**KAARTKLEUREN: Een minimaal aantal kleuren**

Casimir Boot

Universiteit van Amsterdam

emailadres@email.uva.nl

Lonneke Lammers

Universiteit van Amsterdam

emailadres@email.uva.nl

Alwin Lijdsman

Universiteit van Amsterdam

alwinlijdsman@outlook.com

tegenwoordige tijd

geen wij/ zal

**1. Inleiding**

Het minimaliseren van het aantal kleuren wanneer een kaart wordt ingekleurd is een bekend probleem. In 1852 vermoedde Guthrie dat, wanneer elk land wordt ingekleurd waarbij aangrenzende landen niet dezelfde kleur mogen hebben, er maximaal vier kleuren nodig zijn [2]. Dit wordt het Guthrie’s Problem genoemd [2].

Het eerste wetenschappelijke artikel over dit probleem werd geschreven in 1878. In 1976 werd Gurthrie’s probleem opgelost door Kenneth Appel en Wolfgang Haken [1], hiermee werd de Four Color Theorem ontwikkeld [1]. Het was de eerste grote stelling die bewezen werd met behulp van een computer [1].

In dit verslag wordt de Four Color Theorem getoetst met drie verschillende kaarten, een kaart van Pennsylvania, een kaart van een Indiase provincie en een kaart van een sociaal netwerk. Het doel is om met zo min mogelijk kleuren de kaarten in te kleuren. De hoofdvraag van dit verslag luidt daarom als volgt: hoeveel kleuren zijn er minimaal nodig om een kaart in te kleuren? De uitkomt hierop zou volgens de Four Color Theorem vier moeten zijn. Dit zal daarom getoetst worden.



Figuur : Kaart van Penssylvania en kaart van een Indiase provincie

Om de hoofdvraag te beantwoorden zijn in dit verslag twee heuristieken gebruikt die beide tot een correct antwoord komen. Dit om een beter begrip te krijgen van het probleem van het kaartkleuren.

De bruto toestandsruimte grote in dit verslag wordt geschat op , waar k gelijk is aan het aantal kleuren en n gelijk is aan het aantal landen op de kaart. Als ieder land op de kaart een andere kleur toegewezen krijgt, is er altijd een oplossing. De restrictie op de transitie is, zoals eerder gesteld, dat aangrenzende landen nooit dezelfde kleur mogen kennen en dat het minimaal aantal kleuren moet worden gebruikt.

Xxx plaatjes van de kaarten

**2. Methodes**

In deze paragraaf wordt besproken welke methodes zijn gebruikt om te komen tot een antwoord op de hoofdvraag. Om het minimaal aantal kleuren voor het kleuren van een kaart te vinden wordt er in dit verslag gewerkt met een random sampling algoritme en een zelf ontwikkeld gericht algoritme. Als eerste is het random sampling algoritme ontwikkeld besproken. Daarna wordt het gerichte algoritme besproken.

**2.1** **Random sampling algoritme**

Het random sampling algoritme is als eerste ontwikkeld omdat dit algoritme relatief gemakkelijk te ontwikkelen was. Het algoritme werkt als volgt:

Als eerste worden twee CSV-bestanden ingelezen met data over de kaart die geanalyseerd wordt. De CSV-bestanden worden met de hand gemaakt, omdat deze verschillend zij voor iedere kaart. De kaarten die hiervoor zijn gebruikt zijn de kaart van Pennsylvania, een Amerikaanse staat, een regio van India en aantal verbindingen op een sociaal netwerk. Bij de kaart van het sociale netwerk, zijn alle drie de netwerken in elkaar gevoegd. Hierdoor ontstaat er een driedimensionale kaart, want naast “buurlanden” die naast een “land” kunnen liggen, liggen landen nu ook boven en onder elkaar.

In het eerste CSV-bestand worden de namen en te gebruiken afkortingen van de landen opgeschreven. Bijvoorbeeld: ”cra,crawford”. In het tweede CSV-bestand worden vervolgens de buurlanden aangegeven. Bijvoorbeeld: “cra,eri,law,lyo”. Dit betekent dat naast de regio Crawford, de landen Erie, Lawrence en Lycoming liggen.

Door deze gemaakte CSV-bestanden weet het algoritme welke landen aan welke andere landen grenzen. De lijst met landen wordt door elkaar geschud, waardoor er willekeurig wordt begonnen met het kleuren van een land. Dit land wordt als eerste gekleurd en krijgt kleur “1”.

Vervolgens wordt er gekeken naar het tweede land in de lijst. Als dit land grenst aan het eerste land, krijgt dit land de kleur “2”. Zo wordt de gehele lijst afgegaan. Uiteindelijk wordt er gekeken naar wat de hoogste “kleur” was in de kaart. Als er vier kleuren zijn gebruikt is de hoogste waarde “4”.

Het inkleuren van de kaart wordt 1000 maal herhaald. Hieruit komen verschillende maximale waarden. De minimale waarde van deze lijst van maxima, vormt het antwoord op de hoofdvraag. Daarnaast woorden alle maxima in een CSV-bestand gestopt verdere statistische berekeningen mogelijk te maken.

xxxHerhaalbaar? – langer? Misschien veel preciezer iedere stap uitleggen (helpt bij de herhaalbaarheid van het stuk)

**2.2 Gericht algoritme: land met meeste buurlanden als eerste kleuren**

Ondanks het feit dat er goede antwoorden kwamen uit het random sampling algoritme, is er gekozen om een gericht algoritme te ontwikkelen. Dit om te onderzoeken of er een betere methode is om tot het juiste antwoord te komen.

