Was bisher geschah

Abgrenzung der Begriffe: Daten, Information, Wissen, Intelligenz
Symbolische KI: Suchprobleme mit / ohne Gegenspieler

Repräsentation des Kontextes: Zustandsübergangssystem, Zielbedingung, (Verfahren zur Bestimmung der) Werte der Zielknoten

Lösung: Pfad zu einem Zielzustand im Zustandsübergangssystem
Spielstrategie

Lösungsverfahren: Pfadsuche (informiert, uninformiert), MiniMax-Werte, α-β-Suche

Heuristische Funktionen:

- notwendig für informierte Suche
- Eigenschaften
- mitunter automatische Berechnung möglich

Beispiel für Kombination symbolischer und subsymbolischer / statistischer Verfahren

- heuristische Spielbaum-Suche (symbolisch)
- automatische Bewertung der Knoten durch Simulation mehrerer Spiele (statistisch)

Monte-Carlo-Baum-Suche (MCTS)

Entscheidungsunterstützung

Ziel der KI: intelligente Entscheidungen treffen oder vorschlagen (analog menschlichen Experten)

Entscheidung:
Auswahl einer aus mehreren Optionen abhängig von der (aktuellen) Situation, z.B.

- nächster zu expandierender Knoten im Suchbaum
- nächster Spielzug
- Einordnung von Objekten
- Diagnosen
- Kreditwürdigkeit
- Therapieansätze

Bewertung von Objekten / Situationen

Ziel: Bewertung von Objekten (Fällen) anhand bestimmter Merkmale

für Mengen O aller Objekte, W aller möglichen Werte
Funktion $f : O \to W$, z.B.

- O: Knoten in Agenda, W: Priorität $\in \mathbb{R}_{\geq 0}$
- O: mögliche Spielzüge, W: Minimax-Wert $\in \mathbb{R}$
- O: Personen, W: Alter $\in \mathbb{N}$
- O: Belegungen $\beta : P \to \{0, 1\}$, W: Wahrheitswert $\in \{0, 1\}$
- O: digitale Bilder, W: $\in 2^{\{\text{Katze,Hund,Maus}\}}$

Bewertung der Objekte anhand ihrer Merkmale, z.B.

- Position, bisherige und geschätzte zukünftige Kosten
- Eigenschaften des Spielzustandes, z.B. noch nicht blockierte Tripel
- Wahrheitswert $[\![\varphi]\!]_\beta \in \{0, 1\}$ für gegebene Formel $\varphi \in \mathrm{AL}(P)$
- Anordnung der Pixel / Farbwerte in Bildern (Matrizen)

Klassifikationsprobleme

gegeben: Objekt $o \in O$ mit Merkmalswerten m_o
Zuordnung $K \to X_K$ mit $X_K \subseteq M$ (Extension)
jede Klasse definiert durch Merkmalswerte
(z.B. Produkt von Intervallen)

gesucht: Zuordnung Objekt $o \in O$ zu Klasse K mit $m_o \in X_K$

Beispiel:
Klassifikation von Tieren

- Objekte: Tiere,
  z.B. Adler, Eidechse, Goldfisch, Hund, Katze, Maus, Schlange
- anhand von Merkmalen
  z.B. warm/kaltblütig, Anzahl Beine, Haut, lebendgebärend
- in Klassen
  z.B. Säugetiere, Fische, Vögel, Reptilien

Diagnose-Probleme

Spezialfall von Klassifikationsproblemen:

- Objekte: Fälle
- Merkmale: Fragen, Tests
- Merkmalswerte: Antworten
- Klassen: Diagnosen

Anwendungen, z.B. in

- Medizin:
  - Krankheitserkennung
  - Entscheidung für eine Therapie
  - Auswertung von Studien
- Technische Systeme:
  - Konfiguration
  - Feststellung von Störungen
  - Entscheidung für Vorgehen bei Behebung

