

Moderne Methoden der KI

Prof. Dr. sc. Hans-Dieter Burkhard
Vorlesung Sommer-Semester 2007

Modale Logiken

Einführung

Ausgangspunkt: Zustand des Agenten
- Welt-Modell (Umwelt, interne Parameter)
 Vergangenheit - Gegenwart
- Ziele, Aufgaben.Pläne
 Zukunft
Darstellung als Datenbank von Fakten
(logischer Formalismus)

Annahmen eines Agenten müssen nicht korrekt sein

H.D.Burkhard, HU Berlin Vorlesung MMKI
Sommer-Semester 2007 Modale Logik

2

Einführung

Agent glaubt, dass Kronos der Vater von Zeus ist

Wie formalisieren?

Als Relation: belief(agent, vater(kronos,zeus))

- nicht mehr klassischer PK1:

 Jupiter=zeus

 belief(agent, vater(kronos,jupiter))

muss nicht gelten: belief-Logik ist **nicht extensional**

Darstellung durch für möglich gehaltene Welten:
Kartenspiel-Beispiel

H.D.Burkhard, HU Berlin Vorlesung MMKI
Sommer-Semester 2007 Modale Logik

3

Modale Logik

Definition: Ausdrücke für modale Aussagen-Logik

AV := Menge der Aussagenvariablen.

(0) Jedes $p \in AV$ ist ein Ausdruck.

(1) Mit H_1, H_2 sind auch $\neg H_1$ und $H_1 \vee H_2$ Ausdrücke (und ggf. weitere)
Zusätzlich:

(2) Wenn H ein Ausdruck ist, so auch $\Box H$ (und ggf. $\Diamond H$).

Definieren weiter (oder gleich einführen):

$\Diamond p := \neg \Box \neg p$: H gilt möglicherweise

3 Modalitäten: notwendig H , möglich H , (notwendig) nicht H

H.D.Burkhard, HU Berlin Vorlesung MMKI
Sommer-Semester 2007 Modale Logik

4

Modale Logik

Modale Operatoren für *notwendig* und *möglich*

$\Box p$ oder $L p$ für: Es ist notwendig, dass p gilt.

$\Diamond p$ oder $M p$ für: Es ist möglich, dass p gilt.

Anstelle von *Notwendigkeit* auch

- *Verbindlichkeit, Zulässigkeit* bzw. *Verbot* \Leftrightarrow Deontische Logik
- *Wissen = knowledge* \Leftrightarrow Epistemische Logik
- *Annahmen, Überzeugung = belief* \Leftrightarrow Epistemische Logik
 = subjektiv: *für notwendig halten*
- „immer“ = zu jeder Zeit *notwendig* \Leftrightarrow Temporale Logik
 korrespondiert mit „manchmal“ = zu einer Zeit *möglich*

H.D.Burkhard, HU Berlin Vorlesung MMKI
Sommer-Semester 2007 Modale Logik

5

Modale Logik

Möglichkeit bedeutet dabei die Nicht-Notwendigkeit des Gegenteils:

$\neg \Box \neg p$ entspricht $\Diamond p$
 $\neg \text{know } \neg p$ entspricht „nicht wissen“
 $\neg \text{belief } \neg p$ entspricht „nicht annehmen“
 $\neg \text{verboten } \neg p$ könnte entsprechen *erlaubt* $\neg p$
 $\neg \text{immer } \neg p$ entspricht *manchmal p*

In epistemischen Logiken beziehen sich Ausdrücke mit $\Diamond p$ auf Aussagen über „nicht wissen“ bzw. „nicht annehmen“

H.D.Burkhard, HU Berlin Vorlesung MMKI
Sommer-Semester 2007 Modale Logik

6

Zeitbezüge

Am Montag nimmt der Agent an, dass am Dienstag bekannt ist, ob bis Donnerstag alle verfügbaren Flugreisen für den folgenden Montag verkauft sind.
 Am Montag beschließt der Agent, dass am Dienstag festgelegt werden muss, ob bis spätestens Donnerstag neue Angebote an Flugreisen für den folgenden Montag beschafft werden müssen.

wann gilt eine Annahme, wann soll ein Plan ausgeführt werden, wann wurde eine Annahme getroffen bzw. überprüft, wann wurde ein Plan erstellt bzw. überprüft,	wann soll eine Annahme überprüft werden, wann soll ein Plan überprüft werden. ...
---	---

Quelle für Missverständnisse.

Modale Logik

- beliebig verschachtelte Modalitäten
 ich weiß, dass ich nicht weiß ... $\text{know}(\neg \text{know } H)$
 Fairness: immer gilt, dass irgendwann... $\text{immer}(\text{irgendwann } H)$
 z.T. Reduktionsaxiome/-regeln, z.B.: $\Box H \leftrightarrow \Box H$ im System S4
- Verbindungen der Logiken
 er glaubt manchmal, dass er weiß ... $\text{irgendwann}(\text{beliefs}(\text{knows } H))$
- Gegenseitige Modellierung
 er weiß, dass sie weiß ... $\text{know}_a(\text{know}_b H)$

Modale Logik

„Wise Men - Puzzle“ (J. McCarthy)

Der König will seine drei weisen Männer auf die Probe stellen. Sie müssen sich so in einem Kreis aufstellen, dass jeder die anderen sehen und hören kann. Er erklärt, dass er jedem einen Hut aufsetzen wird, dessen Farbe weiß oder schwarz sein kann, wobei aber wenigstens einer der Hüte weiß ist. Tatsächlich setzt er jedem einen weißen Hut auf. Dann stellt er ihnen der Reihe nach die Frage: „Weißt Du, welche Farbe Dein Hut hat?“ Wie antworten die (wahrhaftig) weisen Männer?

