

Tentamen Kennissystemen

25 maart 2008

8:45 – 11:30 uur

Dit tentamen bestaat uit 3 opgaven op 2 pagina's.

Puntenverdeling:

1a	1b	1c	1d	1e	2a	2b	2c	2d	2e	2f	3a	3b	3c	3d	3e	3f	3g
4	5	2	4	6	3	5	4	3	4	5	5	3	6	2	4	4	5

Totaal aantal punten = 74

Lees de vragen nauwkeurig en let goed op dat je alle onderdelen van de vraag beantwoordt (vergeet bijvoorbeeld niet om je antwoord uit te leggen of een voorbeeld te geven wanneer daarom gevraagd wordt). Succes!

Opgave 1: Representatie van kennis

a) Welk(e) van de volgende tekens is/zijn geen symbolen volgens de definitie uit het boek?

Leg uit waarom.

- 1) $\otimes$
- 2) $\mathfrak{F}$
- 3) $\spadesuit$
- 4) $@$
- 5) L

Nummer 2, omdat dat niet iets is dat herkend en gereproduceerd kan worden.

b) Wat zegt de *physical symbol hypothesis* en waarom is deze hypothese van belang voor kennissystemen?

Letterlijk: "A physical symbol system has the necessary and sufficient means for general intelligent action.". Deze hypothese is van belang omdat kennissystemen symbolsystemen zijn. We kunnen er alleen vanuit gaan dat kennissystemen intelligente dingen kunnen doen als we deze hypothese voor waar aannemen.

De basisstructuur van een kennisstelsel bestaat naast de gebruikersinterface uit twee onderdelen.

c) Wanneer we kijken naar een regel-gebaseerd kennisstelsel, in welke van die twee onderdelen bevinden zich dan de regels?

Een KS bestaat uit een kennisbank en inferentie-machine. De regels bevinden zich in de kennisbank, het inferentie-mechanisme zegt hoe de regels gebruikt worden.

Het is zowel een voordeel als een nadeel van regel-gebaseerde systemen dat verschillende soorten kennis in hetzelfde formaat uitgedrukt kunnen worden.

d) Leg uit waarom dit zowel een voordeel als een nadeel is.

Voordeel: heel flexibel en gemakkelijk, omdat alle type kennis op dezelfde manier gerepresenteerd worden.

Nadeel: moeilijk onderhoudbaar omdat alle soorten kennis door elkaar staan en snel onduidelijk kan worden om wat voor kennis het gaat.

- e) Beschrijf drie soorten kennis die uitgedrukt kan worden in regels en geef van elk een voorbeeld.

Zie slide 31 van college 4:

- *Heuristic association*
 - *IF address = hospital THEN health = bad*
- *Causal relation*
 - *IF disease-type = infection THEN temp = high*
- *situation + action*
 - *IF disease-type = infection THEN prescription = anti-biotics*
- *logical relation*
 - *IF temp > 37 THEN fever*

Opgave 2: Medisch registratiesysteem

Voor een ziekenhuis wordt een systeem ontwikkeld waarin artsen de aandoeningen van patiënten en de uitgevoerde verrichtingen kunnen registreren. Op basis hiervan kan het systeem verschillende functies uitvoeren. Eén van de doelen van het systeem is om patiënten te groeperen in verschillende categorieën, bijvoorbeeld “hartpatiënten”, “infectieuze patiënten” of “patiënten in levensbedreigende situaties”.

- a) Wat is de naam van de (standaard) taak die dit systeem uitvoert?

Classificatie.

- b) Welke van de volgende gegevens in het systeem kunnen *in principe* onderdeel zijn van de dataruimte van deze taak? Leg je antwoord kort uit.

- 1) aantal opgenomen patiënten
- 2) gemiddelde hartslag van de patiënt in het laatste uur
- 3) symptomen van een hartinfarct
- 4) diagnose van de patiënt
- 5) behandelend arts van de patiënt
- 6) namen van bacteriën die een infectie kunnen veroorzaken
- 7) leeftijd van de patiënt

*De dataruimte beschrijft de observaties die gedaan kunnen worden. Nummer 2,4,5,7 zijn eigenschappen van individuele patiënten en kunnen dus **in principe** gebruikt worden als observaties. Nummer 1 is een totaal over alle patiënten; nummer 3 en 6 zijn stukjes kennis die gebruikt worden om observaties aan oplossingen te koppelen. N.B. Als 3 en 6 geïnterpreteerd zijn als “eigenschappen aanwezig bij patiënten” is het wel goed gerekend.*

- c) Welke criteria zouden de ontwerpers het best kunnen hanteren voor de oplossing “patiënten in levensbedreigende situaties”:

- 1) *conservative*,
- 2) *positive coverage*,
- 3) *complete explanation*, of
- 4) *probability threshold*.

