

Polygonen: het wankel fundamenteel van ruimtelijk modelleren

Binnen ons 'Geo-Database Management Centre' (GDMC), onderzoeken we verschillende geo-ICT producten, zoals geo-DBMS'en, GIS-pakketten en 'geo'-middenlaagoplossingen. Tijdens onze functionele testen en performance benchmarks kwamen we subtiel, maar fundamenteel, verschillen tegen in de manier waarop polygonen behandeld worden, zelfs daar waar het gaat om de 2D-situatie en polygonen gedefinieerd door rechte lijnsegmenten.

De gevolgen zijn nogal onaangenaam. Zo zou er bijvoorbeeld een verschillend aantal objecten geselecteerd kunnen worden wanneer dezelfde vraag ('query') wordt gesteld aan dezelfde dataset, maar in een ander systeem. Een ander gevolg is dat data verloren kan gaan tijdens het overdragen van het ene systeem naar het andere systeem, omdat polygonen geldig in het ene systeem niet geaccepteerd behoeven te worden in het andere systeem.

Het lijkt allemaal zo eenvoudig, iedereen die met geo-informatie werkt weet wat een polygoon is: een gebied begrensd door rechte lijnsegmenten (en mogelijk met een aantal gaten). Een woordenboekdefinitie van een polygoon: een figuur met meerdere, meestal meer dan vier, hoeken (en zijden) [1]. Een polygoon is het fundamentele geometrische datatype in vele ruimtelijke modellen, zoals die voor topografische gegevens, kadastrale gegevens en bodemgegevens. Waarom is het de belangrijkste geo-ICT-leveranciers nog niet gelukt om dezelfde polygonen te implementeren? Het antwoord is dat in de werkelijkheid de situatie niet zo eenvoudig is als deze op het eerste gezicht lijkt. De twee belangrijkste problemen, die de verschillen tussen systemen veroorzaken, zijn:

1. Mag de buitengrens zichzelf raken en mogelijk ook binnengrenzen (van gaten) raken en zo ja, onder welke omstandigheden?
2. De computer is een eindige digitale machine en daarom kunnen coördinaten soms iets verschillen van de werkelijke gewenste wiskundige waarde. Daarom zijn tolerantiewaarden (epsilons) nodig bij het valideren (goedkeuren) van de geldige polygonen.

Het raken van de buiten- en mogelijk ook de binnengrenzen van een enkele polygoon is gerelateerd aan de topologische analyse van de situatie. Dit is een wiskundige (geometrische) kwestie, zonder implementatieproblemen, zoals de genoemde tolerantiewaarden. Je zou verwachten dat de belangrijkste geometriestandaarden van OpenGIS en ISO hier duidelijk uitsluitel geven. De basis is een recht lijnsegment, gedefinieerd door een begin- en eindpunt. Een polygooninvoer kan worden gegeven door een ongeordende, ongestructureerde lijst van rechte lijnsegmenten. De volgende aspecten moeten geanalyseerd worden voordat besloten kan worden of dit een geldige polygoon representeert (zie ook fig. 2):

1. Welke lijnsegmenten, en in welke volgorde, zijn met elkaar verbonden?
2. Is er één of zijn er meerdere verzamelingen van verbonden rechte lijnsegmenten?
3. Zijn alle verzamelingen verbonden lijnsegmenten gesloten, ofwel: vormen ze grensrings en is elk punt in de ring verbonden met ten minste twee segmenten?
4. In geval dat één punt in vier, zes, acht, etc. segmenten wordt gebruikt, wordt dan de ring opgeknipt in resp. twee, drie, vier, etc. individuele ringen of niet? (Beide keuzen zouden geldig kunnen zijn, het is in ieder geval een situatie die in de werkelijkheid voor kan komen los van het feit hoe de ringen precies gedefinieerd worden, zie fig. 1.)
5. Zijn er kruisende segmenten (mag niet voorkomen in een geldige polygoon)?
6. Is er één ring (de buitenring), die een gebied omsluit dat alle andere (binnen)ringen bevat?
7. Zijn er geen geneste binnenringen (dit zou anders resulteren in niet-verbonden gebieden)?

