

Bachelor klas Informatica, Universiteit Leiden

28 februari 2008

De *voorlopige* projectbeschrijvingen zijn als volgt:

Ramon van Dam **Titel:** Takegaki

Begeleiders: Jeannette de Graaf en Walter Kusters

Omschrijving: De Japanse puzzel “Takegaki” (hier beter bekend als Slitherlink) valt in de categorie van Nikoli puzzels, waar ook de bekende Sudoku en Kakuro varianten onder vallen. Doel van het spel is om op een raster een continue lijn te vinden met hetzelfde begin en einde, die zichzelf nergens kruist. Op enkele plekken staan getallen gegeven en die geven aan hoeveel van zijn randen (maximaal 4) een lijn hebben. Interessante gegevens die onderzocht kunnen worden zijn onder meer:

- het bepalen van de moeilijkheidsgraad van een gegeven puzzel
- NP-volledigheid
- oplossen met simpele logica
- oplossen met geavanceerdere strategieën (neurale netwerken, genetische algoritmen etc.)
- het genereren van nieuwe puzzels met een enkele uitkomst
- aanpakken van puzzels met meerdere uitkomsten

Kerem Denizmen **Titel:** IMDB Parser

Begeleider: Walter Kusters

Omschrijving: De opdracht bestaat uit twee delen, in de eerste deel kijken we of er Wikipedia pagina’s bestaan van schrijvers uit de VN Detective en thrillergids database (www.vnster.nl). De Database moet eerst geparst worden door gewoon de HTML code te doorlopen en de namen er correct uit te halen, daarna zullen deze namen in een file worden gestopt. Na deze operatie zullen we met behulp van de namen kijken of er Wikipedia pagina’s bestaan van deze namen, eerst Nederlands en als die niet bestaat engels. Dit gaan we doen door de namen achter de Wikipedia link te plakken en een web request te doen. Als die bestaat zullen we de link gewoon plakken in een file met de naam erachter, misschien ook het aantal bytes en of het engels of nederlands is, misschien met een optie om de links te sorteren op aantal bytes.

Deel 2 is veel ingewikkelder. Het gaat hier om de links (die zich in de VN detective en thrillergids database bevinden) naar de Internet Movie Database van boeken die verfilmd zijn. Deze pagina's moeten worden geparst en de (onder andere) rating, Director en acteurs moeten eruit worden gevist. Het eerste deel bestaat uit het verkrijgen van de HTML code van de pagina door middel van een request. Nadat men deze heeft verkregen is het de bedoeling om de HTML code correct te parsen en informatie eruit te halen. Mischien is het ook handig om met deze informatie wat statistieken te maken zoals meest populaire Director en acteurs, wat de films met elkaar gemeen hebben in termen van cast, langste film, korste film, welke film maatschappijen enzovoorts. Hier zal nog een precieze lijst van gemaakt moeten worden maar als concept kan men hiermee beginnen. De informatie word weggescreven naar een text file met misschien sorteer opties.

Samengevat: We willen kijken of de schrijvers van de boeken in de VN detective en thrillergids een Wikipedia pagina hebben en we willen informatie van verfilmde boeken uit de VN detective en thrillergids database ophalen van de IMDB pagina's voor de films en daarmee allerlei (interessante) statistieken weergeven.

Bart van der Drift **Titel:** 3D Worlds for Collaborative Work

Begeleiders: Joost Broekens en Walter Kusters

Omschrijving: In this project you will look at one specific aspect of computer supported collaborative work (CSCW), i.e., the potential role of 3D worlds for collaborative remote work. An important question in this project is to what extend such worlds can give the impression of presence. (...) Your project has two phases. A literature survey (1.5 months) and a research phase (1.5 months). In the literature survey phase you will investigate (a) the current state of the art of existing CSCW tools and 3D modeling environments that focus on (or can be used for) presence, and (b) scientific publications on factors that mediate the feeling of presence (of others). In the research part you will implement a prototype 3D environment in which you use those factors influencing presence that have the highest potential. If your prototype is good, you have the opportunity to test it in a real industry environment. You will write a report of your findings which you will have to present at the Telematica Instituut (Enschede).

Frank van Gemeren **Titel:** Extraction van salient points uit real-time video

Begeleider: Michael Lew

Omschrijving: Salient points zijn de "interessante" punten in beelden. Dit is merkbaar doordat er bijvoorbeeld een hoog contrast met de omliggende punten is, maw: veel variatie. Door verschillende transformaties zoals Wavelet Transform kunnen deze punten gevonden worden. Salient

points worden vooral gebruikt bij image recognition, omdat het berekenen van de features van alle pixels teveel rekentijd kost. Saliënt points werken redelijk onafhankelijk van de camera hoek. Het is vooral geschikt voor het herkennen van delen van een plaatje.

