NEWSLETTER

BNVKI AIABN
October 2010
Vol. 27, No. 5
ISSN 1566-8266

Ramanujaniana (part 1)

The Magic of Monte-Carlo Tree Search

The 2010 SKBS-Strukton Prize

News from the Benelux Association for Artificial Intelligence

1729

Editor-in-chief

Srinivasa Iyengar Ramanujan (December 22, 1887 – April 26, 1920), or Ramanujan for short, was an Indian mathematical genius who, with almost no formal training in pure mathematics, made substantial contributions to mathematical analysis, number theory, infinite series and continued fractions. Ramanujan's talent was said, by the prominent English mathematician G.H. Hardy, to be in the same league as legendary mathematicians such as Euler, Gauss, Newton and Archimedes. Ramanujan was fond of infinite series. In particular, he discovered an expression for π, enabling the quick calculation of successive decimals of π. By this formula, π has at the end of the 1980s been calculated up to over a billion digits.

$$\frac{1}{\pi} = \frac{\sqrt{8}}{9801} \sum_{n=0}^{\infty} \frac{(4n)!}{(n!)^4} \frac{[1103 + 26390n]}{(396)^{4n}}$$

Although he died at the very young age of only 32, his achievements are overwhelming. As a consequence, his name is appearing in many fields in mathematics. For example, there are prime numbers obeying some specified constraints, now known as *Ramajunan primes*. There exists a generalization of the Jacobi theta function, now called the *Ramanujan theta function*. There is a famous conjecture, of course called the *Ramanujan conjecture*, relating Fourier coefficients and prime numbers. His name even can be found in relation with the golden ratio ($\varphi = \frac{1+\sqrt{5}}{2}$), for which Ramanujan found the following beautiful expression:

$$\phi = \frac{\sqrt{5}}{1 + \sqrt[5]{5^{3/4}(1/\phi)^{5/2} - 1}} - \cfrac{e^{-2\pi/\sqrt{5}}}{1 + \cfrac{e^{-2\pi\sqrt{5}}}{1 + \cfrac{e^{-4\pi\sqrt{5}}}{1 + \cfrac{e^{-6\pi\sqrt{5}}}{1 + \ldots}}}}$$

Also, there is a *Ramanujan Prize*, honouring a researcher, younger than 45 years old, who has conducted outstanding mathematics research in a developing country. There even is a mathematical journal, entitled *The Ramanujan Journal*, which is completely devoted to areas of mathematics influenced by his work.

A common anecdote about Ramanujan relates to the number 1729. Hardy arrived at Ramanujan's residence in a cab numbered 1729. Hardy commented that the number 1729 seemed to be uninteresting. Ramanujan is said to have stated on the spot that it was actually a very interesting number mathematically, being the smallest natural number representable in two different ways as a sum of two cubes: $1729 = 1^3 + 12^3 = 9^3 + 10^3$. Generalizations of this idea have spawned the notion of "taxicab numbers". A very interesting article on another generalization is written by Henk Visser. The first part of this is reproduced on pp. 108-115 of this issue (the second part will appear in the December issue of the newsletter).

And now for something completely different. Although the BNAIC 2010 is finished meanwhile, we defer reports on it to the December issue (for space reasons), except for a report on the award of the SKBS-Strukton prize and the minutes of the General Assembly. In this meeting it was announced that, in order to reduce costs, the newsletter will undergo some changes, to be effectuated in 2011. Probably the most important one is that the frequency of the newsletter will go back from six to four issues a year. To oppose the inherent drawback that announcements of BNVKI or other relevant events will be untimely, an accompanying up-to-date website will be set up. Further, we also plan to reform the contents of the newsletter. For this, we are very interested in your opinion. What are the sections you (don't) like? What are the types of contributions you miss? What other suggestions for changes do you have?

Please send all suggestions to your editor-in-chief (uiterwijk@maastrichtuniversity.nl). They will be discussed in the board and hopefully will lead to an improved newsletter, which is *your* newsletter after all. I'm looking forward! For the 1729th serious suggestion, an appropriate gift will be arranged.

Wiki's page on Ramanujan:	**http://en.wikipedia.org/wiki/Srinivasa_Ramanujan**
Henk Visser's home page:	**http://www.henkxvisser.nl/**

TABLE OF CONTENTS

Front cover: image of Srinivasa Iyengar Ramanujan, taken around 1900, by an unidentified author.

The deadline for the next issue is: **December 1, 2010**.

BNVKI-Board News

Antal van den Bosch

I think it is safe to say that BNAIC has reached a plateau: the Kirchberg Plateau, that is. BNAIC-2010 was held in the amazingly futuristic Kirchberg quarter of Luxembourg City, where the European Union has some of its major bodies sit in high rising buildings, and where the Luxembourg banking world has decided to cluster. Moreover, in the Musée d'Art Moderne Grand-Duc Jean (Mudam Luxembourg) you can find some pretty alluring and though provoking modern art, and we were fortunate to see the dark black cube of the Melia Hotel on the inside when we had the conference dinner.

BNAIC-2010 was great – it had all the ingredients, two wonderful invited speakers, a great surroundings, and a very lively atmosphere. The board wishes to heartily thank the organizers, a group of people from various departments of the University of Luxembourg, headed by Leon van der Torre, Pascal Bouvry, Eric Dubois, and Thibaud Latour, helped by an active troupe of helping hands. And as much as we loved being in the smallest of our three Benelux countries, we have our heads turned now to BNAIC-2011, which in the General Assembly was determined to be Ghent, Belgium.

In the General Assembly, the board also said good-bye to board members Sien Moens and Richard Booth. Sien and Richard, thank you both for donating time and effort to our association; Sien for the past five years, and Richard for an intense year in which he was the liaison between the board and the BNAIC-2010 organization. Thank you, both.

Ramanujaniana
Part 1

Henk Visser
Haarlem

(*It is the last working day of the year. Math, Comp, and Log are drinking coffee in the canteen of the L.E.J. Brouwer Institute, when Comp's attention is drawn by an interview in a newspaper in which the well-known anecdote of Ramanujan is mentioned. Ramanujan called the number of Hardy's taxicab, 1729, extraordinary, because it is the smallest number that can be twice written as the sum of two cubic numbers.*)

COMP. You are also a number freak, Math. Do you also have such interesting insights?

MATH. I'm not so strongly impressed about Ramanujan's remark to Hardy as most people. After all, someone who knows the first fifteen cubic numbers by heart, as Ramanujan undoubtedly did, can't have failed to see that 1000 + 729 is just 1 more than 1728. Nevertheless, I asked myself if I could invent a similar, but less easily solvable problem. My principle of variation helped me, as you understand. To present it in such a way that the variation comes out clearly: instead of two sums of powers of degree three, I took three sums of powers of degree two.

COMP. So you asked for the smallest number that can be written three times as the sum of two squares?

MATH. Right, and because I know the first forty squares by heart, I trusted that I could solve this problem.

LOG. But how could you be sure that there are any solutions at all?

COMP. What? That's not a problem to me; I know the first four irreducible Pythagorean triples by heart, and already the first three give a solution.

LOG. Show it!

COMP. We have $3^2 + 4^2 = 5^2$, $5^2 + 12^2 = 13^2$, and $7^2 + 24^2 = 25^2$, and the least common multiple of 5, 13, and 25 gives a solution: 325^2. I will write it down for you (*he takes a paper napkin*):

$$(5\cdot13\cdot\underline{3})^2 + (5\cdot13\cdot\underline{4})^2 = (5\cdot13\cdot\underline{5})^2$$
$$(5\cdot5\cdot\underline{5})^2 + (5\cdot5\cdot\underline{12})^2 = (5\cdot5\cdot\underline{13})^2$$
$$(13\cdot\underline{7})^2 + (13\cdot\underline{24})^2 = (13\cdot\underline{25})^2$$

LOG. I am convinced!

MATH. 325^2 is surely not the smallest number with the desired property, but nevertheless it is very special. Comp, I thank you very much for this outcome, because the smallest number that is three times the sum of two squares is, believe it or not, 325!

LOG. I don't believe in mathematical magic, but I admit that it is remarkable!

COMP. Please give the three sums, Math.

MATH. 325 = 324 + 1, 325 = 289 + 36, and 325 = 225 + 100.

LOG. How did you find it?

MATH. The first sum I tried was 365, just for fun, because the year is almost over, and I saw or knew that it is both the sum of 169 and 196, and the sum of 361 and 4, but there was no more to it. Then I tried 325, because it is 324 + 1, similar to Ramanujan's 1728 + 1, and I found the three sums. Thereafter I tried numbers smaller than 325, but this yielded at most only numbers with two sums.

COMP. I will not underestimate your arithmetical capacities, Math, but still I will check your conclusion on my computer if you don't mind …

MATH. Good to hear that you keep working in the vacation period! I also have problems that bother me, so have a good time and see you next year!

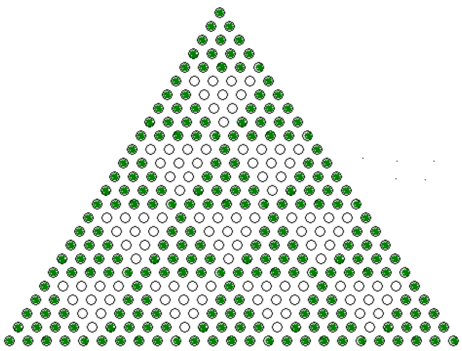
LOG. Are you perhaps trying to invent more Ramanujanean problems, Math, or is there still more to tell about 325?

MATH. Yes and yes, and I can already answer your second question: 325 is the triangular of a square number, so it can be written as the sum of the squares of two successive triangular numbers, just as the famous or notorious number 666.

LOG. I see (*she writes on the napkin*):

$$325 = 15^2 + 10^2, \text{ or } t(s(5)) = s(t(5)) + s(t(4))$$

MATH. This can be represented in a perspicuous figure, as you both know. It becomes my season's greetings picture (*he takes a card out of his pocket*):



Best wishes and a happy new year!

(*They all go their own way.*)

(*When Math is still in his room, trying to bring some order in his papers before leaving the Institute, Comp enters.*)

COMP. I checked your solution, and you were right, 325 is the smallest number with the desired property! But I did more, I looked for the smallest number that is three times the sum of three squares. It appeared to be 101, because 101 = 1 + 36 + 64, 4 + 16 + 81, and 16 + 36 + 49. However, there are so many numbers which are the sum of three squares, 110, 126, 134, 146, and so on, that I wondered what would be the use of this exercise.

MATH. Would it perhaps be better to follow my principle of variation, and look for the smallest number that is four times the sum of three squares?

COMP. I will do my best, please stay in your room for a while! (*He leaves, and returns half an hour later.*)

COMP. The smallest number that is four times the sum of three squares is 161, and the smallest number that is five times the sum of four squares is 126. I don't think it's worthwhile to pursue this further. Therefore I wrote a program in which the principle of generalization is applied to your original problem: to find the smallest numbers that are *n* times the sum of two squares. However, it isn't guaranteed that the outcomes will be optimal. I hereby give you the results in the form I got them. I'm not sure about 11, 17, 19, and 21. Therefore I omitted them. The other missing values gave as yet no solutions. If there are such, they are too large for my program. Moreover it got stuck after 32.