Erster weiser Mann: Ich weiß es nicht."
 Zweiter weiser Mann: Ich weiß es nicht."
 Dritter weiser Mann: Mein Hut ist weiß."

Semantik

$\text{Bel}(H)$ als Gültigkeit von H in allen Welten, die der Agent betrachtet

- „sichtbare“ Welten
- die tatsächliche Welt muss nicht sichtbar sein

$\text{Bel}(\text{Bel}(H))$ usw.:
sichtbar aus sichtbarer Welt

Negation: $\neg \text{Bel}(H)$
in einer sichtbaren Welt gilt H nicht

Kripke-Semantik, „possible-worlds“-Semantik

Rahmen: Mögliche Welten und Sichtbarkeits-Relationen,
formal als Graph („frame“, „Rahmen“):

$G = (Z, R)$ mit Knoten-Menge (Welten, Zuständen) Z
 und Kanten-Menge $R \subseteq Z \times Z$
 R ist Sichtbarkeitsrelation (accessibility relation):
 „ z_0 sichtbar von z “, falls $R(z, z_0)$.
 Ausgangspunkt: Tatsächliche Welt des „Agenten“

Modell: $[G, \beta]$ mit
 Rahmen $G = (Z, R)$ und Belegung $\beta: AV \times Z \rightarrow \{\text{true}, \text{false}\}$
 $\beta(p, z)$ = Wahrheitswert der AV p im Knoten z .
 $\beta(p, z) = \text{true}$ bedeutet: p ist wahr in der „Welt“ z .

Semantik

Erfüllbarkeitsrelation \models
 $\langle M, z \rangle \models H$: $\langle M, z \rangle$ erfüllt den Ausdruck H
 H ist wahr in der Welt z des Modells $M = [(Z, R), \beta]$

Induktiv definiert durch:

$\langle M, z \rangle \models p$	gdw.	$p \in AV$ mit $\beta(p, z) = \text{true}$
$\langle M, z \rangle \models \neg H$	gdw.	nicht $\langle M, z \rangle \models H$
$\langle M, z \rangle \models H_1 \vee H_2$	gdw.	$\langle M, z \rangle \models H_1$ oder $\langle M, z \rangle \models H_2$
$\langle M, z \rangle \models \Box H$	gdw.	$\langle M, z_0 \rangle \models H$ für alle z_0 mit $R(z, z_0)$
$\langle M, z \rangle \models \Diamond H$	gdw.	$\langle M, z_0 \rangle \models H$ für ein z_0 mit $R(z, z_0)$

offenbar gegenseitige Definierbarkeit von $\Box H$ und $\Diamond H$:

$\Diamond H = \neg \Box \neg H$
 $\Box H = \neg \Diamond \neg H$

Semantik für modale Aussagen-Logik

H ist gültig in $z \in Z$ von $M = [(Z, R), \beta]$ gdw. $\langle M, z \rangle \models H$
 H ist gültig in $M = [(Z, R), \beta]$ gdw. $\langle M, z \rangle \models H$ für alle $z \in Z$
 H ist gültig in einer Klasse \mathcal{M} gdw. H gültig für alle $M \in \mathcal{M}$
 H ist allgemeingültig gdw. H gültig in allen Modellen M
 (Schreibweise: $\models H$)

Analog:

H ist erfüllbar in $M = [(Z, R), \beta]$ gdw. $\langle M, z \rangle$ exist. mit $\langle M, z \rangle \models H$
 H ist erfüllbar gdw. Modell M exist. mit H gültig in M

Semantik

Folgerungen:

Bei Kripke-Semantik gilt stets:

- (1) Wenn $\models H$, so $\models \Box H$ (Notwendigkeitsregel)
 (2) (K) $\Box (H_1 \rightarrow H_2) \rightarrow (\Box H_1 \rightarrow \Box H_2)$ (Distributions-Axiom)

(Beweise)

Weitere mögliche Forderungen

(müssen aber nicht gelten)

Axiom T $\Box H \rightarrow H$ (Notwendigkeits-Axiom)
 äquivalent: $H \rightarrow \Box H$
Axiom D $\Box H \rightarrow \Diamond H$
 äquivalent: $\Box H \rightarrow \neg \Box \neg H$
Axiom 4 $\Box H \rightarrow \Box \Box H$
Axiom 5 $\Diamond H \rightarrow \Box \Diamond H$

Weitere mögliche Forderungen

Axiom T $\Box H \rightarrow H$
 äquivalent: $H \rightarrow \Diamond H$
Axiom D $\Box H \rightarrow \Diamond H$
 äquivalent: $\Box H \rightarrow \neg \Box \neg H$
Axiom 4 $\Box H \rightarrow \Box \Box H$
Axiom 5 $\Diamond H \rightarrow \Box \Diamond H$

