

Einführung in die KI

Prof. Dr. sc. Hans-Dieter Burkhard
Vorlesung Winter-Semester 2003/04

Nichtmonotones Schließen

Umgang mit unvollständigem Wissen

- Nicht-Monotones Schließen
- Negation, „Nicht-Wissen“
- Unvollständigkeit
- Abgeschlossenheits-Annahmen
- Default-Schließen
- Revision/Truth Maintenance Systeme (TMS)

Verwandte Gebiete:

- Modellierung als unsicheres Wissen
 - Wahrscheinlichkeiten für Annahmen
 - Modale Logik
- Modellierung als unscharfes Wissen

Unvollständigkeit des Wissens

Entscheiden/Handeln trotz unvollständiger Information.
Entscheiden/Handeln trotz inkonsistenter Information.

Rationalität:
Mit angemessenem Aufwand erfolgversprechende
Entscheidungen treffen

Natürliche Sprache:

- Mitteilung unvollständiger Information.
- Beschränkung auf Wesentliches bzw. Allgemeines.
- Vertrauen bzgl. Mitteilung von Ausnahmen.

Ökonomie von Beschreibungen

Klassische Logik

Spezifisches Verfahren.
Formal einfach.

Monotonie: FI und Abl sind monoton:

$$X \subseteq Y \Rightarrow \text{FI}(X) \subseteq \text{FI}(Y)$$

$$X \subseteq Y \Rightarrow \text{Abl}(X) \subseteq \text{Abl}(Y)$$

Vorteil oder Nachteil?

Ableitungen: Mehr Axiome \Rightarrow mehr Sätze.

Folgerungen:

F folgt aus X gdw. jedes Modell von X ist Modell von F.

Mehr Axiome \Rightarrow weniger Modelle \Rightarrow mehr Folgerungen.

- Explizite Unvollständigkeit (aufwändig).

$$H_0(x) \wedge \neg \text{Ausnahme}(x) \rightarrow H_{00}(x)$$

$$\text{Ausnahme}(x) = A_1(x) \wedge A_2(x) \wedge A_3(x) \wedge \dots$$

- Inkonsistenz nicht darstellbar. $\text{Th}(H \wedge \neg H) = \text{ausd}$

Nicht-Monotonie

Eigentlich Alltags-Verfahren:
Einfach bzgl. Aufwand.
Formal kompliziert.

Natürliches Vorgehen ist nicht-monoton

Umgang mit unvollständige Ausgangsinformation:

- „solange nichts weiter bekannt“ : $H_0(x) \rightarrow H_{00}(x)$
 - „weil nichts weiter bekannt“ : $H_0(x) \rightarrow H_{00}(x)$
- bei „Zusatz-Information“ $A_i(x)$ dann Revision:

$$H_0(x) \wedge A_i(x) \rightarrow \neg H_{00}(x)$$

Umgang mit inkonsistenter Ausgangsinformation:

- Entscheidung für eine Variante (konsistente Teilmenge):

$$\{H_1, \neg H_1, H_2, H_3, \dots, H_n\} \text{ oder } \{\neg H_1, \neg H_1, H_2, H_3, \dots, H_n\}$$

Nicht-Monotonie: Frame-Problem

Beschreibung von Ereignissen und veränderlichen Welten.

Beispiel: *Situationen-Kalkül* (McCarthy, Hayes).

- Situationen beschrieben durch gültige Fakten:
 $\text{Holds}(\text{In}(\text{Anton}, \text{Hörsaal}), \text{Situation085})$
 $\text{Holds}(\text{Color}(\text{Hörsaal}, \text{weiss}), \text{Situation085})$
- Ereignisse verändern Situationen:
 $\text{Situation086} = \text{Result}(\text{Go}(\text{Anton}, \text{Mensa}), \text{Situation085})$
- Axiome beschreiben Veränderungen durch Ereignisse:
 $\forall a, l: s. \text{Holds}(\text{In}(a, l), \text{Result}(\text{Go}(a, l), s))$

Frame-Axiome beschreiben unveränderte Fakten:

$$\forall x, c, a, l: s. \text{Holds}(\text{Color}(x, c), \text{Result}(\text{Go}(a, l), s))$$

Praktisch nicht
handhabbar

Alternativ: allgemeines Axiom für „Persistenz-Default“
Ereignisse verändern Eigenschaften normalerweise nicht.

