

Spatial databases: van research naar GIS-praktijk

Bart Kuijpers en Jan Paredaens
Universitaire Instelling Antwerpen*

Gedurende de afgelopen twee decennia is het aantal computertoepassingen waarin gebruik wordt gemaakt van ruimtelijke informatie alsmaar toegenomen. In deze toepassingen werkt men dan vooral met één-, twee- of drie-dimensionale ruimtelijke objecten. Temporele of historische gegevens kunnen aanzien worden als één-dimensionale ruimtelijke objecten. In robotica-toepassingen beschouwt men dan weer de drie-dimensionale objecten waartussen een robot navigeert. In het merendeel van deze ruimtelijke applicaties wordt echter met twee-dimensionale objecten gewerkt. We denken hier aan CAD-CAM, VLSI-design, architecturale toepassingen, visuele perceptie en autonome navigatie, medische beeldverwerking, maar vooral aan geografische informatiesystemen (GIS).

Omdat in al deze toepassingen de ruimtelijke gegevens dikwijls op een uniforme wijze moeten behandeld worden (dit wil zeggen ondervraagd, aangepast, afgedrukt, enz.) is het gebruik van een *database* als sleutelonderdeel van een software-systeem voor zo'n toepassing aangewezen.

Kunnen databasesystemen, die toch typisch in kantoor- en administratietoepassingen gebruikt worden, het stockeren, manipuleren en bevragen van ruimtelijke gegevens aan? Om deze vraag te beantwoorden blikken we eerst even terug.

De nog jonge geschiedenis van het database-onderzoek wordt gekenmerkt door een gestage evolutie naar steeds complexer wordende toepassingen. In de zeventiger jaren spitste het database-onderzoek zich toe op vlakke, relationele gegevensbanken, gebaseerd op een zeer rudimentair wiskundig model, dat de basis is geworden voor kantoortoepassingen. Onvrede over

*Adres: Universiteit Antwerpen (UIA), Departement Wiskunde en Informatica, Universiteitsplein 1, B-2610 Antwerpen, België. E-mail: {kuijpers,pareda}@uia.ua.ac.be.

de afwezigheid van structuur in dit model, nodig om meer geavanceerde toepassingen aan te kunnen, heeft geleid tot de introductie van complexe objecten en geneste relaties in de tachtiger jaren en tot de studie van object-georiënteerde gegevensbanken in het begin van dit decennium.

Zeer recent tekent er zich echter een totaal nieuwe evolutie in het fundamenteel database-onderzoek af: in plaats van bijna uitsluitend “general purpose” databasemodellen te beschouwen waarin de semantiek toegekend aan objecten en waarden minimaal is, worden steeds vaker eigenschappen van de objecten die men wenst te beschouwen of van de toepassingen die men beoogt, in het database-model verwerkt zodat de datamanipulatietalen hiervan gebruik kunnen maken. Dergelijke databasemodellen hebben uiteraard een enger toepassingsgebied, maar zijn beter in staat de werkelijkheid te modelleren. Het historisch oudste voorbeeld van een dergelijke klasse van “special purpose” databasemodellen zijn de temporele databases. De zeer recente ontwikkeling van constraintdatabases, geometrische en geografische databases, text-gebaseerde databases en Internet-gebaseerde gegevensmodellen is een illustratie van de huidige trend in het database-onderzoek.

De oorzaak van deze snelle evolutie in het database-onderzoek is de proliferatie van complexe databasetoepassingen die betrekking hebben op zeer specifieke datatypes, eigen aan de beoogde toepassing: GIS, navigatie systemen, verwerking van satelietgegevens, het human genome project, CAD-CAM, electronic document processing, on-line encyclopedia, on-line analytic processing (OLAP), multimedia-toepassingen en informatieverwerking op het Internet. Steeds nadrukkelijker worden databases cruciale componenten van zulke toepassingen.

Spatiale databases

We spitsen onze aandacht nu toe op die nieuwe databasetoepassingen die geometrische gegevens behandelen. Güting beschrijft in [3] een *spatiale database* als een databasesysteem dat ruimtelijke datatypes aanbiedt in zijn datamodel en bevragingstaal en dat de implementatie van ruimtelijke datatypes ondersteunt door spatiale indexering te voorzien en efficiënte algoritmes voor de spatial join te geven.