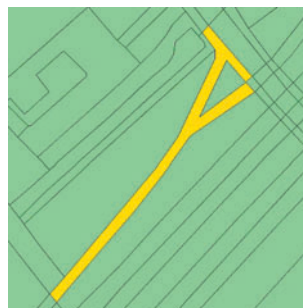


Peter van Oosterom,
co-auteurs
Wilko Quak en
Theo Tijssen,
GIS-technologie,
OTB, TU Delft.
oosterom@
geo.tudelft.nl

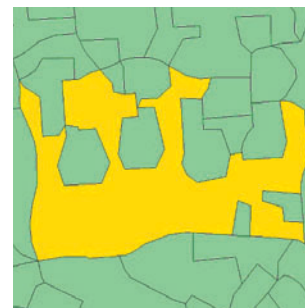
8. Zijn er ringen die elkaar raken (touch)? Dit is gerelateerd aan vraag 4, maar er is ook een situatie denkbaar dat een ring met één van de punten een andere ring raakt in het binnengebied van een rechte lijnsegment.
9. Zijn de ringen goed georiënteerd? Dat wil zeggen: tegen de klok in voor de buitenring (en hiermee een normaalvector definiërend die naar boven wijst zoals gebruikelijk in de computergrafiek: alleen gebieden met de normaalvector in de richting van de gebruiker zijn zichtbaar), en met de klok mee voor binnenringen. Merk op dat dit betekent dat het gebied dus altijd aan de linkerkant van een ring ligt.

Sommige van de vragen zouden gecombineerd kunnen worden in één test in een implementatie die bepaalt of een polygoon geldig is. Als die ongeldig is, kan hij geweigerd worden (met een foutmelding) of kan hij 'gedoogd' worden (met een waarschuwing) door het systeem. Operaties op een dergelijke polygoon kunnen echter vaak niet meer gegarandeerd correct werken en een onvoorspelbaar resultaat opleveren. In dit artikel wordt er vanuit gegaan dat gedurende de gegevensuitwisseling tussen de verschil-

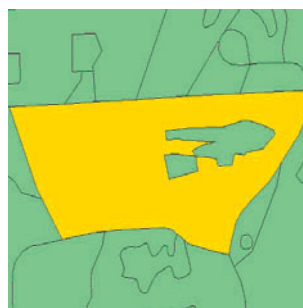
Fig 1.
Echte voorbeelden:
topografische data
1:10.000
(linksboven),
topografische data
1:50.000
(rechtsboven en
linksonder) en
kadastrale data
(rechtsonder).



Top10: buitenring raakt zichzelf in één punt.



Top50,1: twee binnenringen (gaten) raken beide de buitenring.



Top50,2: twee binnenringen raken elkaar.



Kadaster: binnenring raakt buitenring.

lende systemen er voldoende karakters of bytes worden gebruikt in geval van respectievelijk ASCII (zoals GML van OpenGIS en ISO TC211) of binaire gegevensformaten, zodat ongewenste verplaatsing van coördinaten wordt voorkomen (als gevolg van afronding/conversie) en dat het zenden- en ontvangende systeem vergelijkbare mogelijkheden hebben voor het representeren van de coördinaten.

Definitie valide polygoon

De huidige polygoondefinities, zoals in gebruik binnen de computationele geometrie, de ISO-standaard 19107 'Geographic information - Spatial schema' [2] of de OpenGIS 'Simple Feature Specification (SFS) for SQL' (kader) hebben allemaal belangrijke beperkingen en zijn niet in overeenstemming met elkaar [3] (zie overzicht met testpolygoon; fig. 2 en resultaat voor wat betreft geldigheid in de tabel 1). Hierbij moet worden gedacht aan zaken als enclaves, het raken van buiten- en binnenranden, spikes (uitstekende lijnen naar binnen of naar buiten), oriëntatie van randen (met of tegen de klok in), losse delen (multi-polygon), bruggetjes (tussen losse delen of naar enclave), tolerantie (epsilon), etc. Onze definitie is (natuurlijk) gebaseerd op ingrediënten uit de eerder genoemde definities, maar dan nu zonder aspecten over het hoofd te zien.

Een polygoon wordt gedefinieerd door rechte lijnsegmenten (in een plat vlak), allemaal georganiseerd in ringen, waarvan er precies één de buitenrand voorstelt (tegen de klok in georiënteerd) en nul of meer binnenranden (met de klok met georiënteerd). Dit impliceert dat alle punten in ten minste twee lijnsegmenten worden gebruikt. Ringen mogen elkaar niet kruisen, maar het is toegestaan dat ringen elkaar raken of zelfs voor een deel overlappen (met

OpenGIS 'Simple Feature Specification (SFS) for SQL' [OGC99, revision 1.1, 5 mei 1999]

'Een Polygoon is een plat oppervlak, gedefinieerd door één buitenrand en nul of meer binnenranden. Elke binnenrand definieert een gat in de polygoon. De onderliggende wiskundige aannamen (asserties) voor geldige polygoon zijn:

1. Polygoon zijn topologisch gesloten.
2. De randen van een Polygoon bestaan uit een verzameling van Linear-Rings, welke de buiten- en binnenranden bevatten.
3. Ringen mogen elkaar niet kruisen, twee ringen mogen elkaar raken in maximaal één punt:

$$\forall P \in \text{Polygoon}, \forall c1, c2 \in P.\text{Boundary}(), c1 \neq c2, \forall p, q \in \text{Point}, p, q \in c1, p \neq q, [p \in c2 \Rightarrow q \notin c2]$$
4. Een Polygoon mag geen doorsnijdende lijn of uitstekende lijn naar binnen of naar buiten hebben:

$$\forall P \in \text{Polygoon}, P = \text{Closure}(\text{Interior}(P))$$
5. Het binnengebied van elke Polygoon is een verbonden puntenverzameling.
6. Het buitengebied van een Polygoon met 1 of meer gaten is niet een verbonden puntenverzameling. Elk gat definieert een component van het buitengebied.

In de aannamen hierboven hebben binnengebied, buitengebied (resp. Interior, Exterior en ook Closure en Boundary) hun standaard topologische definities. De combinatie van 1 en 3 maken dat een Polygoon een gewone gesloten puntenverzameling is (Regular Closed point set).'

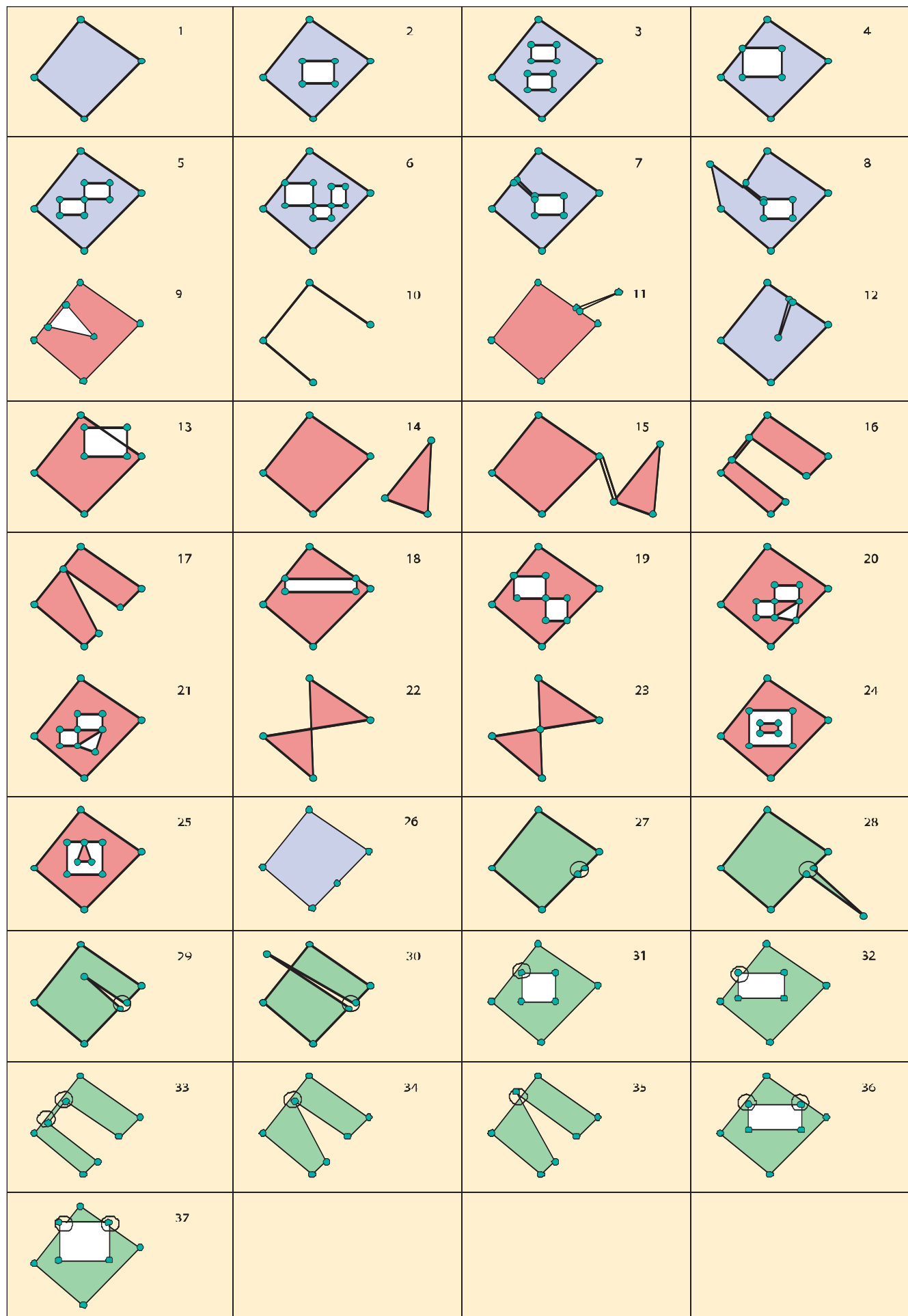


Fig. 2. Overzicht van de gebruikte testpolygoenen.