Nicht-Monotonie: Default-Annahmen

Default:

- Übliche/Normale Eigenschaft
- Fehlerfreie Funktion (Diagnose)
- Hintergrundwissen

„Normaler Ablauf“ für

- Regelsysteme,
- Frame-Systeme,
- ...

Spezielle Behandlung für Ausnahmen:
Priorität für Ausnahmeregel
Überschreiben von Default-Werten

Verträge sind gültig.
Verträge mit Minderjährigen sind ungültig.
Verträge mit Minderjährigen sind gültig, wenn sie im
Beisein eines Vormundes geschlossen werden.

Nicht-Monotonie: Default-Annahmen

Default: Übliche/Normale Eigenschaft

Inkonsistenzen entstehen durch Widersprüche zwischen

- Default und Ausnahme-Fall

$\text{vogel}(X) \rightarrow \text{fliegt}(X)$
 $\text{vogel}(X) \ \& \ \text{pinguin}(X) \rightarrow \neg \text{fliegt}(X)$

- unterschiedlichen Defaults

Quäker sind Pazifisten.
Republikaner sind keine Pazifisten.
Nixon ist Quäker und Republikaner.

Unbekannte Aussagen

Falls Gültigkeit der Aussage H nicht bekannt ist:

Zwei Varianten

Annahme: Aussage H gilt

oder

Annahme: Aussage H gilt **nicht** ($\neg H$ gilt)

Problem der Beweislast.

Nicht-Monotonie: Negation

Closed World Assumption (CWA)

Default-Annahme

- Eine Aussage H gilt **nicht** ($\neg H$ gilt), falls sie
 - nicht bekannt ist
 - nicht gefunden wird (Datenbank)
 - nachweislich nicht bewiesen werden kann (PROLOG)

Annahme der Gültigkeit von $\neg H$ bewirkt Nichtmonotonie:
Zusätzliche Voraussetzungen können
 H bekannt/auffindbar/beweisbar machen,
d.h. $\neg H$ wird ungültig

Unterschied:

- *Unschuldig.*
- *Freispruch mangels Beweises.*

Schließen auf $\neg H$

- Strenge CWA (Datenbanken, OPS-5)
 $\neg H$, falls H nicht in Datenbasis
- Negation by failure (PROLOG mit „finite failure“)
 $\neg H$, falls H nachweislich nicht beweisbar

- CWA in Logik:
 $\neg H$, falls H nicht folgt/nicht beweisbar

Korrekt nur dann,
wenn stets $\neg H$ oder H gültig
(vollständige Theorie).

Im PK1 ist $H \notin FI(X)$ nicht entscheidbar (nicht aufzählbar).

Schließen auf $\neg H$

- Unabhängige Beschreibungen
 H^+ (für H) und H^- (für $\neg H$)
ggf. spezieller Umgang mit Inkonsistenzen erforderlich.

- Dialektische Negation
 $\neg H$, falls „Argumente gegen H sprechen“

Unterschied bei „Negation by failure“

Klassische Logik:

$$\neg H_{00} \notin FI \{ \neg H_0 \rightarrow H_{00}, H_0 \}$$

Negation by failure (speziell PROLOG)

$$\text{not } H_{00} \in FI_{\text{Negation by failure}} \{ \text{not } H_0 \rightarrow H_{00}, H_0 \}$$

Beweis für $\text{not } H_{00}$:

- Versuche H_{00} zu beweisen:
 - Versuche $\text{not } H_0$ zu beweisen:
 - schlägt fehl wegen Axiom H_0
- Beweis für H_{00} fehlgeschlagen

$\text{not } H_{00}$ gültig

Vervollständigung mittels CWA

Theorie Th heie vollstndig,
falls fr jede atomare Grundformel H gilt:
Entweder $H \in Th$ oder $\neg H \in Th$.

Vervollstndigung:

Hinzunahme von fehlenden Formeln (H oder $\neg H$).