***n* = 1**
5 = (s(1)+s(2))

***n* = 2**
65 = (s(1)+s(8)) = (s(4)+s(7))

***n* = 3**
325 = (s(1)+s(18)) = (s(6)+s(17)) = (s(10)+s(15))

***n* = 4**
1105 = (s(4)+s(33)) = (s(9)+s(32)) = (s(12)+s(31)) = (s(23)+s(24))

***n* = 5**
8125 = (s(5)+s(90)) = (s(27)+s(86)) = (s(30)+s(85)) = (s(50)+s(75)) = (s(58)+s(69))

***n* = 6**
5525 = (s(7)+s(74)) = (s(14)+s(73)) = (s(22)+s(71)) = (s(25)+s(70)) = (s(41)+s(62)) = (s(50)+s(55))

***n* = 7**
105625 = (s(36)+s(323)) = (s(80)+s(315)) = (s(91)+s(312)) = (s(125)+s(300)) = (s(165)+s(280)) = (s(195)+s(260)) = (s(204)+s(253))

***n* = 8**
27625 = (s(20)+s(165)) = (s(27)+s(164)) =
(s(45)+s(160)) = (s(60)+s(155)) = (s(83)+s(144)) =
(s(88)+s(141)) = (s(101)+s(132)) = (s(115)+s(120))

***n* = 9**
71825 = (s(1)+s(268)) = (s(40)+s(265)) =
(s(65)+s(260)) = (s(76)+s(257)) = (s(104)+s(247)) =
(s(127)+s(236)) = (s(160)+s(215)) =
(s(169)+s(208)) = (s(188)+s(191))

***n* = 10**
138125 = (s(22)+s(371)) = (s(35)+s(370)) =
(s(70)+s(365)) = (s(110)+s(355)) = (s(125)+s(350)) =
(s(163)+s(334)) = (s(194)+s(317)) =
(s(205)+s(310)) = (s(218)+s(301)) =
(s(250)+s(275))

***n* = 12**
160225 = (s(15)+s(400)) = (s(32)+s(399)) =
(s(76)+s(393)) = (s(81)+s(392)) = (s(113)+s(384)) =
(s(140)+s(375)) = (s(175)+s(360)) =
(s(183)+s(356)) = (s(216)+s(337)) =
(s(228)+s(329)) = (s(252)+s(311)) =
(s(265)+s(300))

***n* = 13**
1221025 = (s(47)+s(1104)) = (s(105)+s(1100)) =
(s(169)+s(1092)) = (s(264)+s(1073)) =
(s(272)+s(1071)) = (s(425)+s(1020)) =
(s(468)+s(1001)) = (s(520)+s(975)) =
(s(561)+s(952)) = (s(576)+s(943)) =
(s(663)+s(884)) = (s(700)+s(855)) =
(s(744)+s(817))

***n* = 14**
3453125 = (s(31)+s(1858)) = (s(110)+s(1855)) =
(s(175)+s(1850)) = (s(350)+s(1825)) =
(s(550)+s(1775)) = (s(625)+s(1750)) =
(s(686)+s(1727)) = (s(815)+s(1670)) =
(s(847)+s(1654)) = (s(970)+s(1585)) =
(s(1025)+s(1550)) = (s(1090)+s(1505)) =
(s(1201)+s(1418)) = (s(1250)+s(1375))

***n* = 15**
1795625 = (s(5)+s(1340)) = (s(52)+s(1339)) =
(s(179)+s(1328)) = (s(200)+s(1325)) =
(s(325)+s(1300)) = (s(380)+s(1285)) =
(s(467)+s(1256)) = (s(520)+s(1235)) =
(s(563)+s(1216)) = (s(635)+s(1180)) =
(s(676)+s(1157)) = (s(800)+s(1075)) =
(s(808)+s(1069)) = (s(845)+s(1040)) =
(s(940)+s(955))

***n* = 16**
801125 = (s(10)+s(895)) = (s(95)+s(890)) =
(s(127)+s(886)) = (s(158)+s(881)) =
(s(193)+s(874)) = (s(230)+s(865)) =
(s(241)+s(862)) = (s(335)+s(830)) =
(s(370)+s(815)) = (s(430)+s(785)) =
(s(458)+s(769)) = (s(463)+s(766)) =
(s(529)+s(722)) = (s(545)+s(710)) =
(s(554)+s(703)) = (s(610)+s(655))

***n* = 18**
2082925 = (s(26)+s(1443)) = (s(134)+s(1437)) =
(s(163)+s(1434)) = (s(195)+s(1430)) =
(s(330)+s(1405)) = (s(370)+s(1395)) =
(s(429)+s(1378)) = (s(531)+s(1342)) =
(s(541)+s(1338)) = (s(558)+s(1331)) =
(s(579)+s(1322)) = (s(702)+s(1261)) =
(s(730)+s(1245)) = (s(755)+s(1230)) =
(s(845)+s(1170)) = (s(894)+s(1133)) =
(s(926)+s(1107)) = (s(1014)+s(1027))

***n* = 20**
4005625 = (s(75)+s(2000)) = (s(147)+s(1996)) =
(s(160)+s(1995)) = (s(336)+s(1973)) =
(s(380)+s(1965)) = (s(405)+s(1960)) =
(s(488)+s(1941)) = (s(565)+s(1920)) =
(s(632)+s(1899)) = (s(700)+s(1875)) =
(s(852)+s(1811)) = (s(875)+s(1800)) =
(s(915)+s(1780)) = (s(1069)+s(1692)) =
(s(1080)+s(1685)) = (s(1140)+s(1645)) =
(s(1197)+ s(1604)) = (s(1260)+s(1555)) =
(s(1325)+s(1500)) = (s(1344)+s(1483))

***n* = 22**
30525625 = (s(235)+s(5520)) =
(s(525)+s(5500)) = (s(612)+s(5491)) =
(s(845)+s(5460)) = (s(1036)+s(5427)) =
(s(1131)+s(5408)) = (s(1320)+s(5365)) =
(s(1360)+s(5355)) = (s(1547)+s(5304)) =
(s(2044)+s(5133)) = (s(2125)+s(5100)) =
(s(2163)+s(5084)) = (s(2340)+s(5005)) =
(s(2600)+s(4875)) = (s(2805)+s(4760)) =
(s(2880)+s(4715)) = (s(3124)+s(4557)) =
(s(3315)+s(4420)) = (s(3468)+s(4301)) =
(s(3500)+s(4275)) = (s(3720)+s(4085)) =
(s(3861)+s(3952))

***n* = 24**
5928325 = (s(63)+s(2434)) = (s(94)+s(2433)) =
(s(207)+s(2426)) = (s(294)+s(2417)) =
(s(310)+s(2415)) = (s(465)+s(2390)) =
(s(490)+s(2385)) = (s(591)+s(2362)) =
(s(690)+s(2335)) = (s(742)+s(2319)) =
(s(849)+s(2282)) = (s(878)+s(2271)) =
(s(959)+s(2238)) = (s(1039)+s(2202)) =
(s(1062)+s(2191)) = (s(1201)+s(2118)) =
(s(1215)+s(2110)) = (s(1290)+s(2065)) =
(s(1410)+s(1985)) = (s(1454)+s(1953)) =
(s(1535)+s(1890)) = (s(1614)+s(1823)) =
(s(1633)+s(1806)) = (s(1697)+s(1746))

***n* = 27**
35409725 = (s(85)+s(5950)) = (s(338)+s(5941)) = (s(650)+s(5915)) = (s(782)+s(5899)) = (s(826)+s(5893)) = (s(901)+s(5882)) = (s(994)+s(5867)) = (s(1339)+s(5798)) = (s(1547)+s(5746)) = (s(1675)+s(5710)) = (s(1790)+s(5675)) = (s(1973)+s(5614)) = (s(2086)+s(5573)) = (s(2210)+s(5525)) = (s(2443)+s(5426)) = (s(2597)+s(5354)) = (s(2725)+s(5290)) = (s(2875)+s(5210)) = (s(3029)+s(5122)) = (s(3094)+s(5083)) = (s(3466)+s(4837)) = (s(3502)+s(4811)) = (s(3563)+s(4766)) = (s(3638)+s(4709)) = (s(3835)+s(4550)) = (s(4069)+s(4342)) = (s(4165)+s(4250))

n = 32
29641625 = (s(67)+s(5444)) = (s(124)+s(5443)) = (s(284)+s(5437)) = (s(320)+s(5435)) = (s(515)+s(5420)) = (s(584)+s(5413)) = (s(835)+s(5380)) = (s(955)+s(5360)) = (s(1180)+s(5315)) = (s(1405)+s(5260)) = (s(1460)+s(5245)) = (s(1648)+s(5189)) = (s(1795)+s(5140)) = (s(1829)+s(5128)) = (s(1979)+s(5072)) = (s(2012)+s(5059)) = (s(2032)+s(5051)) = (s(2245)+s(4960)) = (s(2308)+s(4931)) = (s(2452)+s(4861)) = (s(2560)+s(4805)) = (s(2621)+s(4772)) = (s(2840)+s(4645)) = (s(3005)+s(4540)) = (s(3035)+s(4520)) = (s(3320)+s(4315)) = (s(3365)+s(4280)) = (s(3517)+s(4156)) = (s(3544)+s(4133)) = (s(3664)+s(4027)) = (s(3715)+s(3980)) = (s(3803)+s(3896))

MATH. Marvellous! And now it becomes interesting: can we discover a regularity? The first ten outcomes suggest that we can divide all numbers by 13, or, better, by 65. Moreover, do we have to distinguish between even and odd numbers? This becomes my task for the vacation period! Exciting!

COMP. Have a good time, Math! (*He leaves the room with a shrug of the shoulders.*)

LOG. (*Enters the room a few minutes later.*) Still working, Math?

MATH. Log gave me a tremendous amount of computations, and I am thinking them over.

LOG. Can I see them?

MATH. Here are the first ten items of Comp's list with the smallest numbers that are *n* times the sum of two squares. My outcome 325 is on this list:

***n* = 1**
5 = (s(1)+s(2))

***n* = 2**
65 = (s(1)+s(8)) = (s(4)+s(7))

***n* = 3**
325 = (s(1)+s(18)) = (s(6)+s(17)) = (s(10)+s(15))

***n* = 4**
1105 = (s(4)+s(33)) = (s(9)+s(32)) = (s(12)+s(31)) = (s(23)+s(24))

***n* = 5**
8125 = (s(5)+s(90)) = (s(27)+s(86)) = s(30)+s(85)) = (s(50)+s(75)) = (s(58)+s(69))

***n* = 6**
5525 = (s(7)+s(74)) = (s(14)+s(73)) = (s(22)+s(71)) = (s(25)+s(70)) = (s(41)+s(62)) = (s(50)+s(55))

***n* = 7**
105625 = (s(36)+s(323)) = (s(80)+s(315)) = (s(91)+s(312)) = (s(125)+s(300)) = (s(165)+s(280)) = (s(195)+s(260)) = (s(204)+s(253))

***n* = 8**
27625 = (s(20)+s(165)) = (s(27)+s(164)) = (s(45)+s(160)) = (s(60)+s(155)) = (s(83)+s(144)) = (s(88)+s(141)) = (s(101)+s(132)) = (s(115)+s(120))

***n* = 9**
71825 = (s(1)+s(268)) = (s(40)+s(265)) = (s(65)+s(260)) = (s(76)+s(257)) = (s(104)+s(247)) = (s(127)+s(236)) = (s(160)+s(215)) = (s(169)+s(208)) = (s(188)+s(191))

***n* = 10**
138125 = (s(22)+s(371)) = (s(35)+s(370)) = (s(70)+s(365)) = (s(110)+s(355)) = (s(125)+s(350)) = (s(163)+s(334)) = (s(194)+s(317)) = (s(205)+s(310)) = (s(218)+s(301)) = (s(250)+s(275))

LOG. It's curious that there are no doubles. In fact, the smallest number that is the sum of two squares is not 5, but 2. For *n* = 2, I see immediately 50, because 50 = 49 + 1, *and* 50 = 25 + 25. Comp's solution gives the smallest numbers that are *n* times the sum of two *different* squares! I hope that it doesn't affect your solution for *n* = 3, Math!