Klassifikationsarten und -verfahren

sichere Klassifikation: klar begrenzte Klassen,
Zuordnung Objekt $\rightarrow$ Klasse immer wahr oder falsch
Methoden: z.B. Regelsysteme, Entscheidungsbäume,
Entscheidungstabellen

fallbasierte Klassifikation:
Suche nach Lösungen für ähnliche Fälle
Grundfrage: Wann sind Fälle ähnlich zueinander?
verschiedene Ansätze, z.B. Hamming-Abstand der
Merkmalsvektoren
Verwaltung großer Datenmengen (vorangegangener Fälle)
Machine-learning-Verfahren, verschiedene neuronale Netze

unsichere / unscharfe / heuristische / statistische Klassifikation
Zuordnung Objekt $\rightarrow$ Klasse nicht immer klar, graduell
Methoden: z.B.
heuristische Entscheidungsbäume und -tabellen,
Regelsysteme mit gewichteten Regeln,
Machine-learning-Verfahren, verschiedene neuronale Netze

Sichere Klassifikation

typische Verfahren:

- Mengen von Regeln ohne / mit Alternative
- Entscheidungsbäume, -tabellen, BDDs
- logische Programme (Prolog)
  logisches Programm: Menge von Fakten und Regeln
- deduktive Datenbanken (erlauben rekursive Abfragen):
  extensionale Datenbasis: Datenbank (Menge von Relationen)
  repräsentiert die Faktenmenge
  Regelsysteme als Datalog-Programm (Anfragen)
  intensionale Datenbasis: Menge aller ableitbaren Fakten

Heuristische / statistische Klassifikation

bei unvollständigem, unsicherem, unzuverlässigem Wissen

Lösung (Klasse, Diagnose) ist Verdacht
Lösungen werden bewertet und verglichen, um die / eine beste Lösung zu finden

häufig auch numerische Bewertung von

- Plausibilität
- Zuverlässigkeit, Glaubwürdigkeit
- Häufigkeit

Klassifikation durch Regeln

Kontext: Menge (aussagenlogischer) Regeln
verschiedene Repräsentationsformen möglich, z.B.

- Menge von Regeln
- Formelmenge $\subseteq \mathrm{AL}(P)$
- Entscheidungsbaum
- Entscheidungsdiagramm
- Entscheidungstabelle

Frage (Fall): Menge von Fakten als

- Formel
- Formelmenge
- Zuordnung: Merkmal $\rightarrow$ Wert

Lösung: Klasse (z.B. ja / nein) ,
evtl. mit Modell / Begründung

Problemlöseverfahren: abhängig von der Art der Regeldarstellung

Beispiel Rabatt

1. Wenn Kundenkarte vorhanden, bisheriges Zahlungsverhalten ok und und Bestellwert hoch, wird Rabatt gewährt.
2. Wenn bisheriges Zahlungsverhalten nicht ok oder Bestellwert gering, wird kein Rabatt gewährt.
3. Wenn keine Kundenkarte vorhanden ist, wird kein Rabatt gewährt, aber eine Kundenkarte angeboten.

repräsentiert durch:
Regelmenge, Entscheidungstabelle, Entscheidungsbaum, BDD

Entscheidungstabellen

kompakte Darstellung einer endlichen Menge von Regeln

Aufbau einer Entscheidungstabelle:

- Bedingungsteil
  je Bedingung eine Zeile, je Regel eine Spalte
- Folgerungs- bzw. Aktionsteil
  je Folgerung bzw. Aktion eine Zeile
  in jeder Spalte alle Folgerungen / Aktionen der entsprechenden Regel markiert.
- Regeln werden in der angegebenen Reihenfolge nacheinander ausgewertet (von links nach rechts).

Klassifikation durch Entscheidungstabellen

Kontext: Entscheidungstabelle
(repräsentiert Menge von Regeln)

Frage (konkreter Fall):
Menge von Fakten (Aussagen, Merkmalswerten)
Menge möglicher Aussagen (oder Aktionen)

Lösung: zum konkreten Fall passende Aussagen (oder Aktionen)

Problemlöseverfahren: Suche der gegebenen Aussagenkombination im Bedingungsteil der Entscheidungstabelle, Ablesen der Aussagen (oder Aktionen) im Folgerungsteil

Eigenschaften von Entscheidungstabellen

Entscheidung zwischen mehreren Alternativen
im Beispiel: Annahme / Ablehnung / Rückruf

- vollständig:
  je Kombination von Bedingungen eine Spalte
  (Wahrheitswerttabelle)
- konsolidiert (kompakte Darstellung):
  Zusammenfassung von Spalten mit denselben Aktionen
  (Wildcards in Bedingungen)

