, wird $\neg G$ zu F hinzugefügt. Wenn aus $F \wedge \neg G$ \square ableitbar ist, ist $F \wedge \neg G$ unerfüllbar, dann folgt aber G logisch aus F (Ableitung eines Widerspruchs zum Beweis der Unerfüllbarkeit einer Formelmeng). Da mit R aus jeder unerfüllbaren Formel ein elementarer Widerspruch abgeleitet werden kann, ist der Resolutionalkül widerlegungsvollständig
korrekt: $F \vdash G \Rightarrow F \models G$

vollständig: $F \models G \Rightarrow F \vdash G$

widerlegungs-
vollständig: $F \models G \Rightarrow F \wedge \neg G \vdash \square$

92) Was ist eine Interpretation in der PL 1? ($\Sigma = \text{Func}, \text{Pred}$)

Eine Σ -Interpretation $I = (U_I, \text{Func}_I, \text{Pred}_I)$ besteht aus der nichtleeren Menge U_I , einer Menge Func_I von Funktionen, einer Menge Pred_I von Relationen, mit Stelligkeit n jeweils

93) Was ist die Menge der Terme, $\text{Term}_\Sigma(V)$? ($\Sigma = \text{Func}, \text{Pred}$)

$\text{Term}_\Sigma(V)$ ist die kleinste Menge, die enthält:

(1) x falls $x \in V$

(2) c falls $c \in \text{Func}$ und c hat Stelligkeit 0

(3) $f(t_1, \dots, t_n)$ falls $f \in \text{Func}$ hat Stelligkeit $n > 0$ und $t_1, \dots, t_n \in \text{Term}_\Sigma(V)$

Ein Grundterm über Σ ist ein Term ohne Variablen.

$\text{Term}_\Sigma(\emptyset)$ bezeichnet die Menge der Grundterme