MATH. (*After some scribbling on the blackboard.*) No, that's all right. Nevertheless, I will stay with Comp's outcomes, because they seem to have some nice properties. I factorized all reliable numbers and this resulted in the following table:

***n* = 1**
5 = 5

***n* = 2**
65 = 5·13

***n* = 3**
325 = 5^2·13

***n* = 4**
1105 = 5·13·17

***n* = 5**
8125 = 5^4·13

***n* = 6**
5525 = 5^2·13·17

***n* = 7**
105625 = $5^4·13^2$

***n* = 8**
27625 = 5^3·13·17

***n* = 9**
71825 = $5^2·13^2$·17

***n* = 10**
138125 = 5^4·13·17

***n* = 12**
160225 = 5^2·13·17·29

***n* = 13**
1221025 = $5^2·13^2·17^2$

***n* = 14**
3453125 = 5^6·13·17

***n* = 15**
1795625 = $5^4·13^2$·17

***n* = 16**
801125 = 5^3.13.17.29

***n* = 18**
2082925 = $5^2·13^2$·17·29

***n* = 20**
4005625 = 5^4·13·17·29

***n* = 22**
30525625 = $5^4·13^2·17^2$

***n* = 24**
5928325 = 5^2·13·17·29·37

MATH. Wait, an email message from Comp. (*Math reads it.*)

LOG. What does Comp write?

MATH. He found more numbers, which he already factorized. I'll write them down, together with the values that we already have:

***n* = 27**
$5^2·13^2·17^2$·29

***n* = 28**
5^6·13·17·29

***n* = 30**
$5^4·13^2$·17·29

***n* = 32**
5^3·13·17·29·37

***n* = 36**
$5^2·13^2$·17·29·37

***n* = 40**
5^4·13·17·29·37

***n* = 45**
$5^4·13^2·17^2$·29

***n* = 48**
5^2·13·17·29·37·41

And, last but not least, surprisingly:

***n* = 11**
5^{10}·13

LOG. The powers of 5 are never less than those of the other prime numbers of the form 4*k* + 1. Not all combinations occur. Do you have an explanation, Math?

MATH. I presume that the combinations which are skipped, belong to numbers with one of the required properties, without being minimal of course.

LOG. Let us try $5^2·13^2$. How do we find the desired representations?

MATH. Believe it or not, the answer can be found on Internet, by what is called the computational knowledge machine Wolframalpha. (*Math goes to the site and he finds the answer.*) Look:

4225 has 4 representations as a sum of 2 squares:

$$4225 = 16^2 + 63^2 = 25^2 + 60^2 = 33^2 + 56^2 = 39^2 + 52^2$$

LOG. That's nice, and now $5^3 \cdot 13$.

1625 has 4 representations as a sum of 2 squares:

$$1625 = 5^2 + 40^2 = 16^2 + 37^2 = 20^2 + 35^2 = 28^2 + 29^2$$

MATH. Do you need more examples?

LOG. Perhaps $5^3 \cdot 13^2$ and $5^3 \cdot 13^3$?

MATH. (*He fills in 21125 and 274625.*)

21125 has 6 representations as a sum of 2 squares:

$$21\,125 = 10^2 + 145^2 = 26^2 + 143^2 = 31^2 + 142^2 = 65^2 + 130^2 = 79^2 + 122^2 = 95^2 + 110^2$$

274625 has 8 representations as a sum of 2 squares:

$$274\,625 = 7^2 + 524^2 = 65^2 + 520^2 = 140^2 + 505^2 = 191^2 + 488^2 = 208^2 + 481^2 = 260^2 + 455^2 = 320^2 + 415^2 = 364^2 + 377^2$$

LOG. Again a message from Comp. (*She opens it.*) Interesting! Comp had a brilliant insight! He found regularities by ordering the numbers in a special way, first 5, 5·17, 5·17·29, 5·17·29·37, 5·17·29·37·41, then 5^2, $5^2 \cdot 17$, $5^2 \cdot 17 \cdot 29$, $5^2 \cdot 17 \cdot 29 \cdot 37$, and so on, each time taking the next power of 5 and omitting the factor 13. He encoded the prime factors by their exponents (with the numbers behind the minus signs giving the number of representations as sums of two different squares), and got the following results (*She prints them out.*)

5: 1 - 1
85: 101 - 2
2465: 1011 - 4
91205: 10111 - 8
3739405: 101111 - 16

25: 2 - 1
425: 201 - 3
12325: 2011 - 6
456025: 20111 - 12
18697025: 201111 - 24

125: 3 - 2
2125: 301 - 4
61625: 3011 - 8
2280125: 30111 - 16
93485125: 301111 - 32

625: 4 - 2
10625: 401 - 5
308125: 4011 - 10
11400625: 40111 - 20
467425625: 401111 - 40

3125: 5 - 3
53125: 501 - 6
1540625: 5011 - 12
57003125: 50111 - 24

15625: 6 - 3
265625: 601 - 7
7703125: 6011 - 14
285015625: 60111 - 28

MATH. It seems that the bare powers of 5 must be omitted in order to get complete regularities.

85: 101 - 2
2465: 1011 - 4
91205: 10111 - 8
3739405: 101111 - 16

425: 201 - 3
12325: 2011 - 6
456025: 20111 - 12
18697025: 201111 - 24

2125: 301 - 4
61625: 3011 - 8
2280125: 30111 - 16
93485125: 301111 - 32

10625: 401 - 5
308125: 4011 - 10
11400625: 40111 - 20
467425625: 401111 - 40

53125: 501 - 6
1540625: 5011 - 12
57003125: 50111 - 24

265625: 601 - 7
7703125: 6011 - 14
285015625: 60111 - 28

LOG. Comp missed two outcomes, namely 57003125·41 and 285015625·41. This seems to provide an opportunity for checking the rule that subjoining the next prime factor of the form $4k + 1$ leads to a doubling of the number of sums.

MATH. (*He does a little computing and consults Wolframalpha.*)

57003125·41 = 2337128125

2337128125 has 48 representations as a sum of 2 squares:

Promising!

285015625·41 = 11685640625

(*???*)

That's a pity, Wolfram gives no representations, although it mentions that 11685640625 is the hypotenuse of 16 primitive Pythagorean triples, but that's another topic. Perhaps the predicted number of 56 representations is too large for the computational knowledge machine? Still, the 48 representations of $5^5·17·29·37·41$ seem to give sufficient evidence for trusting the doubling rule. We must thank Comp![1] (*He writes a reply to Comp in which he congratulates him with what he achieved.*)

LOG. Nevertheless, I want a proof!

MATH. I will think it over ... and we have not even solved our original problem of finding regularities in the *minimal* numbers ... Perhaps a comparison between their table and Comp's last one will help.

LOG. It strikes me that the minimal numbers all have the factor 13, whereas this factor is completely absent in the regularities.

MATH. Comp arranged the numbers according to the powers of the factor 5 and the increasing series of factors of the form $4k + 1$. It seems that we must do the same with the minimals, to begin with. First the exponent 2, starting with the value for $n = 3$, my famous 325. I will write it as $m(3) = 325$. Then we get:

$m(3) = 5^2·13$
$m(6) = 5^2·13·17$
$m(12) = 5^2·13·17·29$
$m(24) = 5^2·13·17·29·37$
$m(48) = 5^2·13·17·29·37·41$

LOG. I understand how it continues with 5^3:

$m(8) = 5^3·13·17$
$m(16) = 5^3·13·17·29$
$m(32) = 5^3·13·17·29·37$

MATH. Isn't it simple? I continue:

$m(5) = 5^4·13$
$m(10) = 5^4·13·17$
$m(20) = 5^4·13·17·29$
$m(40) = 5^4·13·17·29·37$

LOG. Yes, and finally:

$m(14) = 5^6·13·17$
$m(28) = 5^6·13·17·29$

MATH. Let's now try the cases in which the factor 13 has the exponent 2, to begin with $n = 9$:

$m(9) = 5^2·13^2·17$
$m(18) = 5^2·13^2·17·29$
$m(36) = 5^2·13^2·17·29·37$

LOG. Then we also have:

$m(15) = 5^4·13^2·17$
$m(30) = 5^4·13^2·17·29$

MATH. So far, so good, but we skipped some values, not only 1, 2, and 4, but also 7, 13, and 22, not to mention 11. We can investigate what happens with them, consulting our friend Wolfram! The values of 1, 2, and 4 are 5, 5·13, and 5·13·17. Let's see what happens with 5·13·17·29, that is, 32045.

32045 has 8 representations as a sum of 2 squares:
$32\,045 = 2^2 + 179^2 = 19^2 + 178^2 = 46^2 + 173^2 = 67^2 + 166^2 =$
$74^2 + 163^2 = 86^2 + 157^2 = 109^2 + 142^2 = 122^2 + 131^2$

LOG. That's OK, and wholly in the spirit of Comp! But 32045 isn't minimal, that was 27625. It's confirmed by Wolfram:

27625 is the smallest number with 8 representations as a sum of 2 squares:
$27\,625 = 20^2 + 165^2 = 27^2 + 164^2 = 45^2 + 160^2 = 60^2 + 155^2 =$
$83^2 + 144^2 = 88^2 + 141^2 = 101^2 + 132^2 = 115^2 + 120^2$

MATH. $m(7) = 5^4·13^2$, but I don't know what to do with it. The next values are already given away, $5^4·13^2·17 = m(15)$, and $5^4·13^2·17^2 = m(22)$. It's true that we can go back from $m(14) = 5^6·13·17$ and $m(28) = 5^6·13·17·29$, to get $5^6·13$, and verify that it has 7 representations as a sum of 2 squares, but $5^6·13$ is clearly larger than $5^4·13^2$.

LOG. We know that $5^{10}·13$ has 11 representations, but is it minimal? Moreover, the relation with $m(22) = 5^4·13^2·17^2$ isn't clear. I'm afraid that we do not get any further. So far the results are impressive, but they leave much to be desired. And now I need a drink.

MATH. Wait a minute! The beautiful value for 11 gives me the idea that we should investigate all preceding numbers of the form $5^k·13$ with the help of Wolfram, in so far as we don't already have their number of representations . I use the letter *n* in stead of *m*, because I expect that some values are not minimal.

[1] All credit to Dr. Jeroen Donkers, who not only wrote the computer programs, but also discovered the surprising theorem.

(*By consulting Wolfram when necessary, Math comes to the following table.*)

$n(5^0 \cdot 13) = 1$
$n(5^1 \cdot 13) = 2$
$n(5^2 \cdot 13) = 3$
$n(5^3 \cdot 13) = 4$
$n(5^4 \cdot 13) = 5$
$n(5^5 \cdot 13) = 6$
$n(5^6 \cdot 13) = 7$
$n(5^7 \cdot 13) = 8$
$n(5^8 \cdot 13) = 9$
$n(5^9 \cdot 13) = 10$
$n(5^{10} \cdot 13) = 11$

MATH. Now we see how Comp's solution for eleven representations is generated!

LOG. Amazing! It gives every number of representations! Maybe we can try to prove the general theorem, before going to the other regularities?

MATH. Seems plausible, but let's first take a drop.

(*He opens a cupboard and makes preparations for a drink.*)

LOG. If we stay in your room, we can investigate the $n(5^k \cdot 13^2)$'s, the $n(5^k \cdot 13^3)$'s, the $n(5^k \cdot 13^4)$'s, and see how far we get on with them.

MATH. (*filling the glasses*) Go ahead!

LOG. (*a couple of minutes later, helped by Wolfram*) Here you are:

$n(5^0 \cdot 13^2) = 1$
$n(5^1 \cdot 13^2) = 3$

$n(5^2 \cdot 13^2) = 4$
$n(5^3 \cdot 13^2) = 6$

$n(5^4 \cdot 13^2) = 7$
$n(5^5 \cdot 13^2) = 9$

$n(5^6 \cdot 13^2) = 10$
$n(5^7 \cdot 13^2) = 12$

$n(5^8 \cdot 13^2) = 13$

$n(5^0 \cdot 13^3) = 2$
$n(5^1 \cdot 13^3) = 4$
$n(5^2 \cdot 13^3) = 6$
$n(5^3 \cdot 13^3) = 8$
$n(5^4 \cdot 13^3) = 10$
$n(5^5 \cdot 13^3) = 12$

$n(5^0 \cdot 13^4) = 2$

$n(5^1 \cdot 13^4) = 5$
$n(5^2 \cdot 13^4) = 7$

$n(5^3 \cdot 13^4) = 10$
$n(5^4 \cdot 13^4) = 12$

$n(5^5 \cdot 13^4) = 15$
$n(5^6 \cdot 13^4) = 17$

MATH. 17 representations in seven steps, incredible! This deserves a toast!