CWA-Vervollstndigung zu X :

$$V_{CWA}(X) := \{ \neg H \mid H \text{ Grundatom} \wedge H \notin FI(X) \}$$

(erfordert Entscheidung ob $H \notin FI(X)$)

$$Th(X) := FI(X)$$

$$\begin{aligned} CWA(X) &:= FI(X \cup V_{CWA}(X)) \\ &= Th(X \cup V_{CWA}(X)) \end{aligned}$$

Vervollständigung mittels CWA

Inkonsistenz bei CWA:

$X = \{ P(a) \vee Q(a) \}$ mit $P(a) \notin FI(X)$ und $Q(a) \notin FI(X)$

folglich:

$\{ P(a) \vee Q(a), \neg P(a), \neg Q(a) \} \subseteq CWA(X)$ -- inkonsistent!

Nicht-Monotonie:

$X = \{ P(a), P(b), P(a) \rightarrow Q(a) \}$

mit $\neg Q(b) \in CWA(X)$

aber $\neg Q(b) \notin CWA(X \cup \{Q(b)\})$

Vervollständigung mittels CWA

Unterschied positive/negative Grundlitterale:

$X = \{ P(a), P(b), P(a) \rightarrow Q(a) \}$

mit $\neg Q(b) \in CWA(X)$

$P_1 =_{Df} \neg P$, $Q_1 =_{Df} \neg Q$

$X_1 = \{ \neg P_1(a), \neg P_1(b), \neg P_1(a) \rightarrow \neg Q_1(a) \}$

mit $\neg Q_1(b) \in CWA(X)$

Vervollständigung mittels CWA

Satz

1. X sei konsistent.

$CWA(X)$ ist inkonsistent

gdw. Grundatome L_1, \dots, L_n existieren mit

$$L_1 \vee \dots \vee L_n \in \text{Th}(X), \text{ aber } L_1, \dots, L_n \notin \text{Th}(X).$$

2. X sei konsistent.

Die Umformung von X in Klauselform führe zu Hornklauseln.

Dann ist $CWA(X)$ konsistent.

3. Für Hornklauseln gilt:

Falls X konsistent, so auch $CWA(X)$.

Auftreten von Nicht-Monotonie

- CWA (und weitere Schlussformen bzgl. Annahme von Negation $\neg H$)
- Überschreiben von Defaults/Standardwerten
- Ausnahmeregeln vs. allgemeine Regeln
- Behandlung des Frameproblems (und verwandter Probleme)
- Behandlung impliziter Annahmen/Kontexte/Hintergründe
- Partielle Modellierung mit Annahmen

Allgemein: Partielle Information wegen

- unvollständigem Wissen
- veränderlichem Wissen
- zu hoher Beschreibungskomplexität

Anpassung von
Inferenzmethoden an
Nichtmonotonie, z.B.

- Regelsysteme,
- Vererbung/Überschreiben

Formale Behandlung von Nicht-Monotonie

Nichtmonotone Logiken

- Default Logiken
- Auto-epistemische Logiken
- Circumscription
- Präferenzlogiken

Belief-Revision, Truth-Maintenance-Systeme

- Protokollierung von Schlussfolgerungen/Abhängigkeiten
- Revision früherer Schlußfolgerungen

Ansatzpunkte für Formalismen

- Unterscheidung:
 - Striktes Wissen
 - Annahmen

- Spezielle Inferenzregeln

Default-Regeln:

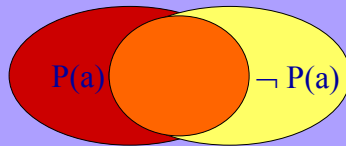
Aus **a** folgt **b**, falls nichts gegenteiliges bekannt ist.

- Modale Operatoren

belief(Geburtsjahr(Napoleon, 1869))

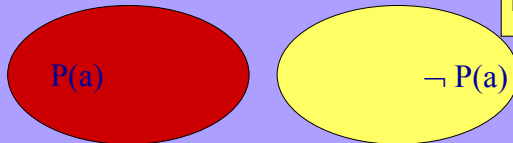
Ansatzpunkte für Formalismen

Inkonsistenz der gesamten Folgerungs-/ Ableitungsmenge.