zichzelf of met andere ringen), zolang maar elk punt in het binnengebied of op de rand van de polygoon maar elk ander punt (in het binnengebied of op de rand van de polygoon) kan bereiken via een pad geheel door het binnengebied van de polygoon, oftewel, hij definieert een verbonden gebied. Hierbij vereisen sommige van de genoemde uitspraken (zoals 'ring raakt andere ring') een tolerantiewaarde in de evaluatie en daarom is deze tolerantie ook expliciet onderdeel van deze definitie.

Het is mogelijk te overwegen om de tolerantiewaarde niet te gebruiken en alleen te kijken naar de digitale representatiewaarde van de coördinaten en de rechte lijnsegmenten die hierdoor zijn gedefinieerd. Stel een situatie voor dat een punt van een binnenring heel dicht bij een buitenring komt (middendeel van lijnsegment). De situatie in de werkelijkheid is zo dat dit punt eigenlijk exact op de buitenring zou moeten liggen. Echter door de eindige hoeveelheid beschikbare getallen in een computer, is het niet altijd mogelijk om de exacte locatie te representeren [4][5], maar zal een nabijgelegen locatie gekozen worden (op basis van 'integer' of 'floating point' in plaats van reële getallen). In dergelijke gevallen is het willekeurig of het punt op het lijnsegment, dan wel links of rechts hiervan zal liggen. Indien er dus geen tolerantie wordt meegenomen, zou het kunnen zijn dat deze situatie wordt geclassificeerd als kruisende ringen (niet toegestaan) of twee disjuncte ringen (niet het geval want ringen zouden elkaar eigenlijk moeten raken). In ieder geval wordt de polygoon-situatieanalyse (ringen en geldigheid) niet goed uitgevoerd. Dit is één van de redenen waarom vele systemen met één of andere tolerantiewaarde werken. Het probleem is nu echter hoe te specificeren wat er precies met de tolerantie moet gebeuren bij het valideren van polygonen (dit deel ontbreekt nog in onze definitie). Een ander voorbeeld dat dit probleem illustreert, is testpolygoon 30: een optie om aan de tolerantie-eis te voldoen is het verwijderen van de 'naar binnen stekende uitsteeksels' en het resultaat zou nu een geldige polygoon zijn. Echter, een andere aanpak zou kunnen zijn het vergroten van de opening tussen de twee eindpunten; het resultaat is nu een niet-geldige polygoon omdat het 'naar binnen stekende uitsteeksel'

nu kruist met één van de andere zijden. Dit voorbeeld toont aan dat het moeilijk is om een niet-ambigue wijze van het werken met een epsilon-tolerantie voor te schrijven gedurende het validatieproces van een polygoon.

Een onverwacht aspect van onze definitie van geldige polygonen is dat een 'uitsteeksel naar buiten' (testpolygoon 11) resulteert in een ongeldige polygoon, omdat het niet mogelijk is om van het midden van het uitsteeksel naar elke ander punt van de polygoon te gaan via het binnenste van de polygoon. Terwijl aan de andere kant een polygoon met 'uitsteeksel naar binnen' (testpolygoon 12) wel een geldige polygoon is omdat het nu mogelijk om vanuit elk punt (ook het midden van het uitsteeksel) elk ander punt van de polygoon te bereiken via het binnenste. Iets soortgelijk als met 'uitsteeksels' doet zich ook voor met 'bruggetjes': terwijl bruggen naar binnen resulteren in geldige polygonen (testpolygonen 7 en 8), resulteren bruggen naar buiten (testpolygonen 15 en 16) in ongeldige polygonen. Zowel in het geval van 'uitsteeksels' als 'bruggetjes' is er een verschil tussen de binnen- en de buitenkant (wat men wellicht asymmetrisch zou kunnen vinden).