Het idee is om een GUI te hebben met links een video (pre-recorded of later real-time webcam) en rechts de uitwerking van de saliënt points (eventueel als overlay). De gebruiker kan een aantal parameters veranderen, wat dan invloed heeft op de saliënt points. Dit is nog niet besproken met Michael Lew, maar ik ga uit van de motion-tracking voorbeelden waarin hetzelfde concept wordt gebruikt.

Johan Groenen **Titel:** Nurikabe

Begeleider: Walter Kusters

Omschrijving: Nurikabe is een NP-complete, binaire determinatie puzzel van Nikoli — de geestelijk vader van de Sudoku. Initieel zijn op een $M \times N$ grid cijfers gedistribueerd, welke aangeven hoe groot het eiland is waarvan de betreffende genummerde cel deel uitmaakt. Voor meer informatie, zie <http://en.wikipedia.org/wiki/Nurikabe>.

Momenteel heb ik een oplosser geschreven die de zoekruimte verkleint door alleen de ruimte van mogelijke eilandvormen te doorzoeken. Volgende stappen zijn: het opstellen van — simpele en meer complexe — logische regels teneinde een logische oplosser te programmeren en het schrijven van een — snelle — generator die — uniform random — een van de mogelijke — gerelaxeerde, bijv. niet uniek oplosbaar — $M \times N$ puzzels genereert.

Hierna wil ik me verdiepen in de verschillende graden van moeilijkheid die al dan niet te definiëren zijn.

Saskia Hiltemann **Titel:** Multicoloured Cellular Automata

Begeleider: Michael Emmerich

Omschrijving: Het maken van een java-applet wat deel zal uitmaken van de educational suite. Applet dat cellular automata implementeert waarbij iedere cel meerdere kleuren (states) kan aannemen. Cellular automata simuleren de mechanismen die zorgen voor bijvoorbeeld strepen van tijgers, patronen op schelpen etc. Iedere cel kan per tijdstap in een andere state terecht komen afhankelijk van de cellen in zijn neighbourhood. Afhankelijk van de beginstate en de overgangsregels kunnen zeer verschillende (en complexe) patronen ontstaan. Voor dit bachelor project maak ik deze applet en bekijk ik (eventueel) het gedrag van een aantal specifieke cellular automata.

Erik Jongma **Titel:** Strand Spaces

Begeleider: Tom Chothia (CWI)

Omschrijving: The Strand Space method is a technique for analyzing cryptographic protocols (particularly authentication and key distribution protocols) at a very high level. The technique (like many others) uses the Dolev-Yao model of cryptographic protocols, and abstracts the real-world cryptographic algorithms into abstract operations. Hence, Strand Spaces focuses on the structure of a protocol, rather than the protocol's use of any particular encryption scheme. Originally designed as a pencil-and-paper proof method, it has also been automated. It has been extended in various ways, and its relationship with other proof methods has been explored. It also seems a handy framework with which to prove general results about protocols.

Joppe Kroon **Titel:** Multi Target Support in Eclipse

Begeleider: Bart Kienhuis

Omschrijving: Om een programma voor een multicore omgeving te kunnen compileren moet gecompileerd worden naar meerdere "targets", een voor elke processor, waarbij een 'object file' gemaakt wordt voor elke processor. Daarnaast is het soms ook het geval dat er voor verschillende processoren gecompileerd moet worden, waarvoor dus verschillende compilers gebruikt moeten worden.

De gewenste ontwikkelomgeving, Eclipse met CDT, schiet op deze punten nog te kort. CDT kan maar voor een enkele target compileren, en dat alleen naar zijn standaard compiler. Het project betreft dan ook het onderzoeken van CDT en het bekijken van de mogelijkheden om deze aan te passen om aan de nieuwe eisen te voldoen.

Susan Laraghy **Titel:** The Stipulation and Improvement of Existing Patterns of DNA

Begeleider: Hendrik Jan Hoogeboom

Omschrijving: Existing sequence patterns of DNA are called "motifs" and have a certain biological significance. These motifs may occur in various positions, whether it be within a sort of genome, or genomes of various sorts but similar genes. The frequency matrices built upon these occurrences are called "position specific scoring matrixes".

Exact details of the project are still unknown. A file of DNA sequences containing hidden motifs will be provided, and furthermore the position specific scoring matrixes that need to be used to find the sequences that resemble the motif sequence above a certain threshold. This in turn demands adjustment of the matrix based on the occurrences obtained.