Für Belief/Knowledge:

Axiom T knows $H \rightarrow H$
 $H \rightarrow \neg \text{knows} \neg H$
 (konsistent mit Realität)
Axiom D Bel $H \rightarrow \neg \text{Bel} \neg H$
 (keine Inkonsistenz in Annahmen)
Axiom 4 Bel $H \rightarrow \text{Bel Bel } H$
 (positive Introspektion)
Axiom 5 $\neg \text{Bel } H \rightarrow \text{Bel } \neg \text{Bel } H$
 (negative Introspektion)

Modallogische Theorien

Syntaktisch bestimmt (Ableitungen)

$\text{Th}(X) = \text{Ab}(X)$

Axiomenmenge
 Ableitungsregeln

Bei Bezug auf Kripke-Modelle:

„Normale Logiken“, enthalten

- Axiom K
- Notwendigkeitsregel

Semantisch bestimmte Theorien:

Für gegebene Klasse \mathcal{M} von Modellen:
 $\text{Th}(\mathcal{M}) = \{ H \mid H \text{ gültig in Klasse } \mathcal{M} \}$

Beschreibung von Eigenschaften der Klasse \mathcal{M} ,

z.B. bestimmte Sichtbarkeitsrelationen

Modallogische Theorien

Syntaktisch bestimmte Theorien: $\text{Th}(X) = \text{Ab}(X)$

- Axiomensystem X : (als Axiomenschema mit allen Einsetzungen)
 - Tautologien des AK
 - Axiom K
 - weitere Axiome
- Ableitungsoperator Ab durch Ableitungsregeln festgelegt:
 - Abtrennungsregel (modus ponens):
 Wenn $\text{Ab}(H_1 \rightarrow H_2)$ und $\text{Ab}(H_1)$, so $\text{Ab}(H_2)$
 - Notwendigkeitsregel:
 Wenn $\text{Ab}(H)$, so $\text{Ab}(\Box H)$

Modallogische Theorien

Syntaktisch bestimmte Theorien: $Th(X) = Ab(X)$

- Spezielle Axiome für modale Operatoren (als Axiomenschemata mit allen Einsetzungen)

Axiom K $\square (H_1 \rightarrow H_2) \rightarrow (\square H_1 \rightarrow \square H_2)$ (Distributions-Axiom)
 Axiom T $\square H \rightarrow H$ (Notwendigkeits-Axiom)
 äquivalent: $H \rightarrow \diamond H$
 Axiom D $\square H \rightarrow \diamond H$ (folgt aus T)
 Axiom 4 $\square H \rightarrow \square \square H$
 Axiom 5 $\diamond H \rightarrow \square \diamond H$

Kombination der Axiome

führt zu speziellen „Systemen“ modaler Logiken, z.B.

System T: $Ab(„Tautologien + K + T“)$

H.D.Burkhard, HU Berlin Sommer-Semester 2007 Vorlesung MMKI Modale Logik

19

Modallogische Theorien

Semantisch bestimmte Theorien:

$Th(M) = \{ H \mid H \text{ gültig in der Klasse } M \text{ von Modellen } \}$

Modell-Klassen M bestimmt durch Eigenschaften der Sichtbarkeitsrelation R in $G = [Z, R]$:

- reflexive Sichtbarkeits-Relation $\forall x \in Z: R(x,x)$
- serielle Sichtbarkeits-Relation $\forall x \in Z \exists y \in Z: R(x,y)$
- transitive Sichtbarkeits-Relation $\forall x,y,z \in Z: R(x,y) \wedge R(y,z) \rightarrow R(x,z)$
- euklidische Sichtbarkeits-Relation $\forall x,y,z \in Z: R(x,y) \wedge R(x,z) \rightarrow R(y,z)$

H.D.Burkhard, HU Berlin Sommer-Semester 2007 Vorlesung MMKI Modale Logik

20

Modallogische Systeme

Alle semantisch über Kripke-Modelle definierten Logiken sind normal.

Ferner gelten die Axiome

- T bei reflexiver Sichtbarkeits-Relation R in $G = [Z, R]$
- D bei serieller Sichtbarkeits-Relation R in $G = [Z, R]$
- 4 bei transitiver Sichtbarkeits-Relation R in $G = [Z, R]$
- 5 bei euklidischer Sichtbarkeits-Relation R in $G = [Z, R]$

Umkehrungen gelten in folgender Form:

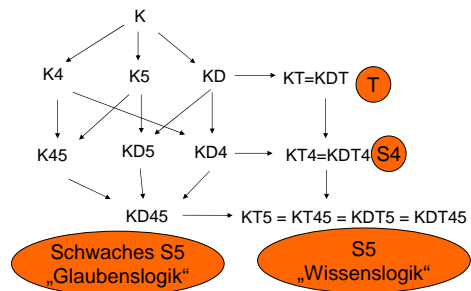
Sei $G=[Z,R]$ ein Rahmen.

Wenn Axiom XXX bei **jeder** Belegung β (in jedem Modell $M=[G,\beta]$) gilt, dann muss R die korrespondierende Bedingung erfüllen.