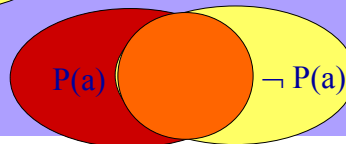


"Auswahl"-Strategie:
spezielle Extension
bevorzugen („Präferenzen“)

Konsistente Teilmengen: „Extensionen“



Skeptische Strategie:
Durchschnitt der Extensionen
verwenden



Default Logik (Reiter 1980)

Default-Regeln:

Falls a gilt, so gilt im allgemeinen auch b .

Aus a folgt b , falls nichts gegenteiliges bekannt ist.

Erweiterung des PK1

durch zusätzliche Ableitungsregeln

der Form (Default-Regel):

$$\frac{P : D_1, \dots, D_n}{C}$$

Außer den Prämissen P ist für die Ableitbarkeit
der Konklusion C die Berechtigung der
Annahmen D_1, \dots, D_n notwendig.

Default-Regel

Falls P ableitbar ist
und für alle D_i gilt: $\neg D_i$ ist **nicht** ableitbar,
so ist C ableitbar.

$$\frac{P : D_1, \dots, D_n}{C}$$
$$C$$
$$\frac{\text{vogel}(X) : \text{fliegt}(X)}{\text{fliegt}(X)}$$
$$\text{fliegt}(X)$$

Spezialfall: Falls $\neg D$ nicht ableitbar, so gilt D .

$$\frac{:D}{D}$$

Unterschied zu CWA:
Falls D nicht ableitbar, so gilt $\neg D$.

„Asymmetrie“: Kontraposition von Defaults muss nicht gelten.

Extensionen der Default-Logik

Extensionen:

X sei konsistente Menge von Formeln,

D sei Menge von Default-Regeln.

Funktion $\Gamma_{X,D}$ über Formelmengen Y ist definiert durch:

$$\Gamma_{X,D}(Y) := \text{Th}(X \cup \{C \mid \frac{P:D_1, \dots, D_n}{C} \in D \wedge P \in \Gamma_{X,D}(Y) \wedge \neg D_1, \dots, \neg D_n \notin Y\})$$

Die Fixpunkte von $\Gamma_{X,D}$ sind die Extensionen:
= Maximale konsistente Mengen.

Beispiele (Brewka): Extensionen

X	D	Fixpunkte (E)
Vog(Tw)	Vog(x):Flt(x) / Flt(x)	$E = \text{Th}(X \cup \text{Flt}(Tw))$
Vog(Tw) Pin(Tw) $\forall x(\text{Pin}(x) \rightarrow \neg \text{Flt}(x))$	Vog(x):Flt(x) / Flt(x)	$E = \text{Th}(X)$
Vog(Tw) Pin(Tw)	Vog(x): Flt(x) / Flt(x) Vog(x): $\neg \text{Flt}(x)$ / $\neg \text{Flt}(x)$	$E_1 = \text{Th}(X \cup \text{Flt}(Tw))$ $E_2 = \text{Th}(X \cup \neg \text{Flt}(Tw))$
Vog(Tw) Pin(Tw)	Vog(x):Flt(x) \wedge $\neg \text{Pin}(x)$ / Flt(x) Vog(x): $\neg \text{Flt}(x)$ / $\neg \text{Flt}(x)$	$E = \text{Th}(X \cup \neg \text{Flt}(Tw))$

Extensionen der Default-Logik

$$\Gamma_{X,D}(Y) := \text{Th}(X \cup \{C \mid \frac{P:D_1, \dots, D_n \in D \wedge P \in \Gamma_{X,D}(Y) \wedge \neg D_1, \dots, \neg D_n \notin Y}{C}\})$$

Eigenschaften der Extensionen E :