(*After quite a number of drinks, they end the session in high spirits, and go home.)*

[To be continued]

The Magic of Monte-Carlo Tree Search

September 29, 2010

Mark Winands
DKE, Maastricht University

On September 29, in collaboration with SIKS and the BNVKI, the Department of Knowledge Engineering (DKE) at Maastricht University organized a symposium called "The Magic of Monte-Carlo Tree Search". The symposium was held in honour of Guillaume Chaslot's Ph.D. defence. The topic of his thesis (and therefore of the symposium as well) was Monte-Carlo Tree Search (MCTS). Approximately 20 participants attended the five talks. An outline of the five talks is given below.

YNGVI BJÖRNSSON (REYKJAVIK UNIVERSITY)
Yngvi Björnsson gave an inspiring talk about effectively using MCTS in General Game-Playing (GGP) programs. The aim of GGP is to create intelligent agents that can automatically learn how to play many different games at an expert level without any human intervention. One of the main challenges such agents face is to automatically learn knowledge-based heuristics in real-time, whether for evaluating game positions or for search guidance. In recent years, GGP agents that use MCTS simulations to reason about their actions have become increasingly more popular. One nice property of GGP as a testbed for research purposes is that one can use such game-playing programs to

study and contrast the effectiveness of MCTS on many different types of games. Also of an interest is that Go and GGP game playing programs, although both prominent examples of an effective use of MCTS, exhibit widely different characteristics. In computer Go, before coming up with a move decision, programs run a high number of simulations (in the hundreds of thousands) that are controlled using pre-defined and (typically) hand-crafted search-control knowledge. Conversely, in GGP, where state-space manipulation is more computing intensive, move decision must be based on much fewer simulations (few thousands), as well as the required search-control knowledge must be learned online during playing games.

PIM NIJSSEN (MAASTRICHT UNIVERSITY)

MCTS is becoming increasingly popular for playing multi-player games. In his talk, Pim Nijssen proposed two enhancements for MCTS in multi-player games. They are called Progressive History and Multi-Player Monte-Carlo Tree Search Solver (MP-MCTS-Solver). He analyzed the performance of these enhancements in two different multi-player games: Focus and Chinese Checkers. Based on experimental results he showed that Progressive History is a considerable improvement in both games. MP-MCTS-Solver, using the standard update rule, is a genuine improvement in Focus.

HENDRIK BAIER (OSNABRÜCK UNIVERSITY)

Hendrik Baier discussed a new simulation strategy for MCTS. The MCTS algorithm builds a search tree by playing many simulated games (*playouts*). Each playout consists of a sequence of moves within the tree followed by many moves beyond the tree. Moves beyond the tree are generated by a biased random *simulation strategy*. The recently published last-good-reply strategy of Peter Drake selects moves that, in previous playouts, have been successful replies to immediately preceding moves. In his talk, Hendrik Baier presented a modification of this strategy that not only remembers moves that recently succeeded, but also immediately forgets moves that recently failed. This modification provides a large improvement in playing strength. He also showed that responding to the previous two moves is superior to responding to the previous one move. Remembering the win rate of every reply performs much worse than simply remembering the last good reply (and worse than not storing good replies at all).

BRUNO BOUZY (UNIVERSITÉ PARIS DESCARTES)

Bruno Bouzy presented a new heuristic search algorithm based on mean values for real-time planning, called Mean-based Heuristic Search for real-time Planning (MHSP). It combines the principles of MCTS with heuristic search in order to create a real-time planner. In contrast to MCTS, at leaf nodes MHSP replaces the simulations (playouts) by heuristic values given by graph-planning techniques. When the heuristic is admissible, the initial mean values of nodes are optimistic, which is a correct way of guiding exploration. In his talk Bruno Bouzy showed promising results of the application of MHSP to the following planning domains: blocksworld, ferry, gripper, and satellite.

GUILLAUME CHASLOT (GOOGLE)

Guillaume Chaslot answered in his talk the question how to enhance MCTS in such a way that a tournament program improves its performance in Go. First, he improved the knowledge in the simulation strategy by learning from move evaluations. Second, he enhanced the selection strategy by proposing progressive strategies for incorporating knowledge. Third, he applied the Cross-Entropy Method to optimize the search parameters of an MCTS program in such a way that its playing strength is increased. Fourth, he designed Meta-MCTS to generate an opening book that improves the performance of an MCTS program.

All five presentations can be downloaded at: www.unimaas.nl/games/symposium2010/

Minutes of the BNVKI/AIABN General Assembly

Tuesday October 26, 2010
Luxembourg

Ann Nowé

Present: Antal van den Bosch (Chair), Virginia Dignum, Jos Uiterwijk, Richard Booth, Ann Nowé, and 24 members.

1. Opening

Chair Antal van den Bosch opens the meeting at 13:45.

2. Minutes of the BNVKI General Assembly of October 31, 2009

The minutes are approved.

3. Announcements

Ann Nowé joined the ECCAI board and has taken over the task of the former treasurer Hendrik Blockeel.

Leon van der Torre says that the people of Luxembourg are positive about joining BNVKI. He clarifies the role of the different partners involved in the organisation of BNAIC 2010: the CRP which is mainly responsible for the logistic support, while the university partners have taken care of the scientific issues. Leon recognises the sponsors.

The organisers received 76 submissions (36 A-papers, 34 B-papers and 6 demos), of which 67 have been accepted (88.16%) (50 oral, 11 posters and 6 demos). On 22/10/10 the organisers received 99 registrations (50 from The Netherlands, 18 from Belgium, 22 from Luxembourg, and 9 others).

Leon van der Torre continues by explaining that the funding by the Luxembourg science foundation implies that BNAIC'10 cannot make profit, so the registration fees were kept low. This year proceedings are available on a USB stick, printed proceedings can be optionally ordered. The board believes that providing electronic proceedings is the way to go, it also is the tendency at big conferences.

Some members suggest that for next editions papers should be made available one week ahead, so people can make printouts, and a booklet with abstracts is also advisable. This idea will be considered for the next BNAIC.

4. Financial report 2009

Virginia Dignum, treasurer, reports on the financial situation of the BNVKI. She explains that the difference between what was estimated and realised, is mainly due to increasing secretarial costs (increase of prices and the fact that there are 2 locations (Maastricht and Tilburg). Before, NWO was supporting the BNVKI via Maastricht. This budget was used to cover the secretarial costs. Now, NWO supports the BNVKI by sponsoring the registration of Dutch students attending the BNAIC.

Although some money from the reserve was used, the assets are still quite large: 40,038 euro. However, the board recognises that the secretarial costs should be reduced and will discuss how this can be done. One idea is that the number of newsletters will be reduced to 4 per year, and will be complemented with a monthly updated agenda on the website.

The auditing committee consisting if Menno van Zaanen and Birna van Riemsdijk approved the expenses.

5. Auditing committee 2010

Treasurer Virginia Dignum proposes the new auditing committee, which checks the financial report to be delivered at the next General Assembly, to consist of Silja Renooij and Mark Hagedoorn. This proposal is accepted by the assembly.

6. Progress report 2010 and plans for 2011

Antal van den Bosch explains that the membership of BNVKI implies a membership of ECCAI, which allows a reduction of the registration fee at ECAI and free access to *AI Communications*.

In The Netherlands the NSVKI, the organisation of AI students, is again active and is publishing the newsletter *De Connectie* and organising the BNAIS student conference. Annette ten Teije is responsible within the board to keep the link with the students.

Antal van den Bosch provides an overview of past Benelux AI events.

- The Magic of Monte-Carlo Tree Search, Symposium, September 29, 2010, Department of Knowledge Engineering, Maastricht University, The Netherlands
- Confluences in Models of Rationality, Workshop, May 30, 2010, Katholieke Universiteit Leuven, Belgium
- BENELEARN 2010, the annual machine-learning conference of Belgium and The Netherlands, May 27-28, 2010, Katholieke Universiteit Leuven, Belgium
- SOKS Symposium on the link between Computational Intelligence and the Semantic Web, April 28-29, 2010, Vrije Universiteit Amsterdam, The Netherlands
- Formal Models of Norm Change 2, Workshop, January 18-19, 2010, University of Amsterdam, The Netherlands
- ESAW 2009, The 10th Annual International Workshop "Engineering Societies in the Agents' World", November 18-20, 2009, Utrecht University, The Netherlands
- SRL-2009 (International Workshop on Statistical Relational Learning), ILP-2009 (19^{th} International Conference on Inductive Logic Programming), and MLG-2009 (7^{th} International Workshop on Mining and Learning with Graphs), July 2-4, 2009, Katholieke Universiteit Leuven, Belgium
- INTETAIN 09, The 3^{rd} International Conference on Intelligent Technologies for Interactive Entertainment, June 22-24, 2009, Amsterdam, The Netherlands
- BENELEARN 09, The 18^{th} Annual Belgian-Dutch Conference on Machine Learning, May 18-19, 2009, Tilburg University, The Netherlands

- GALP, ICR International Symposium on Games, Argumentation, and Logic Programming, April 23-24, 2009, Luxembourg
- Rich Cognitive Models for Policy Design and Simulation, January 12-16, 2009, Lorentz Centre, Leiden, The Netherlands

Antal van den Bosch invites the audience to propose ideas of what the BNVKI can offer extra to its members, to attract also members who didn't attend BNAIC.

Some suggestions from the audience were:

- The board could consider to support the organisation of ECAI or similar events in the Benelux.
- The BNVKI should be represented at local events such as SIREN.
- The BNVKI should be represented in social networks such as LinkedIn, so a member benefits by a higher visibility.
- A list of ECCAI fellows should be available on the BNVKI website.

Finally, a member remarks that the secretarial costs of 60% is high. The board agrees with this and has already ideas to reduce it and in the budget for next year this percentage is lowered.

7. BNVKI board

Antal van den Bosch gives an overview of the current board members and informs the meeting that Sien Moens is stepping out the board, because her term is finished. Also Richard Booth wants to resign; he was the first postdoc member who joined the board and the idea was that postdocs would typically serve a shorter term. At the next general assembly some new board members will be elected.

8. BNAIC'11

The board is happy to announce that Katja Verbeeck and Partick De Causmaeker of the CODeS group, KULeuven, have accepted to organize BNAIC 2011, which will be held on November 3 and 4, 2011 at Ghent, Campus KaHo St. Lieven. Due to lack of time, the presentation is postponed till the closing session.

9. End of meeting

There are no more comments or questions. The meeting is closed at 14:30.

The 2010 SKBS-Strukton Prize

Jaap van den Herik
Director of SKBS

The Foundation for Knowledge Based Systems (SKBS) and the Company Strukton continued their policy of awarding the SKBS-Strukton prize to the best demonstration of the Demo-session of the BNAIC 2010. Since the BNAIC was held as far as in Luxembourg, the company Strukton was unable to attend this BNAIC. The active BNVKI member Bas Obladen (who is a regular visitor since the mid 1990s) and Carlos F. Bosma MSc had to give preference to their business activities. Yet, their interest in new ideas is so intensive that they confirmed to support the SKBS-Strukton prize also in 2010 with an amount of Euro 350, bringing the full prize to Euro 700. We are grateful for this generous gesture and wish the company a prosperous climate for the months to come. We hope to see Obladen and Bosma again in Ghent in Belgium on November 3 and 4, 2011 at the 23rd BNAIC.

The 2010 referee committee consisted of Jaap van den Herik (chair), Han La Poutré (former chair BNVKI), Hendrik Blockeel (Belgium, Leuven), Leon van der Torre (Luxembourg), Birna van Riemsdijk (Delft University of Technology), and Niek Wijngaards (businessman and researcher at Thales). Jaap van den Herik acted as non-voting chair owing to the involvement of Jeroen Janssens and Eric Postma.

The referee committee had to consider six submissions which were eligible for the SKBS-Strukton prize. In Table 1 we list them by topic (in the order of their publication in the Conference Program BNAIC 2010).