H.D.Burkhard, HU Berlin Sommer-Semester 2007 Vorlesung MMKI Modale Logik

21

Modallogische Systeme



H.D.Burkhard, HU Berlin Sommer-Semester 2007 Vorlesung MMKI Modale Logik

22

Syntaktisch definierte Systeme

Durch Kombination der weiteren Axiome T, D, 4, 5 ergeben sich 11 unterschiedliche normale Systeme:

K
 K4
 K5
 KD
 KT = KDT System T
 K45
 KD4
 KD5
 KT4 = KDT4 System S4
 KD45 schwaches System S5
 KT5 = KT45 = KDT5 = KDT45 System S5

H.D.Burkhard, HU Berlin Sommer-Semester 2007 Vorlesung MMKI Modale Logik

23

Semantisch definierbare Systeme

mit Bezug auf Rahmen (statt Modellen wie oben) gilt:

- System T: durch reflexive Sichtbarkeits-Relation R .
- System S4: durch reflexive und transitive Sichtbarkeits-Relation R .
- System S5: durch reflexive, transitive und symmetrische Sichtbarkeits-Relation R . (R ist Äquivalenz-Relation: Klassen gegenseitig sichtbarer Welten).
- schwaches System S5: durch serielle, transitive und euklidische Sichtbarkeits-Relation R .

H.D.Burkhard, HU Berlin Sommer-Semester 2007 Vorlesung MMKI Modale Logik

24

System T

Definiert durch

- semantisch durch reflexive Sichtbarkeitsrelation oder
- syntaktisch als normale Logik mit zusätzlich
Axiom T $\Box H \rightarrow H$ (Notwendigkeits-Axiom)

Bereits in **System T** gilt:

- $\Box \Box H \rightarrow \Box H$ (Umkehrung von **Axiom 5**: $\Box H \rightarrow \Box \Box H$)
- $\Box \Box H \rightarrow \Box H$ (Umkehrung von **Axiom 4**: $\Box H \rightarrow \Box \Box H$)
- $\Box H \rightarrow \Box \Box H$ (Umkehrung folgt aus **Axiom 5**)
- $\Box H \rightarrow \Box \Box H$ (Umkehrung folgt aus **Axiom 4**)

⇒ Reduktionsmöglichkeiten für verschachtelte Modalitäten in **S4**, **S5**.

Epistemische Logiken

Notwendigkeit interpretieren als

- **know p** wissen
- **bel p** annehmen (überzeugt sein, glauben)
- erlaubt sein
- verpflichtet sein (deontische Logik)

Möglichkeit bedeutet dabei die Nicht-Notwendigkeit des Gegenteils.

- $\neg \Box \neg p$ entspricht $\Diamond p$
- $\neg \text{know } \neg p$ entspricht „nicht wissen“
- $\neg \text{belief } \neg p$ entspricht „nicht annehmen“
- $\neg \text{erlaubt } \neg p$ entspricht **verboten** $\neg p$

Bzw. $\Diamond p$ bedeutet „nicht wissen“, „nicht annehmen“ usw.

Wissenslogik

- Axiom K** $\text{knows}(H_1 \rightarrow H_2) \rightarrow (\text{knows } H_1 \rightarrow \text{knows } H_2)$
- Axiom T** $\text{knows } H \rightarrow H$ (Korrektheits-Axiom)
- Axiom D** $\text{knows } H \rightarrow \neg \text{knows } \neg H$ (Konsistenz-Axiom)
- Axiom 4** $\text{knows } H \rightarrow \text{knows knows } H$ (positive Introspektion)
- Axiom 5** $\neg \text{knows } H \rightarrow \text{knows } \neg \text{knows } H$ (negative Introspektion)

Axiome 4 + 5: „Perfekte Selbstkenntnis“

Glaubenslogik

Axiome ohne T

- Axiom K** $\text{bel}(H_1 \rightarrow H_2) \rightarrow (\text{bel } H_1 \rightarrow \text{bel } H_2)$
- Axiom D** $\text{bel } H \rightarrow \neg \text{bel } \neg H$ (Konsistenz-Axiom)
- Axiom 4** $\text{bel } H \rightarrow \text{bel bel } H$ (positive Introspektion)
- Axiom 5** $\neg \text{bel } H \rightarrow \text{bel } \neg \text{bel } H$ (negative Introspektion)

(Möglicher Zusammenhang: $\text{knows } H \leftrightarrow H \wedge \text{bel } H$)

Häufig: **System S5** als Wissenslogik
Schwaches System S5 als Überzeugungslogik.

Omniszenz-Problem

Notwendigkeitsregel:

Wenn **H** allgemeingültig ist, so ist **knows H** allgemeingültig.
D.h.: Wissen umfasst mindestens alle Tautologien!

Weiter:

- Sei **H** Folgerung aus Axiomenmenge $X = \{H_1, \dots, H_n\}$
d.h. $H_1 \wedge \dots \wedge H_n \rightarrow H$ ist allgemeingültig.
- Notwendigkeitsregel: $\text{knows}(H_1 \wedge \dots \wedge H_n \rightarrow H)$ allgemeingültig.
- Axiome sind bekannt: $\text{knows}(H_1 \wedge \dots \wedge H_n)$.
- Gemäss **Axiom K**:
 $\text{knows}(H_1 \wedge \dots \wedge H_n \rightarrow H)$ und $\text{knows}(H_1 \wedge \dots \wedge H_n)$ implizieren $\text{knows } H$.