Geldige en schone polygoon

Het validatieproces (volgens onze definitie van geldige polygonen) zou een stuk eenvoudiger worden als kon worden aangenomen dat *geen enkel punt binnen de epsilon-tolerantie van een ander punt of zijde ligt* (die het zelf niet mede definieert); dit zal een (geldige *schone polygoon* worden genoemd. In testpolygoon 4 (en ook in testpolygonen 31 en 32) houdt dit in dat het segment waarop het punt van de andere ring geacht wordt te liggen, in tweeën wordt gesplitst. Door deze manier van modelleren kan de polygoonvalidatie worden uitgevoerd zonder toleranties. Een ander voordeel van schone polygonen is dat deze ook geen 'uitsteeksel' (naar binnen of naar buiten) meer hebben, omdat de twee eindpunten te dicht bij elkaar liggen (of zelfs gelijk zijn). Op vergelijkbare wijze worden zowel interne als externe 'bruggetjes' verwijderd. Ook worden herhalende punten verwijderd. Alvorens polygonen te versturen, moet het verzendende systeem eerst de epsilon-tolerantie-verwerking uitvoeren. Daarna kan de verzender er zeker van zijn dat het ontvangende systeem ook geldige polygonen zal ontvangen (er van uitgaande dat de coördinaten niet meer dan epsilon worden veranderd tijdens de overdracht).

De grootste afstand waarover een coördinaat kan worden verplaatst terwijl de polygoon nog steeds geldig blijft wordt de *robuustheid* van een polygoonrepresentatie genoemd door Thompson [6]. In testpolygoon 4 (zonder extra punt in buitenring), is de robuustheid nul, want een oneindig kleine verplaatsing (naar buiten) van het punt op de binnenring zorgt ervoor dat de polygoon ongeldig wordt. Echter het toevoegen van een expliciet (gedeeld) punt in de buiten- en binnenring zorgt ervoor dat de robuustheid van de representatie (van dezelfde polygoon) toeneemt tot ten minste de waarde epsilon. De robuustheid is zelfs nog groter, want het is mogelijk om de (gedeelde) punten meer dan epsilon te verplaatsen (de omvang van epsilon is zichtbaar de vorm van de open cirkel in de figuren van testpolygonen 31 en 31). De robuustheid van een polygoon kan worden berekend door de

nr	Oracle	Informix	PostGIS	ArcSDE binary	OGIS-SFS	ISO 19107	Wij
1a							
1b	Verkeerde oriëntatie					Verkeerde oriëntatie	Verkeerde oriëntatie
2							
3							
4a		Zelfsnijdende grens		Ring kruist ring			
4b	Zelfsnijdende grens				regel 3		
5							
6		Zelfsnijdende grens		Ring kruist ring			
7	Zelfsnijdende grens			Ring kruist ring			
8	Zelfsnijdende grens	Zelfsnijdende grens		Ring kruist ring			
9	Ring overlappen	Zelfsnijdende grens		Ring kruist ring	regel 3	?	Rand los van binnengebied
10	Niet gesloten	Niet gesloten	Niet gesloten		regel 1	Geen vlak	Geen vlak
11	Zelfsnijdende grens			Ring kruist ring	regel 4	?	Rand los van binnengebied
12	Zelfsnijdende grens			Ring kruist ring	regel 4		
13	Ring Raken	Zelfsnijdende grens		Drijvende binnenring	regel 3	?	Kruisende ringen
14a	Verkeerde oriëntatie				regel 5	Verkeerde oriëntatie	Twee losse gebieden
14b					regel 5	Twee buitenranden	Twee losse gebieden
15	Zelfsnijdende grens			Ring kruist ring	regel 5	Niet één homogeen gebied	Twee losse gebieden
16	Zelfsnijdende grens	Zelfsnijdende grens		Ring kruist ring	regel 5	Niet één homogeen gebied	Twee losse gebieden
17	Zelfsnijdende grens	Zelfsnijdende grens		Ring kruist ring	regel 5	Niet één homogeen gebied	Twee losse gebieden
18	Ring Raken	Zelfsnijdende grens		Ring kruist ring	regel 5	Niet één homogeen gebied	Twee losse gebieden
19	Zelfsnijdende grens	Zelfsnijdende grens		Ring kruist ring	regel 5	Niet één homogeen gebied	Twee losse gebieden
20		Zelfsnijdende grens		Ring kruist ring	regel 5	Niet één homogeen gebied	Twee losse gebieden
21					regel 5	Niet één homogeen gebied	Twee losse gebieden
22	Zelfsnijdende grens	Zelfsnijdende grens		Ring kruist ring	regel 3	Niet oriënteerbaar	Zelfsnijdende ring
23	Zelfsnijdende grens				regel 3	Niet één homogeen gebied	Twee losse gebieden
24	Verkeerde oriëntatie				regel 5	Twee buitenranden	Twee losse gebieden
25	Zelfsnijdende grens				regel 5	Niet één homogeen gebied	Twee losse gebieden
26							
27							?
28					?	?	?
29					?	?	?
30	Zelfsnijdende grens	Zelfsnijdende grens		Ring kruist ring	?	?	?
31							
32	Ring Raken	Zelfsnijdende grens		Ring kruist ring	?	?	?
33					?	?	Twee losse gebieden?
34					?	?	Twee losse gebieden?
35	Zelfsnijdende grens	Zelfsnijdende grens		Ring kruist ring	?	?	Twee losse gebieden?
36					?	?	Twee losse gebieden?
37	Ring Raken	Zelfsnijdende grens		Ring kruist ring	?	?	Twee losse gebieden?