Since 1999 we have seen many different appearances of the Demo-session. The common characteristic is the emphasis on being "an industrial exhibition". Up to 2006 the prize money was provided by SKBS only. The Foundation for Knowledge Based Systems originates from the late 1980s as a foundation within SPIN (Stimulerings Projectteam In Nederland). The Foundation SNN (Stichting Neurale Netwerken) is another well-known member of the former SPIN. They supported SKBS financially with augmenting the SKBS prize in 2007. In 2008, the industrial partner Strukton announced its willingness to participate in the prize funding. The extra contribution was gratefully accepted. As stated above they continued this policy in 2009 and 2010.

1. Disaster Victim Identification System *Willem Burgers, Wim Wiegerinck, and Bert Kappen* **2. Creating artificial vessel trajectories with Presto** *Jeroen Janssens, Hans Hiemstra, and Eric Postma* **3. Road Network Hierarchy Generation by Distributed Agents for a Routing Application** *Joris Maervoet, Lode Blomme, Katja Verbeeck, and Greet Vanden Berghe* **4. RoboCup.be – A testbed for intelligent soccer strategies** *Tim Vermeulen and Katja Verbeeck* **5. ConfigNow: a knowledge based approach to configuration software** *Hanne Vlaeminck, Kumar Abhinav, Joost Vennekens, Marc Denecker, and Johan Wittocx* **6. The IDP System** *Johan Wittocx, Broes De Cat, and Marc Denecker*

Table 1: The 2010 candidates of the SKBS-Strukton prize.

In 2010, six submissions were exhibited in the demonstration room Salle Nancy where the BNAIC took place. All six were full-fledged demonstrations. It was really a pleasure to walk along the demos and to discuss them with the stand holders. Again, the quality and maturity of presentation, particularly the quality of ideas, has grown considerably over the last year.

The referee committee had a difficult task. The procedure went in shifts: from six we reduced the number of candidates to three and then to two and finally we arrived at one.

The members of the referee committee were invited to score on (a) the originality of the submission, (b) the applicability (scalability), (c) the AI content, (d) the actual applicability in society, (e) the presentation, (f) the applicability in other domains, (g) the applicability by other persons, and (h) open source.

There was a long discussion which finally led to the following winners: Willem Burgers, Wim Wiegerinck, and Bert Kappen for their demo *Bonaparte Disaster Victim Identification System*. The second place was for the IDP system, and the third place for Presto.

Thus, the winning demo was BONAPARTE, which has been used by the NFI (Netherlands Forensic Institute) for identification of the victims of the Tripoli Air Crash in the beginning of 2010. Then it was a β-side program, but now it is accepted as an asset to help the NFI, Schiphol, and the KLPD with their daily work.

Congratulations to the Nijmegen team (RUN), especially to Bert Kappen (SNN) who has won for the second time the SKBS-Strukton prize (SKBS 2003, SKBS-Strukton 2010).

In Table 2 we provide an overview of the winners of the SKBS-Strukton prize so far.

1999 Maastricht M. van Wezel, J. Sprenger, R. van Stee, and H. La Poutré *Neural Vision 2.0 – Exploratory Data Analysis with Neural Networks*
2000 Kaatsheuvel (shared prize) E. Zopfi *HKT* G. Schram *LubeSelect*
2001 Amsterdam Alexander Ypma, Rob Kleiman, Jan Valk, and Bob Duin *MINISOM – A System for Machine Health Monitoring with Neural Networks*
2002 Leuven F. Brazier, D. Mobach, and B. Overeinder *AgentScape Demonstration*
2003 Nijmegen Bert Kappen, Wim Wiegerinck, Ender Akay, Marcel Nijman, Jan Neijt, and André van Beek *Promedas: A Diagnostic Decision Support System*
2004 Groningen Wouter Teepe *The Secret Prover: Proving Possession of Arbitrary Files While not Giving Them Away*
2005 Brussels Gerald de Jong *Fluidiom: The Evolution of Locomotion*
2006 Namur Marion Verduijn, Niels Peek, Peter Rosseel, Evert de Jonge, and Bas de Mol *Procarsur: A System for Prognostic Reasoning in Cardiac Surgery*
2007 Utrecht Tim Harbers, Rob van der Veen, and Marten den Uyl *Sentient Demonstration BNAIC 07: Vicavision*

2008 Enschede (shared prize) Joris Maervoet, Patrick De Causmaecker, and Greet Vanden Berghe *A Generic Rule Miner for Geographic Data* and Dennis Reidsma and Anton Nijholt *Temporal Interaction between an Artificial Orchestra Conductor and Human Musicians*
2009 Eindhoven Tom van Bergen, Maarten Brugmans, Bart Dohmen, and Niels Molenaar *Cobes: The clean, safe and hospitable metro*
2010 Luxembourg Disaster Victim Identification System *Willem Burgers, Wim Wiegerinck, and Bert Kappen*

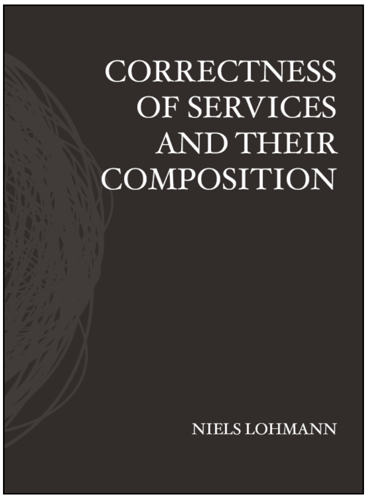
Table 2: Overview of SKBS-Strukton prize winners.

PH.D. THESIS ABSTRACTS

Correctness of Services and their Composition

Ph.D. thesis abstract
Niels Lohmann

Promotores: Prof.dr. W.M.P. van der Aalst, Prof.dr. K. Wolf
Copromotor: Dr. N. Sidorova
Date of defense: September 27, 2010



Service-oriented computing (SOC) is an emerging paradigm of system design and aims at replacing complex monolithic systems by a composition of interacting systems, called services. A service encapsulates self-contained functionality and offers it over a well-defined, standardized interface. This modularization may reduce both complexity and cost. At the same time, new challenges arise with the distributed execution of services in dynamic compositions. In particular, the correctness of a service composition depends not only on the local correctness of each participating service, but also on the correct interaction between them. Unlike in a centralized monolithic system, services may change and are not completely controlled by a single party.

We study correctness of services and their composition and investigate how the design of correct service compositions can be systematically supported. We thereby focus on the communication protocol of the service and approach these questions using formal methods and make contributions to three scenarios of SOC.

The correctness of a service composition depends on the correctness of the participating services. To this end, we (1) study correctness criteria which can be expressed and checked with respect to a single service. We validate services against behavioral specifications and verify their satisfaction in any possible service composition. In case a service is incorrect, we provide diagnostic information to locate and fix the error.

In case every participating service of a service composition is correct, their interaction can still introduce problems. We (2) automatically verify correctness of service compositions. We further support the design phase of service compositions and present algorithms to automatically complete partially specified compositions and to fix incorrect compositions.

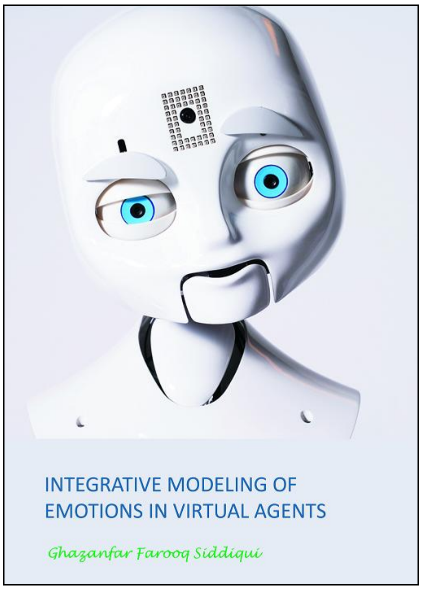
A service composition can also be derived from a specification, called choreography. A choreography globally specifies the observable behavior of a composition. We (3) present an algorithm to deduce local service descriptions from the choreography which – by design – conforms to the specification.

All results have been expressed in terms of a unifying formal model. This not only allows to formally prove correctness, but also makes results independent of the specifics of concrete service description languages. Furthermore, all presented algorithms have been prototypically implemented and validated in experiments based on case studies involving industrial services.

Integrative Modeling of Emotions in Virtual Agents

Ph.D. thesis abstract
Ghazanfar Farooq Siddiqui

Promotor: Prof.dr. J. Treur
Copromotores: Dr. T. Bosse, Dr. J.F. Hoorn
Date of defense: September 28, 2010



In recent years, researchers have become increasingly interested in the application of intelligent virtual agents in various domains. Intelligent virtual agents (IVAs) are autonomous, graphically embodied agents in a virtual environment that are able to interact intelligently with the environment, other IVAs, or with human users. Recently, much research has been dedicated to developing virtual agents with more realistic graphical representations. However, the affective properties of such agents are usually rather limited, and not very human-like. For example, although many IVAs currently have the ability to somehow show emotions by means of different facial expressions, it is quite difficult for them to show the right emotion at the right moment. One step further, it is even more difficult for them to actually understand and react empathically to the emotional state of other agents. This is in conflict with the requirement of virtual agents to closely mimic human affective behavior. Several studies in Social Sciences have shown that this is an important prerequisite for an agent to increase human involvement in a virtual environment. Therefore, existing systems based on IVAs are not as effective as they could be. Properties that they typically lack are the ability to show emotions (not only in terms of facial expression, but also in terms of behavior), in relation to insight in each other's and humans' cognitive and affective states.

To deal with this problem, some authors propose to increase the affective properties of interactive software agents by using knowledge from psychology and cognitive science as a basis for computational modeling of the cognitive and affective processes involved. Recently, a variety of such computational models have been developed for different aspects of human behavior. Examples include models for reasoning processes, visual attention, emotion regulation, mindreading, stress and workload, and moods. If such computational models are available in a formal format, this opens the possibility to equip IVAs with them. However, the step from existing computational models (that are mostly used for simulation purposes) to models that can directly be plugged into a virtual (3D) environment in such a way that the IVAs behave according to the cognitive model, is a nontrivial one. In reality, this step involves an iterative process, consisting of, among others, the following sub-tasks: refinement of the computational model, translation to a specific programming environment, and testing and evaluation of the resulting model in the virtual setting.

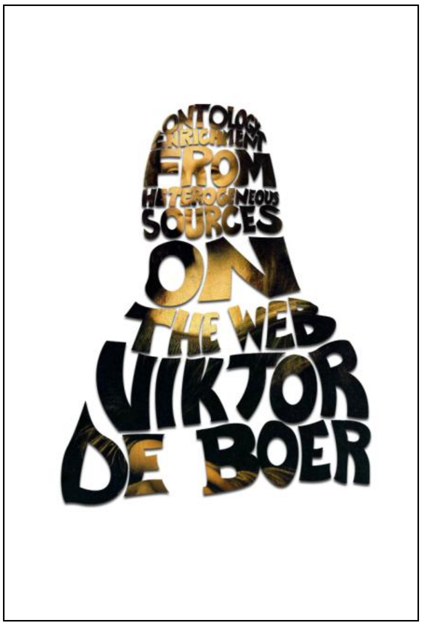
The main research goal of this thesis is to explore how computational models of affect can be integrated within virtual agents. To this end, different theories were taken from different fields (e.g., Social Sciences, Psychology, Economics and Finance) and have been combined to develop integrative models of affect. Most of these models have first been used for simulation, to test whether their overall behavior was satisfactory. For this, various modeling environments have been used, such as LEADSTO and C++. Next, these models have been incorporated within applications related to health care, games and business context. These applications have been developed using JavaScript, in combination with the Vizard toolkit. Finally, for some applications, user tests have been performed. The preliminary results of these tests show that through the developed models, users are more involved in the applications.