D.h.: Mit Axiomenmenge **X** sind alle Folgerungen bekannt!

Omniszenz-Problem

In normalen Logiken (possible-worlds-Semantik) gelten stets **Axiom K** und die Notwendigkeitsregel und damit „Omniszenz“.

Syntaktischer Zugang ohne **Axiom K** und Notwendigkeitsregel muss folglich andere Semantik haben.

Alternative Ansätze:

Levesque: explizites und implizites (ableitbares) Wissen unterscheiden.
Konolige: Wissensbasis + Inferenzregeln bestimmen das Wissen

In der Praxis häufig Beschränkung der Ausdrucksstärke
(z.B. nur Literale)

Epistemische Logiken

In Multi-Agenten-Systemen:

Wissen/Glauben jedes einzelnen Agenten a :

Agent a weiss, dass H gilt: $knows_a H$.

Agent a nimmt an, dass H gilt: $bel_a H$.

Wissen über andere: $knows_a knows_b H$

Gemeinsames Wissen „common knowledge“
(analog: common belief)

(vgl. wise man puzzle)

(vgl. Absprache-Problem: Koordinierte Attacke, siehe Buch Wooldridge)

H.D.Burkhard, HU Berlin Sommer-Semester 2007

Vorlesung MMKI Modale Logik

31

(Statische) Ziele

goal H ebenfalls mittels Kripke-Modell beschreibbar:

$\square p$ entspricht goal p

Eigenschaften, die unbedingt gelten sollen

$\square \neg p = goal \neg p$

Eigenschaften, die keinesfalls gelten sollen

$\diamond p = \neg \square \neg p = \neg goal \neg p$

Eigenschaften, die nicht gefordert werden

Konsistenz bzgl. Annahmen (für den entsprechenden Zeitpunkt) kann gefordert werden, z.B.:

- Goal-Modell als Teilgraph des Belief-Modells oder
- $goal p \rightarrow \neg bel \neg p$ (Möglichkeit des Erreichens)
- $goal p \rightarrow bel(goal p)$ (Agent kennt seine Ziele)

H.D.Burkhard, HU Berlin Sommer-Semester 2007

Vorlesung MMKI Modale Logik

32

Probleme

a) Zeitliche Veränderungen modellieren (\rightarrow temporale Logiken)

b) Seiteneffekte

Abschluss bzgl. logischer Konsequenzen (vgl. Omniscienz-Problem):

Nach Notwendigkeitsaxiom:

$H_1 \rightarrow H_2$ impliziert goal $(H_1 \rightarrow H_2)$

nach K-Axiom:

goal $(H_1 \rightarrow H_2)$ und goal H_1 implizieren goal H_2

d.h. Konsequenzen sind auch Ziele ...

(\rightarrow keine volle logische Inferenz (\rightarrow temporale Logiken))

H.D.Burkhard, HU Berlin Sommer-Semester 2007

Vorlesung MMKI Modale Logik

33

Zeitliche Abläufe

Aktionen zum Erreichen von Zielen an zeitliche Abläufe gebunden

Vorgaben bzgl. Modellierung zeitlicher Abläufe:

- Verzweigt vs. linear
- Zeitpunkte vs. Intervalle
- Diskret vs. kontinuierlich
- Zukunft (Vergangenheit): beschränkt vs. unbeschränkt
- Zyklisch vs. azyklisch

Modelle analog zu Kripke-Modellen $[G, \beta]$

mit Zeitstrukturen $G = [Z, R]$, wobei R zeitliche Nachfolge beschreibt

- R ist transitiv, reflexiv (partielle Ordnung)
- Lineare Zeit: R ist Ordnungsrelation

H.D.Burkhard, HU Berlin Sommer-Semester 2007

Vorlesung MMKI Modale Logik

34

Temporale Logik

Operatoren bzgl. linearer Zeit:

$P \phi$ Irgendwann in der Vergangenheit galt ϕ .

$F \phi$ Irgendwann in der Zukunft wird ϕ gelten.

$H \phi$ Immer in der Vergangenheit galt ϕ .

$G \phi$ Immer in der Zukunft wird ϕ gelten.

P und F analog zu \diamond , H und G analog zu \square .

$P \phi = \neg H \neg \phi$

$F \phi = \neg G \neg \phi$

(aktueller Zeitpunkt wird immer eingeschlossen)

weitere Operatoren:

$X \phi$ ("next") Im nächsten Zeitpunkt gilt ϕ .

