

Deeltentamen II Kaleidoscoop AI, 20 December 2005

Elke vraag is 10 punten waard. Dat levert een maximum van 100 punten op. Je eindcijfer is dus je score gedeeld door 10.

Vraag 1: Natuurlijke taal

Vanwege alle recente vertragingen heeft de NS besloten om de vertragingsberichten in het vervolg automatisch door een computer te laten omroepen. Dit computersysteem moet natuurlijk onder andere Nederlandse tijdsaanduidingen begrijpen. Om dit te realiseren wil de NS graag een grammatica hebben waarmee tijdsaanduidingen beschreven worden. Voorbeelden van geldige tijdsaanduidingen zijn:

- negen uur
- half zeven
- kwart voor acht
- tien voor negen
- dertien over half twaalf
- vijf minuten voor vijf

Vraag a: Maak een grammatica waarmee alle gangbare Nederlandse tijdsaanduidingen worden beschreven, onder andere de bovenstaande voorbeelden.

Vraag b: Gebruik je grammatica bij de vorige vraag om een parseer-boom te maken voor de tijdsaanduiding "negen over negen"

Vraag c: Het parseren van een zin volgens een grammatica kan worden opgevat als een zoekprobleem. Welke eigenschap van een grammatica bepaalt de branchingfactor van de zoekboom die ontstaat bij het parseren?

Vraag 2: Frames

Gegeven is de volgende informatie over de organisatie van onze studierichting:

Elk studie-onderdeel is een aantal studiepunten waard en heeft minstens één verantwoordelijke docent. We onderscheiden twee soorten studieonderdelen, namelijk praktika en colleges. Voor colleges moet zaalnummer en tijd vermeld worden. Praktika hebben geen zaalnummer en tijd, maar wel verplicht een website. Als de website niet vermeld wordt, verwijst een praktikum naar de standaardpagina: <http://www.cs.vu.nl/onderwijs/praktika.html>

Er zijn twee soorten colleges: hoorcolleges, waar een boek bij hoort, en leescolleges, waar een aantal artikelen bij hoort.

Twee voorbeelden van praktika zijn het praktikum Software Engineering (6 studiepunten, onder verantwoordelijkheid van dr. S. E. Neering), en praktikum Kennissystemen (4 studiepunten, onder verantwoordelijkheid van dr. K. S. Temen). Mevrouw Temen heeft haar zaakjes netjes op orde: de webpagina voor haar praktikum is te vinden op <http://www.cs.vu.nl/onderwijs/KSprak.html>. De heer Neering heeft geen webpagina gemaakt.

Vraag a: Representeer de informatie uit bovenstaande tekst in de vorm van een aantal frames. Gebruik hierbij zo goed mogelijk alle verschillende mogelijkheden van frames.

Vraag b: Breid je antwoord bij vraag 2a uit met een voorbeeld van meervoudige overerving (of laat zien dat je antwoord al een meervoudige overerving bevat). Wat is in het algemeen het probleem bij meervoudige overerving? Doet dit probleem zich voor in jouw voorbeeld?

Vraag 3: Resolutie bewijzen

In de film "Monty Python and the Holy Grail" hebben een aantal Middeleeuwse dorpingen zin in een smakelijke heksenverbranding. Ze willen zich er dan ook van overtuigen dat hun dorpsgenote Isolde een heks is. Ze gebruiken hierbij de volgende redenering:

1. Als je een vrouw bent en je brandt, dan ben je een heks
2. Isolde is een vrouw.
3. Als iets van hout gemaakt is, dan brandt het.
4. Als iets drijft, dan is het van hout gemaakt.
5. Een eend drijft.
6. Als twee dingen (x en y) hetzelfde gewicht hebben, en het ene ding x drijft, dan drijft ook het andere ding y.

Ten slotte hebben de dorpingen met behulp van een vastgeroeste weegschaal vastgesteld dat:

7. Isolde weegt even zwaar als een eend.

Helaas voor Isolde volgt hier inderdaad uit dat ze een heks is.

Vraag a: Schrijf bovenstaande feiten [1]-[7] op als logische formules

Vraag b: Vertaal je formules van vraag a naar clausal form

(Als je bij a geen antwoord hebt gegeven, beschrijf dan in het algemeen hoe zo'n vertaling van formules naar clausal form vorm in zijn werk gaat).

Vraag c: Gebruik het resultaat van vraag b om met behulp van resolutie te bewijzen dat Isolde inderdaad een heks is.

(als je bij b geen concrete formules hebt gegeven, beschrijf dan in het algemeen hoe een resolutie-bewijs in zijn werk gaat).

Vraag 4: Version Space Learning

Een verkoper in een computer winkel probeert een klant te adviseren over de aanschaf van een printer. Printers hebben twee eigenschappen (*kleur* en *resolutie*) die elk weer twee mogelijke waardes kunnen hebben: *kleur* = *ja* of *nee*, en *resolutie* = *laag* of *hoog*. Tezamen zijn er dus vier mogelijke soorten printers.

De verkoper probeert er achter te komen welke printer voor een klant het meest geschikt is. Dit doet hij door vragen te stellen aan een klant over diens voorkeur. Op basis van de antwoorden van de klant kan de verkoper hypothesen opstellen over het printerconcept waarin de klant geïnteresseerd is. Voorbeelden van printerconcepten zijn:

- kleur en lage resolutie
- geen kleur en hoge resolutie
- kleur en willekeurige resolutie
- hoge resolutie en kleur doet er niet toe
- elke willekeurige printer

Dit redeneer proces kunnen we weergeven met behulp van het version space learning algoritme.

Vraag a: Organiseer alle mogelijke printerconcepten (NB: alle mogelijke combinaties) in een concept hiërarchie op basis van hun algemeenheid.

Vraag b: In een dialoog met de klant verneemt de verkoper het volgende:

1. Eerst hoort hij dat de klant geïnteresseerd is in een hoge-resolutie zwart-wit printer
2. Daarna hoort hij dat de klant niet geïnteresseerd is in een lage resolutie kleuren printer
3. Ten slotte verneemt de verkoper dat de klant wel geïnteresseerd is in een hoge resolutie kleurenprinter.

Gebruik het version-space algoritme om te berekenen wat de meest specifieke en meest algemene hypothesen zijn die de verkoper kan opstellen over de voorkeur van de klant na elk van de antwoorden.

(Als je bij vraag a geen concept-hiërarchie hebt opgeschreven, beschrijf dan in het algemeen hoe het version-space learning algoritme in zijn werk gaat).