

15 juni 2002

Henk Donkers

Evolverende kaart

Digitale atlassen maken bij plaatsnamen soms een vreemde keus. Wanneer tijdens het cartografische proces het systeem van natuurlijke selectie wordt toegepast, is het resultaat bijna perfect.

Het maken van goede kaarten is zo moeilijk dat computers nog steeds niet in staat zijn om deze automatisch te genereren. Ze hebben nog steeds moeite met het toepassen van cartografische regels zoals het plaatsen van namen zonder dat deze elkaar overlappen. De Utrechtse informaticus Steven van Dijk heeft in zijn proefschrift voor dit probleem een oplossing gevonden en maakt daarbij gebruik van de principes uit de evolutietheorie van Darwin.

Een kaart die gegenereerd wordt door een digitale atlas als de Encarta of de WN Wereld@tlas vermeldt nogal eens andere plaatsen dan een vergelijkbare kaart uit een papieren atlas. Digitale atlassen maken rare selecties van plaatsnamen. Op een kaart van Nederland bijvoorbeeld verschijnen onbetekenende plaatsjes als Stedum, Ens, Klundert en Veen wel op het scherm, maar steden als Arnhem, Apeldoorn, Tilburg, Breda, Leiden en Heerlen niet. Wat er op de kaart komt is namelijk niet afhankelijk van het belang van een plaats, maar van de ruimte die beschikbaar is om een plaatsnaam te vermelden zonder dat deze overlapt met andere plaatsnamen. Het computerprogramma kent een aantal standaardposities voor namen en lettergroottes; plaatsnamen die een andere overlappen worden geëlimineerd. Tenminste in goede digitale atlassen, want in de slechte overlappen ook de namen nog.

De elementen op een kaart kunnen worden verdeeld in puntvormige (zoals steden en dorpen), lijnvormige (zoals rivieren en wegen) en vlakvormige (zoals landen of gebergten). Elk element krijgt een naam die zodanig geplaatst moet worden dat het element te identificeren is. Voor de positionering van namen gelden allerlei cartografische regels. Zo komt een naam van een stad bij voorkeur rechtsboven de stad te staan, wordt de naam van een rivier herhaald als de rivier erg lang is en volgt de naam van een gebied de vorm ervan (de naam Italië volgt bijvoorbeeld de vorm van de laars). De namen mogen elkaar niet overlappen omdat de kaarten dan onleesbaar worden.

Dat betekent een eindeloos geschuif met namen. Cartografisch tekenaars wisten dat proces met een pen of muis tot een goed einde te brengen en prachtige papieren kaarten te produceren. Maar ze worden steeds schaarser omdat er geen nieuwe cartografisch tekenaars meer worden opgeleid. Ze werken wel met computers, maar automatisch kaarten genereren van goede kwaliteit kunnen ze nog steeds niet.

Informaticus Steven van Dijk legt uit wat daarvan de reden is. "Als je veel elementen wilt afbeelden en als de naam van elk element op een heleboel verschillende posities kan staan, moeten er zó veel mogelijke combinaties worden doorgerekend dat zelfs heel krachtige computers dat niet aankunnen. Op zoek naar de beste oplossing slagen ze er niet in een kaart zonder namenoverlap te produceren. Omdat deze exacte methode niet goed werkt, gebruikt men heuristieken. Die nemen niet alle mogelijke oplossingen in overweging, maar alleen de veelbelovende. Digitale atlassen werken met simpele heuristieken, maar die leiden ertoe dat er een heleboel namen niet worden weergegeven die eigenlijk wel op de kaart thuishoren."

Van Dijk heeft nu een geavanceerde heuristiek ontwikkeld die wat meer tijd kost, maar wel tot veel betere resultaten leidt. Hij gebruikt een genetisch algoritme als heuristiek en mikt niet meteen op de gegarandeerd beste oplossing, maar op een aantal redelijk goede die snel te vinden zijn. De meest succesvolle kaarten gebruikt hij als basis voor de ontwikkeling van een serie nieuwe kaarten en hij gaat door tot het resultaat bevredigend is.

vijftig generaties

Van Dijks genetisch algoritme is gebaseerd op de darwinistische theorie van evolutie door natuurlijke selectie. Dat werkt als volgt. Hij laat de computer willekeurig tweehonderd kaarten maken. Dat is de eerste generatie. De computer selecteert daaruit de set kaarten met de beste eigenschappen, dus met weinig overlappings van plaatsnamen. Daaruit ontwikkelt de computer een nieuwe generatie kaarten op basis van de begrippen mutatie en crossing over. Bij mutatie verplaatst de computer een overlappende naam naar een positie in de buurt. Bij crossing over combineert hij stukken van de ene kaart met stukken van een andere kaart. Uit deze nieuwe generatie kaarten selecteert de computer wederom de beste en gaat daarmee opnieuw aan het werk. Zo evolueert de kaart. Na zo'n vijftig generaties is de kaart uitontwikkeld en voor 97 procent perfect.

Van Dijk heeft in eerste instantie een algoritme ontwikkeld voor de plaatsing van namen bij puntvormige elementen. Daar blijkt het goed te werken. Later heeft hij zijn algoritme uitgebreid voor de plaatsing van namen bij lijnvormige elementen. Ook dat bleek te werken. Wel kost het de computer nog altijd een aantal uren om een kaart te produceren. Van Dijk: "Dat is veel sneller dan de oude exacte methode, en de resultaten zijn veel beter dan bij de simpele heuristieken."

Toch zijn de kaarten die Van Dijk produceert nog lang niet zo mooi als die in een papieren atlas. Van Dijk: "Het doel van mijn NWO-project was te laten zien dat je het probleem van map labeling kunt aanpakken met genetische algoritmen en dat je daarbij allerlei aanvullende cartografische principes kunt inbouwen. Dat blijkt te kunnen. Het is nu de beurt aan softwarebedrijven om dat toe te passen."

De digitale atlassen zullen er op korte termijn niet beter van worden omdat daar bij het genereren van kaarten erg de nadruk ligt op snelheid en interactiviteit. Planologen die snel één specifieke, goed leesbare kaart met veel technische details moeten produceren, kunnen er nu al wel mee aan de slag.

Steven van Dijk: Genetic Algorithms for Map Labeling, ISBN 90-393-2864-1.

Het proefschrift is ook te downloaden:

<http://www.cs.uu.nl/people/steven/download/thesis.pdf>