$\phi U \psi$ ("until") irgendwann gilt ψ , (mindestens) bis dann gilt ϕ

H.D.Burkhard, HU Berlin Sommer-Semester 2007

Vorlesung MMKI Modale Logik

35

Temporale Logik

Definierbarkeit von F/G mittels X/U :

$F \phi = true U \phi$

$G \phi = \neg F \neg \phi$

weitere:

unendlich oft: $F^\infty \phi =_{DF} GF \phi$

fast überall: $G^\infty \phi =_{DF} FG \phi$

ϕ vor ψ : $(\phi B \psi) =_{DF} \neg((\neg \phi) U \psi)$

H.D.Burkhard, HU Berlin Sommer-Semester 2007

Vorlesung MMKI Modale Logik

36

Handlungen sind zeitliche Abläufe

- Zusammenhang von Absichten/Plänen mit Realität
- Rationaler Agent soll keine „unvernünftigen“ Absichten haben (Übereinstimmung mit Realität, Persistenz, Abbruch)

Beispiele:

Roboter Willie, Little Nell's Problem, Cheating Husbands

Kombinationen in multi-modalen temporalen Systemen:

z.B. $\text{intend}(\phi) \wedge \text{belief}(\phi \rightarrow \psi) \rightarrow \text{intend}(\psi)$
 $\text{intend}(\phi) \wedge \neg \phi \rightarrow X \text{intend}(\phi)$
 $\text{intend}(\phi) \wedge \text{can}(\phi) \rightarrow F \phi$

desire (AF Gewinnen) & intend (EF Loskaufen) & not belief (AF Gewinn)

Roboter Willie (in Anlehnung an Cohen-Levesque)

Willie bring Bier - Willie beschäftigt mit anderen Dingen,
aber „irgendwann wird er Bier bringen“

Fehler bzgl. Ziele verfolgen

Willie bring Bier - Willi wartet vor dem leeren Kühlschrank

Fehler bzgl. unerreichbare Ziele verfolgen

Willie bring Bier - Willi holt Bier aber verschüttet es unterwegs
absichtlich

Fehler bzgl. Ziele verfolgen

Willie bring Bier und Brötchen - Willi findet keine Brötchen und
verzichtet auf weitere Arbeit

Fehler bzgl. Teil-Ziele verfolgen

Willie bring Bier - Kühlschrank leer, aber Willi bringt Öffner und Glas

Fehler bzgl. Adaption von Teil-Zielen

Cheating Husbands

The queens of the matriarchal city-state of Mamajorca, on the continent of Atlantis, have a long record of opposing and actively fighting the male infidelity problem. Ever since the technologically-primitive days of queen Henrietta I, women in Mamajorca have been required to be in perfect health and pass an extensive logic and puzzle-solving exam before being allowed to take a husband. The queens of Mamajorca, however, were not required to show such competence.

It has always been common knowledge among the women of Mamajorca that their queens are truthful and that the women are obedient to the queens. It was also common knowledge that all women hear every shot fired in Mamajorca. Queen Henrietta I awoke one morning with a firm resolution to do away with the male infidelity problem in Mamajorca. She summoned all of the women heads-of-households to the town square and read them the following statement:

Cheating Husbands

There are (one or more) unfaithful husbands in our community. Although none of you knew before this gathering whether your own husband was faithful, each of you knows which of the other husbands are unfaithful. I forbid you to discuss the matter of your husband's fidelity with anyone. However, should you discover that your husband is unfaithful, you must shoot him on the midnight of the day you find out about it.

What happens?

Cheating Husbands

There are (one or more) unfaithful husbands in our community. Although none of you knew before this gathering whether your own husband was faithful, each of you knows which of the other husbands are unfaithful. I forbid you to discuss the matter of your husband's fidelity with anyone. However, should you discover that your husband is unfaithful, you must shoot him on the midnight of the day you find out about it.

Thirty nine silent nights went by, and on the fortieth night, shots were heard.

Temporale Logik

Semantik:

- Welten (Mengen gültiger Aussagen) zu unterschiedlichen Zeiten
- Sichtbarkeitsrelation an Zeitmodelle angepasst (z.B. linear)

Umschreibung mit klassischen Mitteln z.B.:

$\exists t (\text{later}(t, \text{now}) \wedge \phi(t))$ für $F \phi$

$\forall t (\text{later}(t, \text{now}) \rightarrow \phi(t))$ für $G \phi$

Probleme (?), z.B.:

a) Unterschied bzgl.: $F(\exists x H(x))$ und $\exists x(F H(x))$

kein Unterschied bei Umschreibung im PK1

b) $F F \phi \rightarrow F \phi$ ist allgemeingültig, Umschreibung im PK1 nicht.

Verzweigte Zeit

Operatoren für verzweigte Zeit (bzgl. Zukunft)

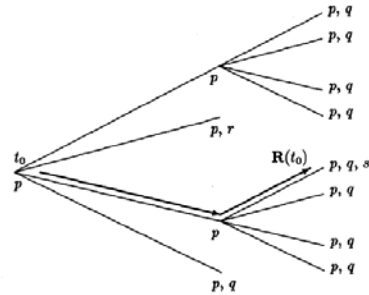
- $A \phi$ für „always“: In jeder Zukunft gilt ϕ .
- $E \phi$ für „exists“: In einer Zukunft gilt ϕ .

In Kombination mit Operatoren für lineare Zeit:

„Computational Tree Logic“ (CTL bzw. CTL*)

Verzweigte Zeit

Beispiel aus M.P. Singh: Multiagent systems



CTL* (Emerson)

(extended Computational Tree Logic, full branching time logic)

- Operatoren für verzweigte Zeit
 - $A \phi$ für „always“: In jeder Zukunft gilt ϕ .
 - $E \phi$ für „exists“: In einer Zukunft gilt ϕ .
- Operatoren bezüglich eines Weges (lineare Zeit)
 - $X \phi$ ("next") Im nächsten Zeitpunkt gilt ϕ .
 - $\phi U \psi$ ("until") Zunächst gilt ϕ , danach ψ .