Tim van Meurs **Titel:** De moeilijkheidsgraad van Kuromasu-puzzels

Begeleider: Walter Kusters

Omschrijving: Het doel van het bachelorproject is het maken van een algoritme dat in staat is om Kuromasu-puzzels op te lossen. Het algoritme zal meerdere methodes gebruiken om de puzzel op te lossen. Sommigen van deze methodes zullen van elkaar afhankelijk zijn. Hierbij worden

eerst de makkelijke methodes gebruikt, als deze niet tot een oplossing leiden wordt overgegaan op moeilijkere methodes, enz. Door achteraf te kijken welke methodes zijn gebruikt kan worden bepaald wat de moeilijkheidsgraad is van de puzzel. Dit wordt gedaan door te kijken naar de complexiteit van iedere methode en de toepasbaarheid van de methode bij het oplossen van de puzzel met pen en papier.

Kuromasu puzzels: (engels: Where is Black Cells)
<http://en.wikipedia.org/wiki/Kuromasu>.

Timo Morsink **Titel:** Project in de maak in combinatie met Grid computing of Google Earth

Begeleider: Lex Wolters

Omschrijving: Waarschijnlijk wordt het een project in samenwerking met de KNMI met Grid Computing als basis. Helaas was er nog niet bekend of er een project nu te regelen is daarbinnen en dat is Lex Wolters nu aan het uitzoeken. Het uiteindelijke project wordt dus pas volgende week besloten. Als er niet iets te vinden is binnen afzienbare tijd (een week) dan ga ik mij waarschijnlijk bezig houden met de sateliet foto's van het KNMI samen te laten werken met Google Earth. Als ik dit ga doen dan wordt het uitzoeken hoe dit het beste kan en waarschijnlijk laten we het programmeerwerk dan voor een software project, omdat het anders een veel te groot project wordt.

Edgar Reehuis **Titel:** SMS-EMOA aanpassen

Begeleider: Michael Emmerich

Omschrijving: LIACS heeft een van de beste algoritmes beschikbaar voor het oplossen van multi-objective optimalisatieproblemen onlangs (mede-)ontwikkeld: het SMS-EMOA. De uitdaging is mede- en tegenstanders in het veld hiervan te overtuigen: hoe beter dan aan de hand van een superieure oplossing voor een “real-life” probleem? Recent is het probleem van het vinden van een kosten- en nut-optimale oplossing voor water distributienetwerken uitgedrukt als een MoO-probleem en is hier een genetisch algoritme op los gelaten (zie Fuzzy hierarchical decision support system for water distribution network optimization: <http://www.informaworld.com/smpp/content~content=a749213315~db=all~order=page>).

De bedoeling is gebruikmakend van deze omschrijving het SMS-EMOA aan te passen aan het probleem en te laten zien dat de oplossing verkregen gebruikmakend van SMS-EMOA, de oplossingen van het alternatieve genetisch algoritme (Pareto) domineert.

Stefan Schrama —

Begeleider: Harmen van der Spek

Jens van de Water **Titel:** 3d database interfaces

Begeleider: Fons Verbeek

Omschrijving: Bij een bestaande medische image database een 3d omgeving creëren waarbij de gebruikers tegelijk in de database kunnen “rondlopen”, kunnen overleggen en informatie kunnen uitwisselen. Plaatjes die met in elkaar in verband staan zullen op innovatieve wijze met elkaar verbonden worden. Het onderzoek zal gericht worden op het collaboratieve aspect van 3d environments.

Stefan Wink **Titel:** Het gebruik van modellen in software projecten en de relatie tussen UML en de uiteindelijke code.

Begeleider: Michel Chaudron

Omschrijving: Bij het ontwikkelen van software worden er vaak vooraf modellen (Waterfall/Star, maar vooral UML) gemaakt. Ik wil onderzoeken hoe nuttig deze modellen zijn, hoeveel tijd moet worden besteed aan het maken van deze modellen en in welke mate ze gebruikt dienen te worden voor een optimaal resultaat. Ook wil ik kijken (in het geval van UML diagrammen) hoe deze modellen helpen bij het schrijven van de software (komen stukken uit het diagram letterlijk in de code terug?).

Sjaak Wolff **Titel:** 3d database interfaces

Begeleider: Fons Verbeek

Omschrijving: Bij een bestaande medische image database een 3d omgeving creëren waarbij de gebruikers tegelijk in de database kunnen “rondlopen”, kunnen overleggen en informatie kunnen uitwisselen. Plaatjes die met in elkaar in verband staan zullen op innovatieve wijze met elkaar verbonden worden. Het onderzoek zal gericht worden op het collaboratieve aspect van 3d environments.

Fee Yun Wong —

Begeleider: Luuk Groenewegen

Luuk Groenewegen en Walter Kusters