- „sicheres Wissen“ enthalten: $X \subseteq E$
- Abgeschlossenheit: $\text{Th}(E) = E$
- Anwendung der Default-Regeln soweit möglich
- Beschränkung auf jeweils damit ableitbare Formeln:

$$\Gamma_{X,D}(E) = E \quad (\text{Minimalität})$$

Extensionen der Default-Logik

$$\Gamma_{X,D}(Y) := \text{Th}(X \cup \{C \mid \frac{P:D_1, \dots, D_n \in D \wedge P \in \Gamma_{X,D}(Y) \wedge \neg D_1, \dots, \neg D_n \notin Y}{C}\})$$

Umständliche Definition ist notwendig:
bei Fixpunkten von

(statt $\Gamma_{X,D}(Y)$)

$$\Gamma_{X,D}(Y) := \text{Th}(X \cup \{C \mid \frac{P:D_1, \dots, D_n \in D \wedge P \in Y \wedge \neg D_1, \dots, \neg D_n \notin Y}{C}\})$$

wäre für $[X,D] = [\emptyset, \{b/a\}]$

außer $\{a\}$ auch $\{\neg b\}$ eine Extension.

Probleme Default-Logik

- Kreisförmige Schlüsse
- Inkonsistenzen

aus $\{ a:H / H, b:H / H, a \vee b \}$ ist H nicht ableitbar

- unerwartetes Verhalten

$[X,D] = [\emptyset, \{ : \neg H / H \}]$ hat keine Extensionen

- Unentscheidbarkeit von „ $\neg D$ nicht ableitbar“
- Default-Logik nicht axiomatisierbar

Auto-Epistemische Logik (Moore, 1985)

Modaler Operator Bel:

$\text{Bel}(p)$ = es wird geglaubt, daß p gilt

Zusätzliche Axiome der Form:

$P \wedge \text{Bel}(D) \rightarrow C$

$P \wedge \neg \text{Bel}(D) \rightarrow C$

Beispiele:

$\text{vogel}(X) \wedge \text{Bel}(\text{fliegt}(X)) \rightarrow \text{fliegt}(X)$

$\text{vogel}(X) \wedge \text{Bel}(\neg \text{fliegt}(X)) \rightarrow \neg \text{fliegt}(X)$

$\text{vogel}(X) \wedge \neg \text{Bel}(\text{fliegt}(X)) \rightarrow \neg \text{fliegt}(X)$

$\text{vogel}(X) \wedge \neg \text{Bel}(\neg \text{fliegt}(X)) \rightarrow \text{fliegt}(X)$

Auto-Epistemische Logik: Extensionen

E ist Extension von einer Axiomenmenge X , falls

$E := \text{Ab}(X \cup \{ \text{Bel}(H) \mid H \in E \} \cup \{ \neg \text{Bel}(H) \mid H \notin E \})$

Beispiel:

$X = \{ \text{vogel}(Tw) \wedge (\neg \text{Bel}(\neg \text{fliegt}(Tw)) \rightarrow \text{fliegt}(Tw)), \text{vogel}(Tw) \}$

$\text{fliegt}(Tw) \in E$

Auto-Epistemische Logik/Default-Logik

Default-Logik (DL) und Auto-Epistemische Logik (AEL) sind in gewisser Weise äquivalent.

$P: D_1, \dots, D_n / C$

entspricht

$\text{Bel}(P) \wedge \neg \text{Bel}(\neg D_1) \wedge \dots \wedge \neg \text{Bel}(\neg D_n) \rightarrow C$

DL-Extensionen entsprechen gewissen Bel-freien AEL-Extensionen.

Circumscription (McCarthy, 1980)

Idee: Gültigkeitsbereich spezieller Prädikate minimal festlegen.

- Bei CWA: Festlegung auf einen minimalen Gültigkeitsbereich mittels Folgerungs-/Ableitungsrelation.
- Bei Circumscription: Festlegung auf einen minimalen Gültigkeitsbereich mittels zusätzlicher Axiome.

Gültigkeitsbereich eines Prädikats P eingrenzbar durch

- positive Festlegungen, z.B.
 - $P(a), P(b), \dots$
 - $\forall x (H(x) \rightarrow P(x)), \dots$
 - „im Zweifelsfalle für den Angeklagten“
- negative Festlegungen, z.B.
 - $\neg P(c), \neg P(d), \dots$
 - $\forall x (H(x) \rightarrow \neg P(x)), \dots$
 - CWA

Auswirkungen
jeweils für $P(e)$?