Syntax CTL*

AV = Menge der Aussagenvariablen.

- AK-Ausdrücke wie üblich.
- Zustandsformeln (Betrachtung der Wege von einem Zustand aus):
 - (Z1) Jedes $p \in AV$ ist eine Zustandsformel.
 - (Z2) Mit ϕ, ψ sind auch $(\neg \phi)$ und $(\phi \vee \psi)$ Zustandsformeln.
 - (Z3) Wenn ϕ Wegformel ist, so sind $E \phi$ und $A \phi$ Zustandsformeln.
- Wegformeln: (Betrachtung der Zustände auf einem Weg)
 - (W1) Jede Zustandsformel ist eine Wegformel.
 - (W2) Mit ϕ, ψ sind auch $(\neg \phi)$ und $(\phi \vee \psi)$ Wegformeln.
 - (W3) Mit ϕ, ψ sind auch $X \phi$ und $\phi U \psi$ Wegformeln.

Bemerkung: CTL ergibt sich mit (Z1)-(Z3) sowie

(W) Wenn ϕ, ψ Zustandsformeln sind, so sind $X \phi$ und $\phi U \psi$ Wegformeln.

Semantik für Temporale Logiken

Allgemein Kripke-Semantik: Modell $[G, \beta]$
mit Zeitstruktur $G = (Z, R)$ und Belegung $\beta : AV \times Z \rightarrow \{true, false\}$

$G = (Z, R)$ mit Knoten-Menge (Welten, Zuständen) Z
und Kanten-Menge $R \subseteq Z \times Z$
 R ist Sichtbarkeitsrelation (accessibility relation):
hier: z_0 zeitlich vor bzw. nach z , falls $R(z, z_0)$

i.a. als (partielle) Ordnung: transitiv, reflexiv, antisymmetrisch

Ausgangspunkt: Aktuelle Welt des „Agenten“

$\beta(p, z)$ = Wahrheitswert der Aussagenvariablen p im Knoten z .
 $\beta(p, z) = true$ bedeutet: p ist wahr in der „Welt“ z

Semantik für CTL*

Erfüllbarkeitsrelation \models
induktiv definieren gemäß Syntax-Definition Z1-Z3, W1-W3

$\langle M, z \rangle \models \phi$: $\langle M, z \rangle$ erfüllt Zustandsformel ϕ
 ϕ ist wahr in der Welt z im Modell $M = [(Z, R), \beta]$

$\langle M, w \rangle \models \phi$: $\langle M, w \rangle$ erfüllt Wegformel ϕ
 ϕ ist wahr auf dem Weg w im Modell $M = [(Z, R), \beta]$

dabei ist $w = z_0 z_1 z_2 \dots$ ein Weg in $G = (Z, R)$

Semantik für CTL *

- (Z1) $\langle M, z_0 \rangle \models p$ gdw. $p \in AV$ mit $\beta(p, z_0) = \text{true}$
 (Z2) $\langle M, z_0 \rangle \models \neg \phi$ gdw. nicht $\langle M, z_0 \rangle \models \phi$
 $\langle M, z_0 \rangle \models \phi_1 \vee \phi_2$ gdw. $\langle M, z_0 \rangle \models \phi_1$ oder $\langle M, z_0 \rangle \models \phi_2$
 (Z3) $\langle M, z_0 \rangle \models E \phi$ gdw. Ein Weg $w = z_0 z_1 z_2 \dots$ in M exist.
 mit $\langle M, w \rangle \models \phi$
 $\langle M, z_0 \rangle \models A \phi$ gdw. Für alle Wege $w = z_0 z_1 z_2 \dots$ in M
 gilt $\langle M, w \rangle \models \phi$
 (W1) $\langle M, w \rangle \models \phi$ gdw. $w = z_0 z_1 z_2 \dots$ und $\langle M, z_0 \rangle \models \phi$
 (W2) $\langle M, w \rangle \models \neg \phi$ gdw. nicht $\langle M, w \rangle \models \phi$
 $\langle M, w \rangle \models \phi_1 \vee \phi_2$ gdw. $\langle M, w \rangle \models \phi_1$ oder $\langle M, w \rangle \models \phi_2$
 (W3) $\langle M, w \rangle \models X \phi$ gdw. $w = z_0 z_1 z_2 \dots$ und $\langle M, z_1 z_2 \dots \rangle \models \phi$
 $\langle M, w \rangle \models \phi_1 U \phi_2$ gdw. $w = z_0 z_1 z_2 \dots$ in M und es gibt $i \geq 0$
 mit $\langle M, z_i z_{i+1} \dots \rangle \models \phi_2$ und
 $\langle M, z_k z_{k+1} \dots \rangle \models \phi_1$ f.a. k mit $0 \leq k < i$

H.D.Burkhard, HU Berlin Vorlesung MMKI
Sommer-Semester 2007 Modale Logik

49

Absichten (Intentions)

Unterschiedliche Begrifflichkeiten (Philosophie):

- Ziele
- Absichten
- Wünsche
- Pläne

Beschreiben: Handlung vs. Zustand

Hier „Absicht“ im Sinne von

- Gewollte Handlungen (self commitment)
(zum zukünftigen Erreichen eines gewollten Zustandes)
- Plan