Übersetzung: Regelmenge $\leftrightarrow$ Entscheidungstabelle

- Regelmenge $\rightarrow$ Entscheidungstabelle
  (i.A. keine vollständige ET)
- Entscheidungstabelle $\rightarrow$ Regelmenge (einfach):
  jede Spalte repräsentiert eine oder mehrere Regeln
  (nach Anzahl der markierten Folgerungen, Aktionen)

Verknüpfung mehrerer Entscheidungstabellen

Menge $\{T_1, \ldots, T_n\}$ von Entscheidungstabellen

spezielle Aktion: Übergang zu einer Tabelle T_i

Ablaufsteuerung durch mehrere Entscheidungstabellen:

- Sequenz: Übergang zur Folge-ET unter allen Bedingungen
- Verzweigung: von Bedingungen abhängiger Übergang zu verschiedenen Folge-ET
- Rekursion: Übergang zur selben ET

Regeln ohne / mit Alternative

Regeln ohne Alternative: Falls A, dann B

als logische Formel: $(A \to B) \equiv (\neg A \vee B)$

Regeln mit Alternative: Falls A, dann B sonst C

als logische Formel:
$(A \to B) \wedge (\neg A \to C) \equiv (A \wedge B) \vee (\neg A \wedge C) = \mathrm{ite}(A, B, C)$
mit dreistelligem Junktor ite

$\{\mathrm{ite}, \mathbb{t}, \mathbb{f}\}$ ist eine Junktorbasis. (ÜA)

Zu jeder aussagenlogischen Formel $\varphi \in \mathrm{AL}(P)$ existiert sogar eine äquivalente Formel ψ, die nur Junktoren aus $\{\mathrm{ite}, \mathbb{t}, \mathbb{f}\}$ enthält und jede Teilformeln mit Wurzel ite die Form $\mathrm{ite}(p, \varphi_1, \varphi_2)$ mit $p \in P$ hat (ite-Formel).

Regeln mit Alternative

(Induktive) Definition der Menge aller ite-Formeln

IA: Jede Aussagenvariable, $\mathbb{t}$ und $\mathbb{f}$ sind ite-Formeln.

IS: Sind $p \in P$ eine Aussagenvariable und φ und ψ ite-Formeln, dann ist auch $\mathsf{ite}(p, \varphi, \psi)$ eine ite-Formel

Beispiele (Tafel):

- $\neg p \equiv \mathsf{ite}(p, \mathbb{f}, \mathbb{t})$
- $p \wedge q \equiv \mathsf{ite}(p, \mathsf{ite}(q, \mathbb{t}, \mathbb{f}), \mathbb{f})$
- $p \rightarrow q \equiv \mathsf{ite}(p, q, \mathbb{t})$
- $b \wedge (a \vee c) \equiv \mathsf{ite}(a, \mathsf{ite}(b, \mathbb{t}, \mathbb{f}), \mathsf{ite}(c, \mathsf{ite}(b, \mathbb{t}, \mathbb{f}), \mathbb{f}))$

Klassifikation durch Entscheidungsbäume

z.B. https://www.kzvlb.de/fileadmin/user_upload/Seiteninhalte/Corona/coronavirus_entscheidungsdiagramm.pdf

https://www.bfarm.de/SharedDocs/Downloads/DE/Medizinprodukte/Entscheidungsbaum_decision_tree_kp.pdf?__blob=publicationFile

- Kontext: Entscheidungsbaum mit
  - inneren Knoten: Merkmale
  - Kanten: Merkmalsausprägungen (Werte)
  - Blättern: Klassen
- Repräsentation des Problems (Objekt, Fall):
  fallspezifische Merkmalswerte
- Repräsentation der Lösung:
  Zuordnung des Objektes (Falles) zu einer Klasse
  z.B. Symptome zu einer Diagnose
- Problemlöseverfahren:
  Beginn in der Wurzel,
  schrittweises Verfolgen der Kanten im Entscheidungsbaum
  entsprechend der Merkmalswerte in jedem Knoten,
  Ablesen der Klasse (Aussage, Aktion) im erreichten Blatt

Binäre Entscheidungsbäume

Binärer Entscheidungsbaum:
Graph $T = (V, E)$ mit den folgenden Eigenschaften:

- Baum T hat eine Wurzel $r \in V$,
- jeder innere Knoten hat Grad 2,
  je einen mit 0 und einen mit 1 markierten Nachfolger
- jedes Blatt hat Grad 0
- jeder Knoten hat genau einen Vorgänger

und Knotenmarkierungen

- innere Knoten: Variablen (Entscheidungskriterien, Merkmale)
- Blätter: Werte oder Aktionen

Beispiel: $b \wedge (a \vee c)$

Jeder binäre Entscheidungsbaum repräsentiert eine Boolesche Funktion.