Tabel 1. Overzicht van de resultaten van de validatie van dezelfde polygonen in verschillende systemen: indien er niets is vermeld dan wordt de testpolygoon geaccepteerd, indien dit niet het geval is dan wordt (de vertaling van) de foutmelding gegeven. (Opmerking bij testpolygoon 10/ArcSDE binary: door conversie/laadprocedure is dit testgeval niet uit te voeren.)

kleinste afstand tussen een punt (node) en een zijde (edge welke niet door dit punt is gedefinieerd) te bepalen. Een 'brute force' algoritme (alle combinaties van punten en zijden proberen) kost $O(n^2)$ rekentijd, terwijl een efficiënter algoritme dit waarschijnlijk kan berekenen in $O(n \log n)$ tijd, waarbij n het aantal punten (of zijden) in de polygoon is. Het concept van robuustheid heeft enige overeenkomsten met de 'indiscernibility' (ononderscheidbaar) relatie tussen twee representaties zoals geïntroduceerd door Worboys [7].

Testen van geo-ICT-systemen

Met een lijst representatieve testpolygonen hebben we vier verschillende geo-DBMS'en getest en de uitkomsten vergeleken met de OpenGIS, ISO en onze eigen definitie van geldige polygonen.

Polygoonvoorbeelden

In de figuren met de testpolygonen worden kleine gesloten cirkels gebruikt om punten te representeren. Soms worden twee van deze cirkels heel dicht bij elkaar getekend, dit betekent dan feitelijk dat de punten identiek zijn. Hetzelfde geldt voor twee lijnen die heel dicht bij elkaar liggen, wat dus betekent dat ze (gedeeltelijk) overlappen. Sommige figuren bevatten lege cirkels, deze geven dan een indicatie van de tolerantiewaarde (in onze voorbeeldpolygonen steeds de waarde 4000). Indien er iets anders

binnen de tolerantiewaarde ligt, dan zou het systeem moeten beslissen de polygoon aan te passen binnen de gegeven tolerantie (en feitelijk een nieuwe representatie generen). Deze nieuwe representatie kan dan geldig of ongeldig zijn afhankelijk van de specifieke configuratie. Merk op dat buitenringen van een polygoon tegen de klok in zijn gespecificeerd en dat de binnenringen met de klok mee gegeven zijn in de testpolygoon. Er zouden meer testvarianten gemaakt kunnen worden door de oriëntatie om te draaien (dit is nu alleen gedaan bij testpolygoon 1, waarbij beide varianten zijn meegenomen). Verder kunnen ringen die zichzelf raken gemodelleerd worden als één ring of als verschillende ringen. In de testpolygoon is steeds de verschillende ringenvariant gekozen, behalve bij testpolygoon 4 waar beide mogelijkheden worden getest: 4a, de verschillende ringenvariant (zonder een expliciet punt waar de ringen elkaar wel exact raken) en 4b, de 'enkele zelfrakende ring'-optie. Er is zelfs een derde variant denkbaar (nog niet getest): twee verschillende ringen maar dan wel in beide ringen het raakpunt (identiek) opnemen. Behalve deze varianten (op oriëntatie en hoe om te gaan met rakende ringen), is er nog een aantal andere testpolygoon te verzinnen. Het is belangrijk om al deze typen polygoon te verzamelen om zo de compleetheid (correctheid) van de relevante standaarden te evalueren.

Systeemtesten

We hebben de testpolygoon in verschillende geo-DBMS'en getest: Oracle, Informix, PostGIS (PostgreSQL) en ArcSDE binary. Van deze implementaties voldoen Informix en PostGIS aan de OpenGIS-specificaties voor een polygoon (SFS, level 2). Het OpenGIS Well Known Text (WKT, 'bekende tekst') formaat is gebruikt bij het invoegen van de records in de verschillende systemen. Hieronder een voorbeeld van een 'insert' SQL statement zoals is gebruikt om een correct polygoon (testpolygoon 2) toe te voegen (met uitzondering van ArcSDE, waar het laadproces anders verloopt):

```
insert into test_polygon values ('2', ST_PolyFromText('polygon(
(33300 19200, 19200 30000, 8300 15000,
20000 4200, 33300 19200),
(25000 13300, 17500 13300, 17500 19200,
```

```
25000 19200, 25000 13300))', 128992));
```