In the first part of the thesis, the extent to which an agent becomes involved with another agent, or stays at a distance from it, is modeled, by formalizing in a computational manner the (previously informally described) I-PEFiC model. These extents depend upon different aspects of the other agent such as ethics (good or bad), aesthetics (beautiful or ugly), epistemics or realism (how realistic or unrealistic the other agent is), similarity (resemblances between the two agents) and affordances the other agent offers as an aid or obstacle for task performance of the agent. Second, the integration of three models of affect is addressed. The approach taken in this part is to select three of the more influential models, which share that they can be used to enhance believability of virtual characters: CoMERG (Computational Model of Emotion Regulation based on Gross theory), EMA (Emotion and Adaptation) and I-PEFiC^{ADM} (I-PEFiC extended with a module for Affective Decision Making). Then, we focus on modeling affective states of a person in economical context. This economical context is considered a large-scale multi-agent system consisting of thousands or millions of other agents. In this part it is described how a person's involvement with these other agents in the form of individual investment decisions depends on a personal risk profile, the state of greed of the person, and the state of the world economy. Finally, in the last part of the thesis, applications of virtual agents that show emotions in the domains of health care, business, and games are presented and evaluated.

Ontology Enrichment from Heterogeneous Sources on the Web

Ph.D. thesis abstract
Victor de Boer

Promotor: Prof.dr. B.J. Wielinga
Copromotor: Dr. M. van Someren
Date of defense: September 30, 2010



This thesis is about semi-automatic methods that extract structured information from unstructured documents on the Web.

The Semantic Web is a proposed extension of the current World Wide Web (WWW). Where the WWW is a web of documents connected through hyperlinks, the Semantic Web is a web of interlinked data. The documents on the WWW are designed to be read and interpreted by humans. Since the information on the Semantic Web appears in a formalized form, it can be interpreted and reasoned with by computer applications.

An example of the advantages of having information available in this formalized form can be seen in the MultimediaN E-Culture demonstrator. This application uses background knowledge to connect metadata descriptions of objects of Dutch cultural heritage institutions to each other. The fact that the terms that are used as metadata in the object descriptions are interlinked allows the application to perform search and browsing tasks that can not be performed by existing text-based search systems. If, for instance, we know that Vincent van Gogh is considered a Post-Impressionist, we can deduce that his paintings have some relation to this style and to other works of other Post-Impressionists.

Problem and Approach

The information or knowledge used in the Semantic Web can be divided into two categories: The first category is that of meta-knowledge that describes which classes exist in a domain and which relations hold between these classes. For the cultural-heritage domain these could include classes such as "Artist", "Art_Style" and "Work". Possible relations are "has_painted" or "has_style". This meta-knowledge is formalized in ontologies. The second category is that of the specific instances of the classes and relations in the domain. These instances are stored in a knowledge base. Here we can find which artists exist, to which specific style they belong and which artworks they have produced.

In this thesis, we focus on enriching knowledge bases defined by ontologies. We describe methods that automatically identify new instances of relations and classes. Since in a lot of domains the target information is already available on the WWW in human-readable form, the methods use the WWW as the source from which to extract the target information. Information Extraction (IE) is the standard name for the task of extracting structured information from unstructured documents.

Current IE methods use extraction techniques that make use of extensive Natural Language Processing (NLP) or are based on wrappers that exploit explicit structures in documents. NLP-based methods try to model the natural-language sentences in order to understand the syntactic structure of the sentence. Using the syntactic structure, the semantic information can be extracted. This approach has the downside that the learned syntactic model is language- and structure-dependent. Wrapper-based methods try to extract the target information from structures such as tables and lists. The performance of these methods is also dependent on the availability of such structures in the documents.

In this thesis, we introduce a number of methods that are designed to be less dependent on language and document structure. The methods use simple matching techniques to identify potential target information in documents and use simple term co-occurrence statistics to determine the likelihood of the information being correct. By using simple matching and co-occurrence methods, the proposed Information Extraction methods are able to extract and aggregate information from many documents. This increases the robustness of the methods and allows us to exploit the redundancy of information on the Web. The proposed methods all extract a working corpus of documents from the Web, for which the co-occurrence statistics are calculated.

In chapters 2, 3 and 4 of this thesis, we present IE methods that extract instances of relations and classes defined in ontologies. We evaluate these methods by having them extract cultural-heritage instances that are to be used by the MultimediaN E-culture application. The effectiveness of the methods are further evaluated using similar tasks in other domains.

A METHOD FOR EXTRACTING RELATION INSTANCES

In Chapter 2, we describe a method that extracts instances of relations (of the form [Subject, relation, Object]) from the WWW. We use a semi-supervised approach, where we assume that we have a very small collection of example relation instances which we aim to automatically expand. The method further assumes that all instances of the subject and object class of the relation are known: with this method, we do not extract new class instances, only instances of relations.

An example of this task is the extraction of instances of the relation [Art_Style, has_artist, Artist]. Given are lists of both art styles and artists (both are indeed available in the MultimediaN E-culture knowledge base) and a small seed set of relation instances ([Post-Impressionism, has_artist, Vincent_van_Gogh], [Post-Impressionism, has_artist, Gauguin], etcetera). The specific task is to expand this list of relation instances.

The method described in Chapter 2 follows a number of steps: First a working corpus of documents is created. For this, the Google search engine is queried with a search term that matches one subject instance (for example “Post-Impressionism”). A limited number of resulting documents is saved in memory. In the next step, the method identifies instances of the relation’s object (for example: Artists names) in the working corpus’ documents. The found instances are part of candidate relation instances. Next, for every document, a Document Score is determined: the total number of object instances that are in the seed set identified in that document divided by the total number of object instances in that document. This score represents the level to which the target relation is represented in the document. For each candidate object instance, the Document Scores are added and normalized, resulting in an Instance Score. This Instance Score is the likelihood that a candidate relation instance is correct. The relation instance with the highest associated Instance Score is added to the seed set, at which point all scores are re-calculated. The iterative process expands the seed set of relation instances step-by-step and this makes it possible to start with a small seed set: in the beginning, the ‘easy’ correct expansions will be extracted. As the number of iterations grows, so does the seed set and new instances will be identified using more and more knowledge. To stop this iterative process, we introduce the Drop Factor: when we observe a relative drop in the Instance Score below some threshold value, the method halts and presents the current seed set as its final result.

We evaluated this so-called ‘redundancy method’ using experiments in two domains. In the first experiment, we extracted instances of the relation [Art_Style, has_artist, Artist] introduced earlier. We extracted relation instances for ten different art styles starting with initial seed sets of three artists. We manually evaluated 400 extracted relation instances. The results show that the Drop Factor allows us to select either higher precision (between 0.62 and 0.92) or recall (between 55 and 247 correct relation instances). We compared the rankings for the 400 relation instances based on our

Instance Score with those of a popular semantic distance measure, the Normalized Google Distance. The results showed that the ranking based on Instance Scores is considerably better. In a second experiment, we used the method to extract instances of the relation [Football_Club, has_player, Football_Player]. The results of this experiment showed that the method achieves the same performance level as for the first experiment. The optimal drop factor was also the same for the two tasks.

EXTRACTING TIME PERIODS

In Chapter 3, we describe a method for a more specific extraction task. Here the goal is to extract time periods for (historical) concepts. In the MultimediaN E-culture vocabularies, this temporal information is not available (for example, for art styles), even though such information can be of great value.

The method that we present has two main phases. In the first phase, we first extract a working corpus of documents by querying the Google search engine with the target concept's label. The method then identifies potential occurrences of years (four digit numbers) in all documents. The frequency distribution of these years is normalized by the frequency distribution of years on the whole Web. After this normalization step, we are left with a frequency distribution that is representative of the search term. We fit a Normal distribution to this data. The parameters of this model allows us to determine the start and end years of a time period.

We evaluated this first phase by extracting time periods for art styles, wars, historical periods and artists. We compared the results to a manually created gold standard. The results show a relatively low error that is also stable for the four types of concepts.

In the second phase the extracted time periods, represented by the model parameters are transformed into instances of a structured vocabulary. The structured vocabulary used is the TIMEX2 annotation standard that provides a normalized representation of time expressions. In Chapter 3, we present a number of rewriting rules that convert a frequency model into a TIMEX2 instance. We evaluated these rules by applying them to the extracted periods for the instances of the four classes previously introduced. Manual evaluation shows that 88.75% of the TIMEX2 periods where considered either completely or partly correct. Again, this performance was stable for the different concept types.

PATTERN SPECIFICITY FOR EXTRACTION PATTERNS

In Chapters 2 and 3, we use very simple, general methods that are able to extract information from different varying documents. In Chapter 4, we take a closer look at the effect of using more general or more specific extraction methods. We do this by introducing a extraction method that uses a text-analysis application, tOKo. With tOKo, we can search in corpus documents using patterns, in which semantic classes can be used. Using these, we can extract relation instances such as [Art_Style, has_artist, Artist] from the documents.

To test the effect that the specificity of the patterns has on the performance of the extraction method, we use patterns of varying specificity. Patterns that are more specific make less errors and therefore result in a higher precision, while more general patterns result in a higher recall, usually at the cost of precision. However, when more general patterns are used, information from different documents can be aggregated. When a threshold value on the frequency of pattern hits is used, a high precision can be attained. Overall, when measuring the performance using the F-measure, using general patterns produces better results. The patterns were evaluated in multiple domains, for multiple relation instantiation tasks.

COMBINING BACKGROUND KNOWLEDGE AND INFORMATION SOURCES

In the evaluation of the previously described methods, we found that the use of background knowledge for filtering out faulty candidate instances can raise the performance of the overall extraction task. A second improvement can be obtained through the combination of multiple information sources. This corresponds to the idea in those chapters that using redundancy and multiple sources is beneficial to the performance of IE methods.

In Chapter 5 we present a method that uses background knowledge and multiple information sources to lower the likelihood of faulty candidate relation instances (and to effectively filter them out) on one hand and to raise the likelihood of correct candidate instances on the other hand. We use background knowledge rules that produce a likelihood score for a candidate relation instance using the background knowledge available in the ontology and knowledge base. In Chapter 5, we present a number of example rules.

One of these rules states for example that it is unlikely that a relation of the type "participates_in" holds between a subject and an object that have non-overlapping time periods (e.g. it is unlikely that an artist belongs to an artstyle that started after his death). The rules generate a likelihood score. The different information sources also produce a likelihood score for a candidate relation instance. In Chapter 5, we discuss a number of methods through which all these scores can be combined.

Through a number of experiments, we show that combining information from different sources with available background knowledge indeed raises the quality of a set of candidate relation instances. Errors that are made by one method can be corrected by a different source of information such as temporal or spatial background knowledge. More incorrect candidate relation instances are added to the knowledge base and the scores of more incorrect instances drop below the threshold and are filtered out.

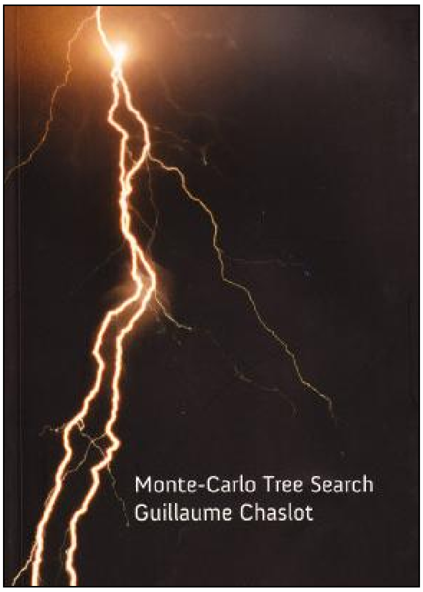
Conclusion

One observation that we made from the evaluation of the methods presented in this thesis is that the simple, general extraction methods do indeed extract information from different sources. By aggregating this information from heterogeneous sources and using threshold values on frequencies of occurrences, good precision scores can still be obtained. We here use the redundancy of information on the Web to our advantage. The simple extraction methods are indeed less dependent on the language and structure of the documents and can be re-used for different extraction tasks in different domains. Evaluation showed that the performance for the methods is stable in different domains. By using multiple sources, more information can be extracted, leading to a higher recall. At a meta-level, multiple Information Extraction sources can also be combined, leading to even more exploitation of redundancy. In combination with available background knowledge, this can also be used to filter candidate information, in order to achieve higher recall and precision.