Übersetzung: Regelmenge $\leftrightarrow$ Entscheidungsbaum

- Entscheidungsbaum $\rightarrow$ Regelmenge (einfach):
  Jeder Pfad von der Wurzel zu einem Blatt repräsentiert eine Regel.
  Innere Knoten und die dazu ausgewählten Kanten bestimmen die Voraussetzungen, das Blatt die Folgerung.
- Regelmenge $\rightarrow$ Entscheidungsbaum (weniger einfach), z.B. über ite-Formel oder Entscheidungstabelle

Allgemeine Entscheidungsbäume

Verallgemeinerung binärer Entscheidungsbäume:

- Knotenmarkierung: Variable mit mehreren möglichen Werten,
- innere Knoten mit beliebig vielen Nachfolgern
  (mehrere Alternativen möglich)

Allgemeiner Entscheidungsbaum:
Baum (gerichteter azyklischer Graph) mit den Eigenschaften

- eine Wurzel,
- jeder Knoten hat genau einen Vorgänger,
- Blätter, Knotenmarkierung: Klassen
- innere Knoten mit Knotenmarkierung: Merkmal,
  (Frage nach Merkmalswert)
  ausgehende Kanten markiert mit möglichen Werten oder
  Wertebereichen dieses Merkmals (Antworten)
  Jeder innere Knoten hat so viele Nachfolger wie mögliche
  Alternativen (Werte oder Wertebereiche)

Effektivere Darstellung binärer Entscheidungsbäume

Beispiel: $\varphi \in \mathsf{AL}(\{x, y, z, u\})$ mit
$\mathsf{Mod}(\varphi) = \{W_{0000}, W_{0011}, W_{1100}, W_{1111}\}$

Schrittweise Anwendung (ebenenweise von Blättern zur Wurzel) der folgenden Operationen auf einen Entscheidungsbaum T (oder ein BDD):

Verschmelzung von Knoten , die Wurzeln gleicher Teilbäume sind
(Vereinigung der Kantenmarkierungen)

Kontraktion von Kanten zwischen einem Knoten und seinem Kind, falls dieses sein einziges Kind ist
(Vereinigung der Knotenmarkierungen, gleichzeitige Entscheidung nach mehreren Kriterien)

Ergebnis: markierter gerichteter azyklischer Graph (DAG), hat meist weniger Knoten als T

Binary Decision Diagrams

kompakte Darstellungsform für Boolesche Funktionen
(und endliche Mengen)

Binary Decision Diagram (BDD):
markierter gerichteter azyklischer Graph (DAG) mit

Markierungen:

- innere Knoten: Variablen
- Kanten: 0 oder 1

und den folgenden Eigenschaften:

- BDD hat eine Wurzel,
- jeder innere Knoten hat Grad 2,
  je einen mit 0 und einen mit 1 markierten Nachfolger
- jedes Blatt hat Grad 0

ROBDDs

Variablenordnung: totale Ordnung $<$ auf Variablenmenge

BDD heißt genau dann bzgl. der Variablenordnung $<$ geordnet, falls auf jedem Pfad von der Wurzel zu einem Blatt die Variablen in aufsteigender Reihenfolge der Variablenordnung vorkommen.

Beispiel (Tafel): φ mit $\mathrm{Mod}(\varphi) = \{W_{0000}, W_{0011}, W_{1100}, W_{1111}\}$ mit Variablenordnungen $[x, y, z, u]$ und $[x, z, y, u]$

BDD heißt genau dann OBDD (ordered binary decision diagram), wenn es bzgl. einer Variablenordnung geordnet ist

OBDD heißt genau dann ROBDD (reduced ordered binary decision diagram), wenn keine der Reduktionsoperationen (Knotenverschmelzung, Kantenkontraktion) möglich ist.

ROBDD bilden für jede Variablenordnung kanonische Repräsentation (bis auf Isomorphie eindeutig)
Größe abhängig von Variablenordnung