Oracle heeft als enige een aparte validatiefunctie waarin een tolerantiewaarde kan worden opgegeven (in het voorbeeld is de tolerantie waarde 4000):

```
select id, sdo_geom.validate_geometry_with_context(geom,4000)
as valid from test_polygon
where sdo_geom.validate_geometry_with_context(geom,4000) <>
'TRUE';
```

Tabel 1 geeft een overzicht van de reactie van de verschillende systemen. Deze tabel bevat ook het verwachte resultaat volgens de ISO- en OpenGIS-definities en volgens onze eigen definitie. De waarnemingen en interpretaties leiden tot de volgende algemene conclusies. Oracle Spatial (version 9.2.0.3.0) heeft een tolerantieparameter, die kan worden gebruikt bij vele operaties. In deze testen is de tolerantiewaarde steeds 4000. Experimenten met andere tolerantiewaarden geven (soortgelijke) resultaten. Indien Informix (Spatial DataBlade Release 8.11.UC1, Build 225) een polygoon met verkeerde oriëntatie vindt, dan wordt deze intern omgedraaid zonder waarschuwing. PostGIS 0.6.2 (on PostgreSQL 7.1.3) ondersteunt alleen de GeomFromText-functie (en niet de PolyFromText functie) en de geometrie wordt niet gevalideerd.

Conclusie

Zoals ook tijdens onze testen en benchmarking van verschillende geo-ICT-producten al was gebleken, is er nog geen consistent gebruik van een polygoondefinitie. Dit is zowel het geval voor de standaarden (specificaties) zelf als voor de implementatie hiervan in producten. Zowel op basis van de ISO19107-standaard alsook op basis van de OpenGIS (SFS for SQL)-implementatiespecificatie, is het soms erg moeilijk of zelfs onmogelijk om te bepalen of een polygoon geldig is of niet. Volgens onze evaluatie met de verzameling testgevallen, zijn de resultaten van (on)geldige polygoon niet altijd hetzelfde volgens de geharmoniseerde ISO- en OpenGIS-standaarden. Bovendien, zowel ISO als OpenGIS-definities zeggen niets over het belangrijke aspect van de tolerantiewaarde. Dit is een belangrijke reden waarom wij onze eigen definitie van een polygoon (met gaten) hebben gegeven. Deze is vervolgens verfijnd tot het begrip geldige schone polygoon, die geschikt zijn voor het uitwisselen tussen verschillende systemen.

Niet alleen in theorie (ISO en OpenGIS), maar ook in de praktijk (naast de genoemde Geo-DBMSsen is ook een 'middleware' product getest, LaserScan Radius Topology, wat hier niet verder is beschreven) bleken er belangrijke verschillen te bestaan bij de validatie van polygoon. Het behoeft geen verdere uitleg dat dit vervelende problemen zal geven bij het uitwisselen van gegevens, en in het ernstigste geval zelfs verlies van gegevens. Wij roepen daarom de standaardisatie-organisaties en de Geo-ICT-leveranciers op dit probleem aan te pakken en onze polygoondefinitie te overwegen.

In dit artikel werden alleen simpele polygoon op een plat vlak besproken. Echter, er kunnen in de wereld van Geo-ICT-producten (en -standaarden) toch nog een aantal complexere situaties zich voordoen:

- Multi-polygonen (d.w.z. polygonen met meer dan één buitenring, die niet met elkaar verbonden zijn);
- Ook niet-lineaire randen in de grens (b.v. cirkelbogen);
- In 3D-ruimte, maar polygonen beperkt tot op een plat vlak;
- In 3D-ruimte, maar polygonen beperkt tot op het oppervlak van een polyhedron (dus vlak in stukken);
- In 3D-ruimte, op niet-platte oppervlakken, specifiek de aardse ellipsoïde (of geoïde).

Voor al deze gevallen moet een onambigue en complete definitie (inclusief tolerantieaspecten) beschikbaar zijn. Testgevallen zouden gedefinieerd moeten worden en vervolgens zouden de producten moeten worden geëvalueerd. Nadat we dan klaar zijn met polygonen, moeten we vervolgens verder gaan met andere ruimtelijke datatypen zoals polyhedrons [Are03]. Ten slotte, tot nu toe is alleen de validatie van polygonen besproken, maar wat gebeurt er met deze (on)geldige polygonen gedurende het uitvoeren van operaties?