Circumscription (Mc Carthy, 1980)

Minimaler Gültigkeitsbereich mittels zusätzlicher Axiome.

Beispiel:

verheiratet(Peter).

verheiratet(Petra).

als weiteres Axiom:

$\forall x(\text{verheiratet}(x) \rightarrow x = \text{Peter} \vee x = \text{Petra})$

allgemein:

(Axiomen-)Schema Prädikaten-Circumscription für ein Prädikat P bezüglich einer Formel H .

(Axiomen-)Schema Prädikaten-Circumscription

Definition:

P sei n -stelliges Prädikatensymbol, H Formel ohne freie Variable.

$H(Q)$ entstehe aus H durch Ersetzung des Prädikatensymbols P durch ein n -stelliges Prädikatensymbol Q .

Das Schema der Prädikaten-Circumscription von P bezüglich H ist die Menge aller mit unterschiedlichen Q möglichen Formeln

$$\begin{aligned} & (H(Q) \wedge \forall x_1 \dots \forall x_n (Q(x_1, \dots, x_n) \rightarrow P(x_1, \dots, x_n))) \\ & \rightarrow \forall x_1 \dots \forall x_n (P(x_1, \dots, x_n) \rightarrow Q(x_1, \dots, x_n)) \end{aligned}$$

Instanzen dieses Schemas gemeinsam mit $H(P)$ und den sonstigen Axiomen für Ableitungen benutzen:

Geignetes Q erzwingt P mit minimalen Gültigkeitsbereich.

Alternativ z.B.: Circumscription-Axiom in PK2

Präferenzen

- Präferierte Theorie:
 - Präferenz-Relation bzgl. der maximalen konsistenten Teiltheorien einer inkonsistenten Theorie.
- Präferierte Modelle:
 - Präferenz-Relation über den Modellen.
 - Folgerungsrelation nur bzgl. der präferierten Modelle.
 - Präferenz von Modellen z.B. durch
 - Modelle mit wenig Ausnahmen
 - Modelle mit vielen Standard-Annahmen

Circumscription bedeutet Präferenz von Modellen mit minimaler Gültigkeitsbereich für Prädikat P .

Revisions-Mechanismen

„Belief Revision“:

Im Inferenz-Prozess werden aus Axiomen/Annahmen A_1, A_2, \dots Schlussfolgerungen H_1, H_2, \dots gezogen.

Bei Revision von A_i kann Rechtfertigung für H_j entfallen:
ggf. muß auch H_j revidiert werden.

Z.B. in nicht-monotonen Systemen:

- Konsistenz-Prüfung bei veränderten Voraussetzungen
- Überprüfung der bisher erfolgten Schlüsse

„Belief Update“:

Durch Zustandsänderung notwendige Aktualisierungen.

Revisions-Mechanismen

Wissensbasis $B \cup \{p\}$ sei inkonsistent

$\{a, b, c, a \wedge b \rightarrow c, \neg c\}$

Syntaktische Revision:

Ziel: konsistente Teilmenge $B' \subseteq B \cup \{p\}$ auswählen

Kriterien für Auswahl:

- Zuverlässigkeit der Information
- Wichtigkeit der Information

$\{b, a \wedge b \rightarrow c, \neg c\}$

$\{a, a \wedge b \rightarrow c, \neg c\}$

$\{a, b, \neg c\}$

$\{a, b, c, a \wedge b \rightarrow c\}$

Semantische Revision:

Betrachtung von Modellen

Truth Maintenance Systeme

Truth Maintenance Systeme (TMS)

(auch „Reason Maintenance Systeme“ - RMS)

protokollieren die Ableitungsprozesse bzw.

die Abhängigkeiten und

ermöglichen damit die Revision von Schlussfolgerungen.

- JTMS - Justification Based TMS (Doyle, 1979)
Protokolliert die Begründungen:
Regelanwendungen mit unmittelbaren Voraussetzungen
- ATMS - Assumption Based TMS (de Kleer, 1984)
Protokolliert die jeweils zugrunde liegenden Axiome