Een spatiaal databasesysteem moet

- een elegant kader ter beschikking stellen waarin thematische (of klas-sieke) informatie met ruimtelijke informatie kan gecombineerd worden;

- zo algemeen mogelijk zijn en niet specifiek ontworpen zijn voor één toepassingsgebied;
- een formeel gedefiniëerde semantiek hebben die gesloten is onder verzameling theoretische, geometrische en topologische operatoren en die gedefiniëerd is in termen van een eindige representatie;
- onafhankelijk zijn van een specifiek database management systeem (DBMS) maar wel co-opereren met een willekeurig DBMS;
- gebruik maken van efficiënte implementatietechnieken, in het bijzonder voor operaties op n -dimensionale objecten; en
- moet een up-to-date visuele user interface hebben en een gateway naar multimedia.

De eerste spatiale database systemen ondersteunen niet al deze criteria. Hun hoofdbekommernis was om bestaande, veelal relationele database management systemen op nogal triviale manieren uit te breiden door ruimtelijke datatypes te voorzien en door SQL uit te breiden op een ad-hoc manier. Er was, en er is nog steeds, een gebrek aan fundamentele inzichten betreffende ruimtelijke datatypes. In traditionele database systemen is het duidelijk welk deel van de information retrieval behandeld wordt door het DBMS, en welk deel door de applicatie-software wordt beheerd. Een projectie, bijvoorbeeld, is deel van de bevragingstaal, de berekening van een standaard deviatie is dat niet. Hoe kunnen we zulk onderscheid maken in de context van spatiale databases? Welke zijn de datamodellen die we als geschikt kunnen beschouwen voor spatiale DBMS en welke zijn de typische features en eigenschappen van spatiale datamanipulatietalen?

Wij geloven, dat een dieper inzicht in de karakteristieken die spatiale database systemen onderscheiden van traditionele DBMS hierbij fundamenteel is. Om het met de woorden van [2] te zeggen:

“The challenge for the developers of DBMSs with spatial capabilities lies not so much in providing yet another special-purpose data structure that is marginally faster when used in a particular application, but in defining abstractions and architectures to implement systems that offer generic spatial data management capabilities and that can be tailored to the requirements of a particular domain”.

Bij dit onderzoek is kennis nodig van tenminste drie verschillende domeinen: database- en informatiesystemen, geografie en cartografie, abstracte en computationele meetkunde en topologie. Bijdragen van elk van deze domeinen zijn onontbeerlijk om tot een solide theorie van spatiale informatiesystemen te komen.

Het polynomiale constraintmodel voor spatiale databases

We herhalen hier een voor de hand liggende opmerking die typisch voor ruimtelijke gegevens geldt en die minder van toepassing is op de eerder geciteerde toepassingen die gebruik maken van discrete gegevens: ruimtelijke gegevens zijn in het algemeen oneindige verzamelingen van punten die uiteraard op een eindige en effectieve manier dienen voorgesteld te worden (in het geheugen van een computer) zodat een efficiënte behandeling van deze ruimtelijke gegevens mogelijk wordt.

Hiervoor zijn verschillende voorstellen gedaan. Er zijn uitbreidingen voorgesteld van het relationeel model met (ad-hoc) datatypes zoals punt, lijn en polygoon waarmee men voor specifieke GIS-toepassingen database oplossingen kan bieden. Sinds een vijftal jaren is er echter een uitbreiding van het relationele model onderzocht waarbij men toelaat een zeer algemene klasse van ruimtelijke figuren te beschouwen, namelijk alle figuren die tot de *elementaire meetkunde* behoren. Figuren worden dan beschreven door middel van zogenaamde polynomiale constraints, met name logische combinaties van veeltermongelijkheden die betrekking hebben op de coördinaten van de punten die tot de figuur behoren.

In de context van het polynomiale constraintmodel bestaat een spatiale database uit een eindig aantal spatiale relaties waarbij iedere relatie bestaat uit een eindig aantal tuples van de vorm

$$(TA_1, \dots, TA_n; SA),$$

waarbij de TA_i klassieke, thematische attribuutwaarden zijn en SA een spatiale attribuutwaarde is. Dit laatste attribuut wordt eindig voorgesteld door een polynomiale constraintformule. Als voorbeeld geven we het tuple

$$(\text{sector, geel}, (x^2 + y^2 < 1) \wedge (\neg x < 0 \wedge \neg y < 0))$$

waarbij als ruimtelijke figuur het kwart van de eenheidscirkel in het eerste kwadrant van het vlak wordt gegeven samen met een waarde van het naamen van het kleurattribuut.

Hetzelfde constraintparadigma kan bovendien worden gebruikt om bevragingstalen te construeren waarmee geometrische gegevensbanken kunnen worden gemanipuleerd. Zo kan men bijvoorbeeld door de “calculus” formule

$$(\exists a)(\exists b)(\exists c)(\neg(a = 0 \wedge b = 0) \wedge ((\forall x)(\forall y)(S(x, y) \leftrightarrow ax + by + c = 0))),$$

de vraag uitdrukken of een vlakke ruimtelijke figuur S een rechte lijn is.