Unterschied:

- Erreichen eines Zustands (achievement)
- Sichern von Bedingungen (maintenance)

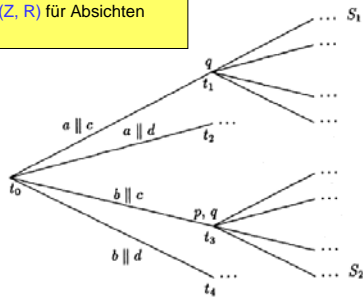
H.D.Burkhard, HU Berlin Vorlesung MMKI
Sommer-Semester 2007 Modale Logik

50

Modell für Absichten

Beispiel aus M.P. Singh: Multiagent Systems

Intentions als präferierte zeitliche Abläufe
z.B. mit Zeitmodell $G = (Z, R)$ für Absichten
 $\text{intend } p := AF p$



H.D.Burkhard, HU Bei
Sommer-Semester 20

Modell für Absichten

Konsistenzforderung bzgl. Annahmen z.B. beschreiben durch :

- Intention-Modell \subseteq Belief-Modell (ebenfalls als Zeitmodell)

$$G_{\text{intend}} \subseteq G_{\text{bel}}$$

$$\text{intend } p (= A_{\text{intend}} F_{\text{intend}} p) \rightarrow E_{\text{bel}} F_{\text{bel}} p$$

Weitere Notationen bzgl. Aktionen und Agenten:

does, done, succeeds, ...

(Wissen über) Fähigkeiten:

can, ...

H.D.Burkhard, HU Berlin Vorlesung MMKI
Sommer-Semester 2007 Modale Logik

52

Absichten (Intentions)

Rationalitätseigenschaften:

- Möglichkeit zum Erfolg (Konsistenz zu Annahmen bzgl. Zukunft, Future directed belief)
Agent soll nichts unerreichbares beabsichtigen und ggf. aufgeben
(Problem: wann genau aufgeben?)
- Konsistenz von parallelen Absichten eines Agenten (folgt aus 1)
- Konsistenz bzgl. notwendiger Mittel zu einem gewollten Zustand
- Fähigkeiten zur möglichen Erreichung des gewollten Zustands müssen vorhanden sein

Begrenzte Fähigkeiten, aber optimistisch

H.D.Burkhard, HU Berlin Vorlesung MMKI
Sommer-Semester 2007 Modale Logik

53

Absichten (Intentions)

Commitment (Verpflichtung) und Persistenz (Dauerhaftigkeit):

(Cohen-Lesque: Intention = choice with commitment)

- Commitment des Agenten auch bei veränderter Umwelt
Aber: Nicht mehr bei Unerreichbarkeit des gewollten Zustands
- Aufhören, wenn gewollter Zustand erreicht ist.

Absichten sind nicht notwendig erfolgreich

(future belief: gewollter Zustands möglich, aber nicht notwendig)

H.D.Burkhard, HU Berlin Vorlesung MMKI
Sommer-Semester 2007 Modale Logik

54

Absichten (Intentions)

Strittig:

Abschluss bzgl. logischer Konsequenz bzw. angenommenen Folgen (believed effects)

Sind logische Folgerungen der Absichten auch beabsichtigt?

(Seiteneffekte)

Verantwortung für Seiteneffekte?

Verselbständigung der Seiteneffekte (Schmerz)?

Wünsche (Desires)

Schwächere Anforderungen bzgl. Rationalität als bei Absichten.

Erlaubt sind:

- Inkonsistenz bei mehrfachen Absichten
- Inkonsistenz bzgl. Annahmen
- Inkonsistenz bzgl. Mitteln

- In einigen Theorien als Vorstufe zu Intention (BDI),
- In anderen nicht vorhanden

Fazit

Benötigen Modelle

(Entwickler, anderer Entwickler, Nutzer, Agent: Unterschiedliche Perspektiven!)

Agent benötigt Datenstrukturen für

- Annahmen (belief)
- Ziele, ... (goal, ...)

Die Datenstrukturen müssen

- Weltsituationen beschreiben (zu unterschiedlichen Zeiten)
- zeitliche Aktualisierungen erlauben (eingeschränkte Inferenz, i.a. nicht volle Logik)
- bestimmte Konsistenzbedingungen erfüllen (welche?)

Fazit

Multi-modale temporale Logiken beschreiben Probleme, z.B.

- verzweigende Zeit
- mögliche Inkonsistenzen
- Konsequenzen (erwünschte, unerwünschte)

Umsetzung in klassischer Softwaretechnik durch

- Datenstrukturen
- Inferenzverfahren (i.a. schwächer als in modaler Logik)

Fazit

Agentenprogrammierung als Ausführung logischer Spezifikationen:

Concurrent MetateM (Michael Fisher)

- Arbeit analog zu Regelsystemen: Update entsprechend Umwelt
- Regeln als temporale Formeln: $P \rightarrow F$
- Interpreter protokolliert bisherigen Verlauf (history):
 - Vorbedingungen bzgl. history testen
 - Nachbedingung als commitment: Agent führt Aktionen aus
- Programm ist Spezifikation des Verhaltens.
- Agent führt Spezifikation aus und erzeugt iterativ ein Modell der Spezifikation