Het gerichte algoritme heeft dezelfde basis als het random sampling algoritme. Het wijkt af van het random sampling algoritme bij het bepalen van de volgorde van de te kleuren landen. Waar het random algoritme willekeurig alle ingelezen landen door elkaar schudt, zoekt het gerichte algoritme eerst naar de landen met de meeste aangrenzende buurlanden en zet deze vervolgens vooraan in de lijst om gekleurd te worden. Landen die een gelijk aantal buren hebben worden wel in een willekeurige volgorde naast elkaar gezet.

Ook dit algoritme wordt 1000 maal herhaald. Die maximale waarden worden geteld en in een CSV-bestand gestopt voor verdere statistische analyse.

**3. Resultaten**

In deze paragraaf wordt besproken welke resultaten er uit het onderzoek naar de meeste kaartkleuren is gekomen. De uitkomsten worden besproken per kaart, waarbij er een vergelijking wordt gemaakt met tussen het random sampling algoritme en het gerichte algoritme. Onder statistieken wordt in dit verslag gezien: het antwoord op de hoofdvraag, oftewel het minimaal aantal kleuren, de gevonden maximale waarden en de gemiddelde waarden.

**3.1 Kaart van Pennsylvania**

Als eerste is de kaart van Pennsylvania geanalyseerd. Bij het random sampling algoritme kwam hier een minimumaantal kleuren van vier uit. Dit aantal is overeenkomstig met de Four Color Theorem.

**4. Conclusies**

*Hier schrijf je je conclusies, eventuele overdenkingen (hoe zou het nog beter kunnen, is het algoritme ook in andere gebieden toepasbaar).*

4 color theorem blijkbaar niet voor 3d kaarten

hoog aantal aangrenzen landen eerst kleuren lijkt effect te hebben op de oplossing, meer kans op een goede oplossing

Verder onderzoek: andere aspecten die ook effect zouden kunnen hebben, zoals aantal “landenpunten”.

Nadeel aan ons onderzoek: er is kans op een dubbele oplossing, vooral in het gericht zoeken. Want 1000x genereren maar geen check op gelijke oplossingen. (misschien hillclimber moeten doen?)

**5. Referenties**

*Als je literatuur hebt gebruikt, hier toevoegen. Als je eraan refereert in de tekst, zet je op die plek alleen [1], zodat mensen achterin de details kunnen vinden. Als je geen literatuur gebruikt, weglaten.*

*Wat ook nog kan is een dankwoord, bijvoorbeeld voor mensen die wel geholpen hebben maar geen auteur zijn, mensen die je een inzicht hebben gegeven, of administrators die je even hun supercomputer hebben laten gebruiken. Altijd naam en bedrijf noemen en zorgen dat de bedankte persoon zich er goed over voelt.*

*Als je zowel een dankwoord als een referentiesectie hebt: de referentiesectie is \*altijd\* het laatste onderdeel van je verslag.*

[1] Artificial Intelligence, a modern approach, Russel & Norvig,3rd Edition, Addison-Wesly, pg 287-387.

[2] “A packing problem with applications to lettering of maps” Michael Formann and Frank Wagner (1991) SCG '91 Proceedings of the seventh annual symposium on Computational geometry

1. *http://mathworld.wolfram.com/Four-ColorTheorem.html*
2. *http://mathworld.wolfram.com/GuthriesProblem.html*

***DONE maar goed om te checken:***

1. ***Inleiding***

*Hier schrijf je een inleiding die in elk geval kort, bondig en compleet de hele vraagstelling bevat. Ook moet er een inschatting komen van de toestandsruimtegrootte (belangrijk, daar deden we het voor), en eventuele restricties op transities in beschreven worden, maar nog niet de methodes die je gebruikt om de toestandsruimte te doorzoeken.*

Eternity II is een edge-matching puzzle uitgegeven door Tomy in november 2012. Het doel is bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla $1.000.000 te winnen.

Fig. 1: een kleine-schaal versie van Eternity II met centrumstukken, randstukken en hoekstukken. Stukken mogen 90o geroteerd worden in beide reichtingen, en twee aangrenzende stukjes ‘passen’ als de rakende kanten hetzelfde symbool hebben. De puzzel is opgelost als alle stukjes passen.

*Figuren moeten altijd genummerd zijn, en \_eigenlijk\_ moet er ook altijd naar verwezen worden in de tekst. Een goede figuur maken is een kunst. Het kan enorm bijdragen aan de toegankelijkheid van je artikel, maar een slechte figuur doet afbreuk. Iedere figuur heeft een bijschrift (of ‘caption”). Bijschriften zijn zo kort mogelijk, maar niet korter dan dat (moeilijke nuance). Als richtlijn kun je voor een eenvoudige figuur als deze met twee tot drie korte duidelijke zinnen klaar zijn.*

Ieder stukje kan op 4 manieren geplaatst worden bla bla bla bla bla bla bla bla bla bla bla. De bruto toestandsruimte van Eternity bevat daarom 81226 \* 1027 elementen. Stukjes 13 en 71 zijn identiek, net als bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla. De netto toestandsruimte bevat daarom 2341 \* 1023 elementen, dit aantal is te groot om met een exhaustive method binnen redelijke tijd tot een goed einde te komen.