Andere zeer belangrijke klassen van “general purpose” datamanipulationstalen zijn procedurele talen (zoals algebraïsche talen) en deductieve of regelgebaseerde talen (zoals Datalog). De integratie met constraintdatabases is echter verre van triviaal. Bestaande procedurele talen voor ruimtelijke gegevensbanken zijn vaak gebaseerd op enkele ad-hoc gekozen geometrische operaties, die met behulp van technieken uit de computationele meetkunde efficiënt worden geïmplementeerd. Integratie van constraintdatabases met procedurele talen kan enerzijds een meer objectieve set operaties opleveren voor procedurele talen en anderzijds een efficiënte implementatie voor declaratieve talen. De integratie met deductieve talen, die sterk hebben bijgedragen tot het inzicht in de concepten en de implementatie van kennisbanken, is momenteel nog erg problematisch. De beschrijving van de semantiek, de detectie van het determinisme en de eliminatie van oneindige berekeningen vormen hier de grote struikelstenen.

We kunnen hier besluiten door te stellen dat deze werkwijze met constraints leidt tot de integratie van technieken uit constraint programming in het database-onderzoek. De aldus bekomen constraintdatabases hoeven bovendien niet beperkt te blijven tot de geometrische context, die hiervan slechts één—zij het voorname—interpretatie is.

Een zeer groot probleem is uiteraard de efficiënte implementeerbaarheid van dit datamodel. Hoewel het volledige model in theorie implementeerbaar is, steunen alle bestaande implementaties op lineaire polynomen, die wel efficiënt berekenbaar zijn. Spijtig genoeg moet men dan de spatiale gegevens benaderen met lineaire figuren wat niet altijd eenvoudig is en het propageren van fouten als gevolg heeft. In sommige prototypes kan men nu echter bepaalde krommen van een hogere graad gebruiken, zoals cirkel segmenten, wat de boven vermelde benadering realistischer maakt.

Spatiale genericiteit

Een ander probleem bij het definiëren van datamanipulatietalen voor spatiale databases vloeit voort uit het feit dat de voor de hand liggende declaratieve talen geen rekening houden met het geometrisch karakter van de ruimtelijke gegevens. Deze talen formuleren immers constraints op de coördinaten, wat impliceert dat behoudseigenschappen moeten gelden onder geschikte coördinatentransformaties, afhankelijk van de beschouwde meetkunde (projectief, affien, Euclidisch, enz.). Deze behoudseigenschappen houden in dat de objecten in een geometrische gegevensbank maar tot op zekere hoogte geïnterpreteerd worden. Ook hier dringt een integratie met de theorie van “general purpose” gegevensbanktalen zich op, die de gegevens in het geheel niet interpreteren. Om deze laatste eigenschap te vatten, hebben Chandra en Harel [1] reeds in 1980 het concept van *genericiteit* ingevoerd, dat thans algemeen erkend wordt als het essentiële onderscheid tussen gegevensbanktalen en algemene programmeertalen. Recent zijn wij nagegaan hoe de genericiteit kan aangepast en gediversifieerd worden voor geometrische gegevensbanken.

Chandra en Harel hebben onderzocht welke databasebevestigingen men in de context van het relationeel databasemodel als “redelijk” kan beschouwen. Zij noemen een databasebevestiging redelijk als het een Turing-berekenbare functie is die bovendien generisch is. Met een generische functie bedoelen Chandra en Harel een functie waarvan het resultaat onafhankelijk is van de interne representatie van de gestockeerde data (meer formeel betekent dit dat zulk een functie invariant is onder isomorfismen, m.a.w., permutaties van het domein van de database).

Ook voor spatiale databases is het begrip “genericiteit” onderzocht en het blijkt dat dit in de ruimtelijke context opsplitst in een hiërarchie van genericiteitsbegrippen die elk afhankelijk zijn van de geometrie waarmee men de ruimtelijke data wenst te interpreteren en van de geometrische eigenschappen die gebruikt worden in een specifieke toepassing.

We illustreren dit aan de hand van drie voorbeelden van uiteenlopende toepassingen.