Samenvatting

Polygonen: het wankel fundament van ruimtelijk modelleren

Ruimtelijke modellen zijn zowel in 2D als in 3D vaak gebaseerd op polygonen. Veel Geo-ICT-producten ondersteunen ruimtelijke datatype, zoals een polygoon gebaseerd op de OpenGIS 'Simple Features Specification'. OpenGIS en ISO hebben een overeenkomst om hun specificaties en standaarden te harmoniseren. In dit artikel zijn ook enkele implementaties met elkaar vergeleken. De testresultaten onthullen belangrijke verschillen tussen de verschillende implementaties. Gebaseerd op deze ervaringen stellen we een onambigue definitie van een polygoon voor. Hoewel goed gedefinieerd, kunnen deze polygonen nog steeds problemen veroorzaken bij de uitwisseling tussen verschillende systemen (onder andere door conversie en afronding van coördinaten), daarom wordt ook het concept van geldige schone polygonen geïntroduceerd.

TREWOORDEN

GIS-technologie, gegevensverwerking, onderzoek

Summary

Polygons: the unstable foundation of spatial modelling

Spatial models are often based on polygons both in 2D and 3D. Many Geo-ICT products support spatial data types, such as the polygon, based on the OpenGIS Simple Features Specification. OpenGIS and ISO have an agreement to harmonise their specifications and standards. In this paper we discuss the relevant aspects related to polygons in these standards and compare several implementations. The test results reveal significant differences in the implementations, which causes interoperability problems. Based on these experiences we propose an unambiguous definition for polygons, which makes polygons again the stable foundation it is supposed to be in spatial modelling and analysis.

KEYWORDS

GIS-technology, data processing, research

Résumé

La qualité des données est une priorité chez Essent Netwerk

Cet article décrit l'information géographique de l'infrastructure des câbles et conduits du fournisseur d'énergie Essent Netwerk Zuid. On traite e.a. la réalisation et la mise à jour des fichiers numériques avec la

conversion des cartes de gestion, l'emploi des SIG et les notifications au Centre d'information des câbles et conduits (KLIC). Le suivi de ces notifications, les expériences avec le centre KLIC et les souhaits de Essent Netwerk Zuid au sujet du centre KLIC sont traités, ainsi que les expériences des entrepreneurs de travaux publics. Enfin on traite quelques sujets du développement de la problématique à l'échelle nationale, ainsi que l'autorégulation et la législation potentielle.

MOTS CLÉS

Technologie SIG, traitement des données, recherche

Literatuur

- [1] The Shorter Oxford English dictionary, 1973.
- [2] ISO/TC 211/WG 2, ISO/CD 19107, *Geographic information – Spatial schema*, 2003.
- [3] Oosterom, Peter van, Wilko Quak and Theo Tijssen, *Polygons: the unstable foundation of spatial modeling*, In: *proceedings of the ISPRS workshop on 'Spatial, temporal and Multi-Dimensional Data Modelling and Analysis'*, 2-3 Oktober 2003, Québec, Canada.
- [4] Goldberg, David, *What Every Computer Scientist Should Know About Floating-Point Arithmetic*, ACM Computing Surveys, Vol. 23, March 1991, pages 5-48.
- [5] Güting, Ralf Hartmunt and Schneider, Markus, *Realms: A foundation for spatial data types in database systems*. In: D.J. Abel and B.C. Ooi, editors, *Proceedings of the 3rd International Symposium on Large Spatial Databases (SSD)*, volume 692 of *Lecture Notes in Computer Science*, pages 14-35. Springer-Verlag, 1993.
- [6] Thompson, Rodney, *PhD research proposal 'Towards a Rigorous Logic for Spatial Data Representation'*. Department of Geographical Sciences and Planning, The University of Queensland, Australia, November 2003.
- [7] Worboys, Mike, *Some Algebraic and Logical Foundations for Spatial Imprecision*, in: Goodchild, M. and Jeansoulin, R (ed), *Data Quality in Geographic Information: from error to uncertainty*, Hermes, 1998.
- [8] Arens, Calin, Jantien Stoter and Peter van Oosterom, *Modelling 3D spatial objects in a Geo-DBMS using a 3D primitive*. *Proceedings of the 6th AGILE 2003*, Lyon, Frankrijk.

Dankwoord

We danken de TDN en het Kadaster voor het beschikbaar stellen van hun datasets, en Oracle, Informix, PostgreSQL/PostGIS, ESRI ArcSDE) voor het beschikbaar stellen van hun producten. Voorts willen wij de reviewers en de deelnemers aan de ISPRS-workshop en het Dagstuhl-seminar danken voor hun kritische kanttekeningen, die in dit artikel zijn verwerkt. Deze publicatie vindt plaats in het kader van het onderzoeksspeerpunt Sustainable Urban Areas, uitgevoerd door de Technische Universiteit Delft.