Monte-Carlo Tree Search

Ph.D. thesis abstract
Guillaume Chaslot

Promotor: Prof.dr. G. Weiss
Copromotores: Dr. M.H.M. Winands, dr. B. Bouzy, dr.ir. J.W.H.M. Uiterwijk
Date of defense: September 30, 2010



This thesis studies the use of Monte-Carlo simulations for tree-search problems. The Monte-Carlo technique we investigate is Monte-Carlo Tree Search (MCTS). It is a best-first search method that does not require a positional evaluation function in contrast to αβ search. MCTS is based on a randomized exploration of the search space. Using the results of previous explorations, MCTS gradually builds a game tree in memory, and successively becomes better at accurately estimating the values of the most promising moves. MCTS is a general algorithm and can be applied to many problems. The most promising results so far have been obtained in the game of Go, in which it outperformed all classic techniques. Therefore Go is used as the main test domain.

Chapter 1 provides a description of the search problems that we aim to address and the classic search techniques which are used so far to solve them. The following problem statement guides our research.

Problem statement: *How can we enhance Monte-Carlo Tree Search in such a way that programs improve their performance in a given domain?*

To answer the problem statement we have formulated five research questions. They deal with (1) Monte-Carlo simulations, (2) the balance between exploration and exploitation, (3) parameter optimization, (4) parallelization, and (5) opening-book generation.

Chapter 2 describes the test environment to answer the problem statement and the five research questions. It explains the game of Go, which is used as the test domain in this thesis. The chapter provides the history of Go, the rules of the game, a variety of game characteristics, basic concepts used by humans to understand the game of Go, and a review of the role of Go in the AI domain. The Go programs MANGO and MOGO, used for the experiments in the thesis, are briefly described.

Chapter 3 starts with discussing earlier research about using Monte-Carlo evaluations as an alternative for a positional evaluation function. This approach is hardly used anymore, but it established an important step towards MCTS. Subsequently, a general framework for MCTS is presented in the chapter. MCTS consists of four main steps: (1) In the *selection step* the tree is traversed from the root node until we reach a node, where we select a child that is not part of the tree yet. (2) Next, in the *expansion step* a node is added to the tree. (3) Subsequently, during the *simulation step* moves are played in self-play until the end of the game is reached. (4) Finally, in the *backpropagation step*, the result of a simulated game is propagated backwards, through the previously traversed nodes.

Each step has a strategy associated that implements a specific policy. Regarding selection, the UCT strategy is used in many programs as a specific selection strategy because it is simple to implement and effective. A standard selection strategy such as UCT does not take domain knowledge into account, which could improve an MCTS program even further. Next, a simple and efficient strategy to expand the tree is creating one node per simulation. Subsequently, we point out that building a simulation strategy is probably the most difficult part of MCTS. For a simulation strategy, two balances have to be found: (1) between search and knowledge, and (2) between exploration and exploitation. Furthermore, evaluating the quality of a simulation strategy has to be assessed together with the MCTS program using it. The best simulation strategy without MCTS is not always the best one when using MCTS. The backpropagation strategy that is the most successful is taking the average of the results of all simulated games made through a node.

Finally, we give applications of MCTS to different domains such as Production Management Problems, Library Performance Tuning, SameGame, Morpion Solitaire, Sailing Domain, Amazons, Lines of Action, Chinese Checkers, Settlers of Catan, General Game Playing, and in particular Go.

The most basic Monte-Carlo simulations consist of playing random moves. *Knowledge* transforms the plain random simulations into more sophisticated *pseudo-random* simulations. This has led us to the first research question.

Research question 1: *How can we use knowledge to improve the Monte-Carlo simulations in MCTS?*

Chapter 4 answers the first research question. We explain two different simulation strategies that apply knowledge: urgency-based and sequence-like simulation. Based on the experience gathered from implementing them in INDIGO and MOGO, respectively, we make the following three recommendations. (1) Avoiding big mistakes is more important than playing good moves. (2) Simulation strategies using sequence-like simulations or patterns in urgency-based simulations are efficient because they simplify the situation. (3) The simulation strategy should not become too stochastic, nor too deterministic, thus balancing exploration and exploitation.

Moreover, we develop the first efficient method for learning automatically the knowledge of the simulation strategy. We proposed to use *move evaluations* as a fitness function instead of learning from the results of simulated games. A coefficient is introduced that enables to balance the amount of exploration and exploitation. The algorithm is adapted from the tracking algorithm of Sutton and Barto. Learning is performed for 9×9 Go, where we showed that the Go program INDIGO with the learnt patterns performed better than the program with expert patterns.

In MCTS, the selection strategy controls the balance between exploration and exploitation. The selection strategy should favour the most promising moves (exploitation). However, less promising moves should still be investigated sufficiently (exploration), because their low scores might be due to unlucky simulations. This move-selection task can be facilitated by applying knowledge. This idea has guided us to the second research question.

Research question 2: *How can we use knowledge to arrive at a proper balance between exploration and exploitation in the selection step of MCTS?*

Chapter 5 answers the second research question by proposing two methods that integrate knowledge into the selection step of MCTS: progressive bias and progressive widening. Progressive bias uses knowledge to direct the search. Progressive widening first reduces the branching factor, and then increases it gradually.

We refer to them as "progressive strategies" because the knowledge is dominant when the number of simulations is small in a node, but loses influence progressively when the number of simulations increases.

First, the progressive strategies are tested in MANGO. The incorporated knowledge is based on urgency-based simulation. From the experiments with MANGO, we observe the following. (1) Progressive strategies, which focus initially on a small number of moves, are better in handling large branching factors. They increase the level of play of the program MANGO significantly, for every board size. (2) On the 19×19 board, the combination of both strategies is much stronger than each strategy applied separately. The fact that progressive bias and progressive widening work better in combination with each other shows that they have complementary roles in MCTS. This is especially the case when the board size and therefore branching factor grows. (3) Progressive strategies can use relatively expensive domain knowledge with hardly any speed reduction.

Next, the performance of the progressive strategies in other game programs and domains is presented. Progressive bias increases the playing strength of MOGO and of the Lines-of-Action program MC-LOA, while progressive widening did the same for the Go program CRAZY STONE. In the case of MOGO, progressive bias is successfully combined with RAVE, a similar technique for improving the balance between exploitation and exploration. These results give rise to the main conclusion that the proposed progressive strategies are essential enhancements for an MCTS program.

MCTS is controlled by several parameters, which define the behaviour of the search. Especially the selection and simulation strategies contain several important parameters. These parameters have to be optimized in order to get the best performance out of an MCTS program. This challenge has led us to the third research question.

Research question 3: *How can we optimize the parameters of an MCTS program?*

Chapter 6 answers the third research question by proposing to optimize the search parameters of MCTS by using an evolutionary strategy: the Cross-Entropy Method (CEM). CEM is related to Estimation-of-Distribution Algorithms (EDAs), a new area of evolutionary computation. The fitness function for CEM measures the winning rate for a batch of games. The performance of CEM with a fixed and variable batch size is tested by tuning 11 parameters in MANGO. Experiments reveal that using a batch size of 500 games gives the best results, although the convergence is slow. A small (and fast) batch size of 10 still gives a reasonable result when compared to the best one. A variable batch size performs a little bit worse than a fixed batch size of 50 or 500. However, the variable batch size converges faster than a fixed batch size of 50 or 500.

Subsequently, we show that MANGO with the CEM parameters performs better against GNU GO than the MANGO version without. In four self-play experiments with different time settings and board sizes, the CEM version of MANGO defeats the default version convincingly each time. Based on these results, we may conclude that a hand-tuned MCTS using game engine may improve its playing strength when re-tuning the parameters with CEM.

The recent evolution of hardware has gone into the direction that nowadays personal computers contain several cores. To get the most out of the available hardware one has to parallelize MCTS as well. This has led us to the fourth research question.

Research question 4: *How can we parallelize MCTS?*

Chapter 7 answers the fourth research question by investigating three methods for parallelizing MCTS: leaf parallelization, root parallelization and tree parallelization. Leaf parallelization plays for each available thread a simulated game starting from the leaf node. Root parallelization consists of building multiple MCTS trees in parallel, with one thread per tree. Tree parallelization uses one shared tree from which games simultaneously are played.

Experiments are performed to assess the performance of the parallelization methods in the Go program MANGO on the 13×13 board. In order to evaluate the experiments, we propose the strength-speedup measure, which corresponds to the time needed to achieve the same strength. Experimental results indicate that leaf parallelization is the weakest parallelization method. The method leads to a strength speedup of 2.4 for 16 processor threads. The simple root parallelization turns out to be the best way for parallelizing MCTS. The method leads to a strength speedup of 14.9 for 16 processor threads. Tree parallelization requires two techniques to be effective. First, using local mutexes instead of a global mutex doubles the number of games played per second. Second, the virtual-loss enhancement increases both the games-per-second and the strength of the program significantly. By using these two techniques, we obtain a strength speedup of 8.5 for 16 processor threads.

Modern game-playing programs use opening books in the beginning of the game to save time and to play stronger. Generating opening books in combination with an αβ program has been well studied in the past. The challenge of generating automatically an opening book for MCTS programs has led to the fifth research question.

Research question 5: *How can we automatically generate opening books by using MCTS?*

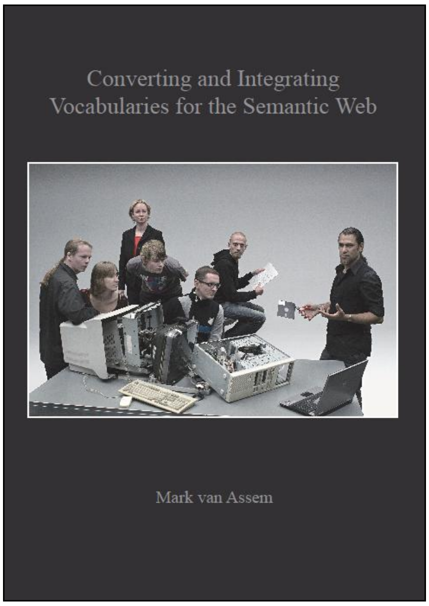
Chapter 8 answers the fifth research question by combining two levels of MCTS. The method is called Meta Monte-Carlo Tree Search (Meta-MCTS). Instead of using a relatively simple simulation strategy, it uses an entire MCTS program (MoGo) to play a simulated game. We describe two algorithms for Meta-MCTS: Quasi Best-First (QBF) and Beta-Distribution Sampling (BDS). The first algorithm, QBF, is an adaptation of greedy algorithms that are used for the regular MCTS. QBF favours therefore exploitation. During actual game play we observe that despite the good performance of the opening book, some branches are not explored sufficiently. The second algorithm, BDS, favours exploration. In contrast to UCT, BDS does not need an exploration coefficient to be tuned. The algorithm draws a move according to its likelihood of being the best move (considering the number of wins and losses). This approach created an opening book which is shallower and wider. The BDS book has the drawback to be less deep against computers, but the advantage is that it stayed longer in the book in official games against humans. Experiments on the Go server CGOS reveal that both QBF and BDS were able to improve MoGo. In both cases the improvement is more or less similar. Based on the results, we may conclude that QBF and BDS are able to generate an opening book which improves the performance of an MCTS program.

The last chapter of the thesis returns to the five research questions and the problem statement as formulated in Chapter 1. Taking the answers to the research questions above into account we see that there are five successful ways to improve MCTS. First, learning from move evaluations improves the knowledge of the simulation strategy. Second, progressive strategies enhance the selection strategy by incorporating knowledge. Third, CEM optimizes the search parameters of an MCTS program in such a way that its playing strength is increased. Fourth, MCTS benefits substantially from parallelization. Fifth, Meta-MCTS generates an opening book that improves the performance of an MCTS program. Yet, we are able to provide additional promising directions for future research. Finally, the question of understanding the nature of MCTS is still open.