In sommige applicaties, zoals temporele databases waarin men met punten en tijdsintervallen werkt, is de ordening van de punten in de tijd en de lengte van tijdsintervallen van belang. Het is daarom “redelijk” hier enkel translaties van de één-dimensionale ruimte als permutaties van het onderliggend domein te beschouwen. “Geef de gebeurtenissen die 10 minuten na gebeurtenis x hebben plaatsgegrepen” is een voorbeeld van zo’n gener-

ische vraag, maar de vraag “Geef de gebeurtenissen in 1998” is niet generisch van dit type.

In vele GIS applicaties (zoals kadaster applicaties) is de exacte lengte en grootte van ruimtelijke objecten van belang. Voor zulke applicaties zijn isometrieën (afstandbewarende transformaties) van het vlak het best geschikt om als permutaties te beschouwen. GIS-databasebevragingen zullen dan antwoorden geven die enkel Euclidische eigenschappen van de ruimtelijke data betreffen en niet afhankelijk zullen zijn van het gekozen coördinatensysteem. “Geef alle steden op minder dan 1 kilometer van de Schelde” is hier generisch maar “Geef alle Steden ten noorden van Antwerpen” is niet generisch.

In nog andere toepassingen is men enkel geïnteresseerd in de relatieve positie van ruimtelijke objecten en hun onderlinge samenhang. Hierbij denken we bijvoorbeeld aan modelleringen zoals we die terugvinden op schematische trein- of tramkaartjes. Het is niet de bedoeling van zulke kaartjes om informatie over exacte afstanden van ruimtelijke objecten te bevatten. Deze treinkaarten ondersteunen enkel vragen over de onderlinge samenhang (verbindingen) van ruimtelijke objecten (stations). We zullen dan ook eisen dat bevragingen in dit geval enkel invariant zijn onder topologische transformaties van het vlak. In dit kader is “Geef alle steden in Noord-Holland” generisch en “Geef alle steden op minder dan 10 kilometer van Schiphol (in vogelvlucht)” niet generisch.

In de spatiale databaseliteratuur van de voorbije jaren is er uitgebreid aandacht besteed aan deze verschillende types van spatiale genericiteit. Zo zijn er bijvoorbeeld databasebevragingstalen ontwikkeld die juist (de eerste-orde, respectievelijk, de berekenbare) bevragingen aankunnen van een bepaald genericiteitstype. Deze talen geven dan een gebruikersinterface waarbinnen een gebruiker enkel vragen van een bepaald type kan stellen.

Zoals reeds eerder vermeld, zijn voor GIS toepassingen vooral databasebevragingen van belang die invariant zijn onder isometrische transformaties van het vlak. Er is bewezen dat deze vragen voor databases in het polynomiale constraintmodel geformuleerd kunnen worden in een punt-gebaseerde taal (dus in een taal die *niet* gebruik maakt van coördinaatvariabelen) met behulp van de predicaten: **between**, **equidistance**, en **unitdistance**.

Tenslotte geven we een voorbeeld van een isometrie-generische vraag naar alle namen van bewoners van huizen (in een kadasterdatabase bijvoorbeeld) die juist 1 kilometer van elkaar liggen en die kan dan geformuleerd worden

door de volgende uitdrukking:

$$(\exists n)(\exists m)(\text{Lives}(n; p) \wedge \text{Lives}(m; q) \wedge \mathbf{unitdistance}(p, q)),$$

waarbij $\text{Lives}(n; p)$ een spatiale relatie is die naast de naam van bewoners ook de locatie van hun huizen bevat.

Wij hopen de lezer overtuigd te hebben dat achter het gebruik van GIS-systemen, het aanklikken van plaatjes en het manipuleren via menus er een belangrijke theorievorming aanwezig is welke het onderwerp is van intens wetenschappelijk onderzoek. Dit onderzoek is wereldwijd en legt vooral de verbanden tussen databases, geografische eigenschappen, Euclidische meetkunde en temporele formalismen. De nog jonge ontwikkeling van de GIS-systemen zal steeds belangrijker worden vooral door te proliferatie van applicaties en het stijgende belang van gebruikersvriendelijkheid. Beide auteurs maken deel uit van ADREM, een onderzoeksgroep aan de Universiteit van Antwerpen, die zich onder andere met deze theorievorming bezig houdt en waarover meer informatie op `win – www.uia.ac.be/adrem/` beschikbaar is.

References

- [1] Chandra A., Harel D. Computable queries for relational database systems. *Journal of Comp. and System Sciences*, 21, 2, 156–178, 1980.
- [2] Guenther O., A. Buchmann. Research Issues in Spatial Databases. *Sigmod Record* Vol 19,4, 61–68, 1990.
- [3] Güting R. An Introduction to Spatial Database Systems. *The VLDB Journal*, 3, 4, 1994.