

VERSLAG CONFERENTIE "VERY LARGE DATA BASES"

C. Meijs¹

Een kleine 500 informatici uit 30 landen waren van 22 t/m 25 augustus j.l. in de RAI te Amsterdam bijeen. Onderwerpen die op het programma stonden hadden vooral betrekking op de object-georiënteerde technologie, deductieve databases en de integratie van kunstmatige intelligentie en databases.

De programmacommissie had 45 publicaties geselecteerd welke werden gepresenteerd in de voordrachten. Naast deze voordrachten, waarin de resultaten van fundamenteel onderzoek gepresenteerd werd, waren er ook enkele voordrachten waarin gevestigde grootheden hun visie gaven over actuele onderwerpen voor de praktijk omgeving.

Relationeel DBMS.

Chris Date ging uitvoerig in op referentiële integriteit in relationele DBMS. Het onderwerp referentiële integriteit en vreemde sleutels is reeds lang een heet hangijzer in de wereld van database ontwerpers en gebruikers. Omwille van de begripsvorming volgen hier eerst een nadere uitwerking van de begrippen referentiële integriteit en vreemde sleutels:

- Een database voldoet aan referentiële integriteit als alle vreemde sleutels consistent zijn met de bijbehorende primaire sleutels in de overige tabellen (relaties).
- Een vreemde sleutel van een tabel heeft als mogelijke waarden:
 - de waarde van een primaire sleutel in de andere tabel(len)
 - of de null waarde

M.a.w. referentiële integriteit is een kwalificatie die we kunnen toekennen aan een database of aan een DBMS. Vreemde sleutel is een eigenschap van een (set) attributen van een tabel.

Voorbeeld:

Een tabel Werknemers met als primaire sleutel het personeelsnummer en een Afdelingentabel met afd# als primaire sleutel.

Het attribuut afd# in de tabel T1 vormt een vreemde sleutel, die verwijst naar de tabel T2. In dit geval wordt dus voldaan aan de eis van referentiële integriteit.

| T1 = Werknemers | | | T2 = Afdelingen | |
|-----------------|------|------|-----------------|---------|
| Pers. nr. | Naam | Afd# | Afd# | Naam |
| 100 | jans | 10 | 10 | verkoop |
| 200 | kok | - | 20 | inkoop |
| 300 | deet | 20 | | |
| 350 | jans | 20 | | |

Nadat Chris Date ingegaan was op de historie en basisconcepten, kwam hij met aanbevelingen voor een strategie, die ervoor moet zorgen dat een database referentieel integer blijft. Verder was hij verheugd dat IBM in haar relationeel DBMS (DB2) het 'vreemde sleutel' concept ondersteunt sinds 1988 (in versie 2).

Gegevensmodelleren.

Een onderzoeksgebied dat duidelijk al vruchten heeft afgeworpen is het gegevensmodelleren. Na het relationele model van Codd in de zeventiger jaren zijn er sindsdien verschillende krachtige technieken voor (semantische) gegevensmodellering ontwikkeld. Shir Nijssen presenteerde een effectieve methode voor de informatie-analyse. De methode is gebaseerd op het vastleggen van relaties in een conceptueel gegevensmodel. Hierbij wordt uitgegaan van de informatiebeschrijving in natuurlijke taal. Middels synthese kan vanuit dit ontwikkelde conceptueel gegevensmodel gemakkelijk een relationeel schema in de optimale normaalvorm verkregen worden. Dat wil zeggen dat normalisatie van het verkregen gegevensmodel overbodig is geworden. De door Prof. Nijssen gepropageerde methode (NIAM) wordt al redelijk veel in de praktijk toegepast.

Deductieve en Object georiënteerde Databases

Deductieve databases zijn gebaseerd op 1^o orde logica. In het betreffende onderzoek wordt vaak een logische programmeertaal (PROLOG) gebruikt voor het redeneren tbv. o.a. queries, integriteitsvoorwaarden en views. De koppeling van deze programmeeromgeving met een database geeft in principe interessante mogelijkheden: response aan hoog-expressieve queries, checken van validatietregels in de Knowledge Base etc.

Object georiënteerd is een veel gehoord en geliefd onderwerp in de database wereld. Het idee is dat vele objecten (in o.m. CAD/CAM, hypertext) voor relationele systemen te complex zijn. De opzet is om krach-

1) C. Meijs is werkzaam bij de vakgroep Informatica van de Landbouwniversiteit Wageningen, via Email: luwrvd::meijs of meijs@rcl.wau.nl

tige en flexibele systemen te ontwikkelen. Enkele kenmerken van object georiënteerde databases zijn:

- Data abstractie voor de constructie van nieuwe typen:
Extensie mogelijkheden zijn groot (behalve naast de relationele typen bijv. ook stacks)
- Overerving ('inheritance') bij type hiërarchieën:
Bevordert het gebruik van gemeenschappelijke code
- D.m.v. verzenden van boodschappen worden bepaalde modules met code uitgevoerd.
Er is echter (nog) geen eenduidigheid over de verdere karakteristieken van de object-georiënteerde programmeeromgevingen.

Ook zijn de verwachtingen omtrent de (mogelijke) resultaten nog duidelijk verdeeld.

Zal de object-georiënteerde benadering in de negentiger jaren net zo geaccepteerd worden als de relationele technologie nu? Deze vraag wordt door de meeste deskundigen ontkennend beantwoord. Enkele argumenten:

- Er is geen significante onderliggende theorie voor de onderbouwing (zoals de relationele algebra voor de relationele systemen).
- De mixture van logische en implementatie concepten in de object-georiënteerde databases past niet in de algemeen aanvaarde drie-schema benadering volgens ANSI/SPARC.

Tegenover deze nadelen staat dat enkele object-georiënteerde concepten prima toepasbaar zijn bij de implementatie van complexe objecten. Als programmeertaal wordt hiervoor momenteel vaak C⁺⁺ of Smalltalk gebruikt.

Kunstmatige Intelligentie

In vele onderzoeken wordt getracht om succesvolle concepten uit de kunstmatige Intelligentie te integreren in de databases (en vice versa). Er zijn tussen deze twee vakgebieden verschillende overeenkomsten waar te nemen. Bijvoorbeeld semantische netwerken in AI en semantische datamodellen in DB.

Daarnaast zijn er ook verschillen zoals de systeemcontrole, welke in Expertsystemen vaak vrij simpel is en die in DB complex is (recovery, security etc.).

Zoekmethoden daarentegen zijn in DB simpel (1e orde logica) terwijl ze in Expertsystemen veelal uit hogere orde logica bestaan en heuristisch van aard zijn.

Michael Brodie verwacht veel van de 'kruisbestuiving' tussen AI en DB, maar het succes is volgens hem sterk afhankelijk van de vooruitgang die geboekt gaat

worden in de Logica en de Cognitieve Wetenschappen. Deze onderzoeksgebieden zullen het opstellen van formele modellen moeten ondersteunen.

Tot Slot

Een toekomstbeeld dat Vannevar Bush in 1945 schilderde is tegenwoordig werkelijkheid geworden. We kunnen zittend achter een werkstation verschillende informatiebronnen aanroepen.

Voor de toekomst is de uitdaging om m.b.v. netwerken verschillende heterogene databases te koppelen voor de bouw van Knowledge Bases. Wellicht zullen hierbij verschillende technieken en concepten zoals op de conferentie gepresenteerd genoemd, gebruikt gaan worden.

Vrijdag 25 augustus werd de succesvolle conferentie besloten. De volgende (16^e) VLDB conferentie zal in augustus 1990 in Australië gehouden worden. □