

Hertentamen Kennissystemen

15 februari 2011

18:30 – 21:00 uur

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

Puntenverdeling:

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

Totaal aantal punten = 83

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!

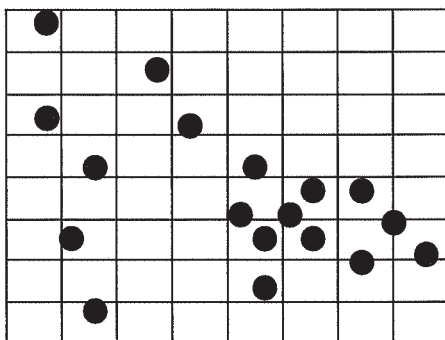
Opgave 1: Bring-me-home app

Stel dat je een toepassing voor je mobiele telefoon wilt maken die in staat is om je na een avondje stappen thuis te brengen. Nouja, het programma geeft je advies over hoe je thuis kunt komen, gegeven je huidige locatie en informatie over de bushaltes, bustijden, wandelroutes en plekken waar je 's avonds beter niet kunt komen.

- Leg uit wat het verschil is tussen data, informatie en kennis en illustreer dat aan de hand van een voorbeeld voor deze toepassing.
- Welk generiek taak model kun je onderscheiden in deze toepassing?
- Een deel van de kennis die nodig is voor deze taak bestaat uit kennis over onderdelen (*components*). Beschrijf vier aspecten van *components* die vaak beschreven worden en geef van elk een voorbeeld voor deze applicatie.

De specificatie van de oplossing kan worden gegeven in functionele termen, bijv. "een snelle route" of "een veilige route".

- Wat heb je in ieder geval nodig om de taak uit te kunnen voeren als de specificatie van de oplossing gegeven is in functionele termen?
- Gegeven de volgende verspreiding van bushaltes op de kaart. In hoeveel vakken zou een quad-tree representatie deze ruimte opdelen wanneer de drempelwaarde (*threshold*) < 4 is? Leg uit hoe je er aan komt.



- Beschrijf het algoritme dat gebruikt maakt van deze representatie om de dichtstbijzijnde bushalte te vinden.

Opgave 2: Classificatie van bedrijven

Een grote bank heeft een classificatiesysteem voor het classificeren van bedrijven in 5 verschillende categorieën van "financiële betrouwbaarheid" (de oplossingen). Betrouwbare bedrijven, met weinig

schulden en een hoge omzet, krijgen makkelijker leningen dan minder betrouwbare bedrijven. Om de betrouwbaarheid van bedrijven te bepalen worden meer dan 100 verschillende criteria gebruikt (de dataruimte).

De winst en de omzet van een bedrijf kan op veel verschillende manieren gerepresenteerd worden. In dit systeem wordt de winst en de omzet altijd teruggerekend naar een bepaalde standaard representatie.

- Wat is een andere naam voor een dergelijk standaard representatie (dus wanneer bepaalde kennis altijd terug wordt vertaald naar een standaard representatie)?
- Geef een voordeel en een nadeel van deze aanpak.
- Leg uit wat het verschil is tussen een oplossing die een dataset D verklaart (*explains*) en een oplossing die consistent is met de dataset D.
- Teken op basis van de volgende tabel met covering relaties een oplossingshiërarchie (niets ingevuld = ?).

		Oplossingen (S)						
		1	2	3	4	5	6	7
Dataset D	1	1			1	1		
	2			1		1		
	3							
	4	1		1	1	1	1	0
	5		0					0
	6				1	1	1	
	7			0		0		

Men heeft er voor gekozen om dit systeem te implementeren m.b.v. een regelgebaseerd systeem. In zo'n systeem kun je op twee manieren redeneren: vooruit en achteruit.

- Leg uit welke manier hier het best kan worden gebruikt en waarom (lees de beschrijving hierboven nog eens goed).

Helaas zijn er heel wat onzekere factoren bij het bepalen van de betrouwbaarheid van bedrijven.

- Geef de namen van de twee technieken die in het college zijn behandeld om met onzekerheid om te gaan en beschrijf in enkele zinnen wat ze inhouden.
 - Welke techniek zou je kiezen om in dit regelgebaseerde toe te passen? Leg uit waarom!
- Met behulp van beide technieken kun je iets zeggen over de zekerheid van bepaalde feiten.
- Leg het verschil uit tussen de betekenis van de uitdrukking " $P(A) = 0$ " (ene techniek) en " $CF(A) = 0$ " (andere techniek).

Opgave 3: Topsalarissen

Er is de laatste tijd veel te doen over zogenaamde topsalarissen, bijvoorbeeld van bankdirecteuren of van de directeur van de staatsloterij. In die discussie is het een probleem dat het begrip "topsalaris" nogal vaag is.

- Geef de naam van een techniek die gebruikt wordt om met vage begrippen om te gaan. Leg uit wat die techniek inhoudt.

Deze techniek maakt het mogelijk om een vaag begrip te visualiseren m.b.v. een grafiekje.

- Representeer het begrip "topsalaris" m.b.v. deze techniek en leg uit wat op de verschillende assen van de grafiek staat.
- Schets hoe zo'n grafiekje er uit ziet voor "benaderende" definities, zoals bijvoorbeeld "modaal salaris".

Je kunt de functie van gecombineerde vage begrippen bepalen op basis van de functies van de vage begrippen waaruit de combinatie is opgebouwd.

- Leg uit hoe je op basis van twee begrippen A en B en hun functies f_A en f_B de functie van de vereniging (*union*) "A of B" kunt bepalen. Schets hoe het grafiekje voor het vage begrip "modaal- of topsalaris" eruit zou zien.

Einde tentamen.