Leg uit waarom.

Positive coverage beschouwt iets als een mogelijke oplossing wanneer er één of meer aanwijzingen zijn, wat in dit geval een veilige heuristiek is. Je wilt immers geen patiënten missen. Probability threshold kan ook gebruikt worden wanneer er een mate van bedreiging berekent kan worden.

Een andere taak van het systeem is om het rooster voor de verpleegkundigen te maken, waarbij rekening gehouden wordt met hun werktijden, expertise, de aandoeningen van de opgenomen patiënten. De roosterindeling gaat in slots van 4 uur.

d) Welke (standaard) taak voert het systeem in dit geval uit?

Planning (of configuratie).

e) Zou je voor deze taak kiezen voor een continue of discrete tijdsrepresentatie? Leg je antwoord uit.

Discreet, want de indeling gaat in slots van 4 uur. Het is dus niet nodig om kleinere onderscheidingen te kunnen maken.

De arbo-wetgeving schrijft voor dat een dienst een maximale duur heeft, bijvoorbeeld 12 uur.

f) Hoe noem je het deelmodel waarin deze kennis thuishoort? Geef een ander voorbeeld van kennis die in dit deelmodel zou kunnen worden beschreven.

Arrangement model. Ander voorbeeld: verplichte lunchpauzes.

Opgave 3: File-management

Stel dat Rijkswaterstaat een systeem wil ontwikkelen dat moet helpen om verkeersopstoppingen te voorkomen. De taak die dit systeem uitvoert is diagnose: er wordt een opstopping geconstateerd, de oorzaak wordt achterhaald, en er zal geprobeerd worden om het probleem op te lossen, bijvoorbeeld door alternatieve routes te suggereren.

a) Beschrijf en leg uit welke voordelen van het toepassen van een kennissysteem in dit geval relevant zijn.

Larger distribution of knowledge (in space): kennis over omleidingen worden verspreid onder automobilisten;

Easy to modify: nieuwe wegen kunnen makkelijk worden toegevoegd;

Incomplete or uncertain information: vage concepten als “langzaamrijdend” spelen een rol;

Andere voordelen met goede argumentatie zijn ook goed gerekend.

b) Er worden bij diagnose twee soorten oorzaken onderscheiden. In het geval van file kan “sneeuwval” een oorzaak zijn. Welk type oorzaak is dit?

De twee soorten oorzaken zijn “components” en “illness”. Sneeuwval is een voorbeeld van een illness, omdat het iets is wat het hele systeem doortrekt.

c) Wat beschrijft het systeemmodel bij een diagnosetaak? Geef twee voorbeelden van kennis die in dit kennissysteem in het systeemmodel zou zitten.

Causale relaties en het gedrag van een systeem. Voorbeelden in dit domein: een kaart van het wegennet, effect van opstoppingen op andere snelwegen, etc.

In dit systeem is de capaciteit van de weg een belangrijke factor. Deze wordt berekend uit de gegevens in het systeem over het aantal rijbanen van een weg en de toegestane snelheid.

d) Hoe heet kennis die niet direct is opslagen maar eerst moet worden berekend?

Afgeleiden of impliciete kennis.

e) Wat moet je tegen elkaar afwegen om te bepalen of je kennis al of niet op deze manier opslaat?

Kosten van opslag van afgeleide kennis, benodigde rekentijd om kennis af te leiden, frequentie van gebruik.

Vage begrippen zoals “motregen”, “hevige neerslag” of “laagstaande zon” spelen in dit domein een rol. Deze begrippen kunnen worden beschreven m.b.v. *fuzzy sets*.

f) Beschrijf voor alle drie de bovengenoemde begrippen wat op de x-as en de y-as van *fuzzy set* grafiekjes zou staan.

Op de y-as staat altijd de “mate van geldigheid” van het begrip. Op de x-as van de eerste twee begrippen staat “hoeveelheid regen” (bijvoorbeeld in mm/uur), op de x-as van “laagstaande zon” staat “hoek van de zon” (bijvoorbeeld t.o.v. de aarde).

g) Is het mogelijk de functie voor het vage begrip “hevige neerslag of laagstaande zon” af te leiden uit de functies voor bovenstaande begrippen? Zo ja, beschrijf of teken hoe dat moet. Zo nee, leg uit waarom niet.

Nee, want de x-as van beide begrippen is anders, dus er is geen zinvolle combinatie mogelijk.

Einde tentamen.