Converting and Integrating Vocabularies for the Semantic Web

Ph.D. thesis abstract
Mark van Assem

Promotor: Prof.dr. A.Th. Schreiber
Copromotor: Dr. J.R. van Ossenbruggen
Date of defense: October 1, 2010



Institutions such as libraries, museums and other archives have been collecting books, paintings, statues and other objects for centuries. To manage these collections, cataloguers have described each object with respect to its title, author, subjects, materials and other attributes. This process is called "indexing", and simplifies the process of searching through the collections. An object description created during indexing is essentially a set of attribute-value pairs. Such a description might consist e.g. of pairs *author=Rembrandt*, *title=Anatomy Lesson*, *date=1632*, *type=painting*, *subject=group portrait*. Such descriptions are also called *metadata* (data about the actual object).

The values for the pairs are usually taken from *vocabularies*. Vocabularies are lists of concepts with definitions and play a key role in indexing and search. Firstly, they offer a set of agreed upon concepts that cataloguers can pick from. Secondly, concepts provide a convenient place to group synonymous terms (e.g. “clair-obscure” and “chiaroscuro” which both refer to Rembrandt’s painting style). Thirdly, the concepts usually have a unique identifier, which allows the cataloguer to indicate the correct concept even though the concept has an ambiguous term (e.g. “painting” as the process of applying a protective coating to an object vs. “painting” as the process of creating an expressive or communicative image). Fourthly, the concepts are often placed into a hierarchy (e.g. “origami” below “Japanese art”) which simplifies search (a search for books on Japanese art will also return books on origami).

To aid the indexing process, each institution not only prescribes a set of vocabularies to be used, but also the attributes and their names. Attributes are called elements, and the set of allowed attributes is together called the *metadata element set*.

With the advent of the Web people and institutions have started sharing their data. This has the potential benefit that all information on a particular topic, say the paintings of Vincent van Gogh, can be queried as if they were stored in one system. This requires that the data is *integrated*: it must conform to a particular format and structure that the search system understands. Two obstacles to integration are the different data formats in use (the *syntactic integration problem*) and the different terms in use to denote similar concepts (the *semantic integration problem*). One example of the latter problem is when one institution has a concept called “clair-obscure” and another has a concept called “chiaroscuro”. Another example is when one institution uses a metadata element called “creator” while another has an element called “author”. These concepts and elements are highly similar and a search for clair-obscure paintings by Rembrandt needs to query through both concepts and elements.

The Semantic Web is a research area that proposes particular solutions to these problems. Firstly, it proposes to use a family of web-based knowledge representation languages that have RDF as underlying model (RDFS and the several flavours of OWL). Conversion of data to this family of languages solves a substantial part of the syntactic integration problem. Secondly, the languages have a few simple mechanisms to relate similar concepts to each other, solving part of the semantic integration problem. For example, it is possible to state that “creator” is equivalent to “author”, so that a query on either element will automatically include the results obtained from querying with the other.

In this thesis we assume that the approach and languages proposed by the Semantic Web community are useful for achieving integration, and aim to apply these in the context of the cultural-heritage domain. The problem of converting the original data sets to RDF/OWL has not been investigated much. In this thesis we focus mostly on conversion of vocabularies. Our problem statement is as follows: *How can existing vocabularies be made available to Semantic Web applications?* Problems that need to be solved include understanding the original syntactical format in which the vocabulary is expressed, understanding the conceptual model that lies behind it, linking this conceptual model to that of RDF/OWL, and finding an appropriate way to convert the former model into the latter. These tasks are far from being automated, but the demand for proper conversions will increase in the coming years. Therefore, this thesis has focused on developing *methods* for conversion of vocabularies. Methods are step-wise processes with guidelines that can be followed by people performing the conversion task. Several choices have to be made during the process that affect the resulting representation. Conversion can be performed for the benefit of one particular application and tuned to its specific needs, but conversions can also aim at a representation that is as complete and reusable as possible for any application. The main contribution of this thesis is the development of two separate methods to cater to both situations.

In Chapter 2 we develop a first version of a generic method for conversion. The assumption is that a faithful and complete conversion of the vocabularies results in a representation that is useful for most applications. A method consisting of several steps and guidelines was drafted, and then applied to two case studies: conversions of the MeSH and WordNet vocabularies. These helped to improve the method; they showed which additional guidelines were needed to adequately handle these cases. We deliberately chose two complex vocabularies so that a broad range of vocabulary features were covered. Two tailor-made schemas for each vocabulary and a complete conversion of their content to these schemas are the outcome.

Another way to convert vocabularies suitable for a wide range of applications is to use a standard, widely supported vocabulary schema. In Chapter 3 we developed a method aimed at the emerging SKOS standard. We chose three vocabularies as use cases: a simple one (GTAA), an intermediately complex one (IPSV) and a complex case (MeSH). We found that SKOS was suitable for converting GTAA and IPSV (making use of RDF/OWL abilities to specialize a schema), but MeSH could not be covered completely because SKOS did not allow a concept's terms to be represented as instances themselves.

In Chapter 4 we returned to the problem of generic conversion as approached in Chapter 2. The assumption that our method results in vocabularies useful for many applications was tested by comparing our generic conversion of WordNet to a conversion developed with application use cases in mind (in the W3C Semantic Web Best Practices Working Group). The comparison of the two WordNets showed how our generic method should be changed to cater for a wider range of applications. However, another outcome of the chapter is that a generic method cannot cater to all requirements an application might have, because it may require that content is left out or structured differently than in the original source. We also improved the WordNet conversion to SKOS by applying a newly developed extension that allows terms to be represented as instances themselves. This solves most of the problems noted in the MeSH conversion to SKOS.

Given the results from Chapter 4, we developed a new method that can be used to cater to specific applications in Chapter 5. We adapted the generic method by introducing specific steps to determine the requirements and use cases that need to be covered. The principle of complete and faithful conversion of the original source was dropped. The case study is the MultimediaN E-Culture search and browsing application, for which the AAT, TGN and ULAN were converted. This case study pointed out that the conversion made by the E-Culture team without our method missed several pieces of information needed by the application use cases.

In Chapter 6 we continued to investigate conversions targeted at specific applications. In this chapter we concentrated on alignment applications, which take two or more vocabularies in RDF/OWL as input and produce mapping relations between concepts of the vocabularies. Our analysis showed that these applications cannot handle representations as produced by our methods. We provided a conversion technique to mitigate this problem. The analysis is part of a study on new evaluation techniques for alignments of vocabularies. Alignment is a central ingredient of integration as promoted by the Semantic Web community. The outcome of the study is that our proposed techniques are better tuned to evaluating the quality of an alignment for a particular application than existing techniques.

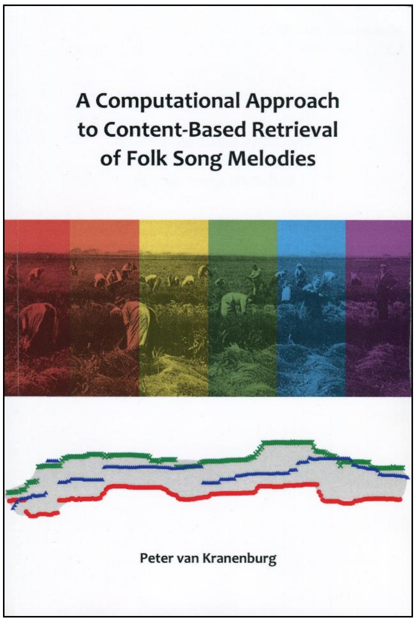
Integration of collections relies on integration of both vocabularies and metadata element sets. In Chapter 7 we study how an existing metadata element set can be represented in RDF/OWL in a way that is interoperable with vocabulary representations as advocated in this thesis. The metadata element set is called VRA and caters specifically to cultural heritage. We show how VRA can be implemented as a specialization of the more generic Dublin Core element set. Linking VRA with Dublin Core allows integration of collections from different domains (television archives, libraries, etcetera) into one search system. We also show how VRA can be specialized to reflect that e.g. the Rijksmuseum uses e.g. ULAN as range of the "creator" element (we term this feature *collection-specific value ranges*).

In summary, this thesis contributes to the integration of metadata collections in three ways. Firstly and chiefly through the development of conversion methods for vocabularies and through contributing actual conversions made with the methods. Secondly, through investigating how metadata schemas can be represented in a way that allows using them together with vocabulary representations produced by the methods. Thirdly, by contributing a study on how alignments can be evaluated on their usefulness for particular applications.

A Computational Approach to Content-Based Retrieval of Folk Song Melodies

Ph.D. thesis abstract
Peter van Kranenburg

Promotores: Prof.dr. R.C. Veltkamp, Prof.dr. L.P. Grijp
Copromotor: Dr. F. Wiering
Date of defense: October 4, 2010



In order to develop a Music Information Retrieval system for folk song melodies, one needs to design an adequate computational model of melodic similarity, which is the subject of this Ph.D. thesis. Since fundamental understanding of melodies in oral culture as well as fundamental understanding of computational methods to model similarity relations between folk song melodies both are necessary, this problem needs a multidisciplinary approach.

Chapter 2 reviews the relevant academic background of both Folk Song Research (as sub-discipline of Ethnomusicology) and Music Information Retrieval. It also presents an interdisciplinary collaboration model in which Computational Musicology serves a 'man-in-the-middle' role. The particular task of Computational Musicology is to design computational models of concepts from Musicology. In the context of this thesis, the concept of *tune family* is the most important.

An important step towards the understanding of the concept of tune family is the method to annotate similarity relations between melodies that is presented in Chapter 3. It has been developed in close collaboration with musicological domain experts to make aspects of their intuitive similarity assessments explicit. 360 melodies in 26 tune families were 'manually' annotated, resulting in an Annotated Corpus that is a valuable resource for the study of melodic similarity and for the evaluation of computational models of melodic similarity. From the annotations we conclude that the relative importance of the various dimensions (global and local rhythm, global and local contour, motifs, lyrics) varies to a large extent in individual comparisons. Furthermore, the dimension of motifs seems the most important for recognizing melodies. This means that in many cases melodies are judged to be related based on shared characteristic melodic motifs.

In Chapter 4, 88 low-level, global, quantitative features of melody are used to discriminate between tune families. It appears that such features can be used to recognize melodies within a relatively small corpus (such as the 360 melodies in the Annotated Corpus), but that these features lose their discriminative power in a larger dataset of thousands of melodies.

Chapter 5 uses the same kind of features in another musical domain: baroque organ fugues. Global features are used to assess authorship problems of fugues that are in the catalog of J.S. Bach. Several hypotheses from musicological literature about the authorship of several fugues are supported by findings using this method. The various degrees of success of the same computational method in the previous and the current chapters show that computational methods cannot 'blindly' be applied to musicological questions.

In Chapter 6, the potential of alignment algorithms for folk song melody retrieval is studied by incorporating musical knowledge in the algorithm in the form of appropriate, musically motivated, substitution scoring functions. This approach leads to good retrieval results both for a small (360 melodies) and a large (4830 melodies) dataset. Furthermore, domain experts were able to classify 'problematic' melodies using the results of the alignment algorithms.

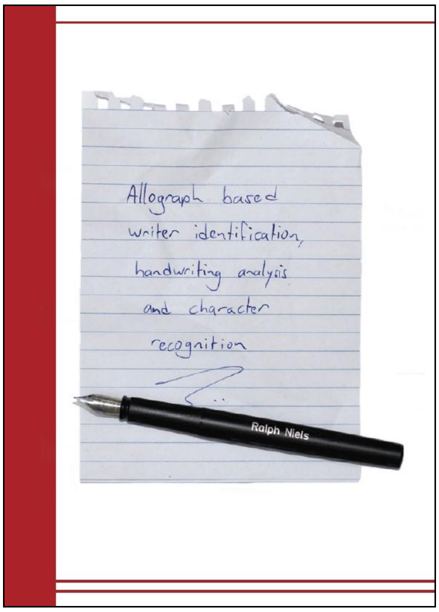
The final chapter introduces a local, motif-based approach to the retrieval of folk song melodies using audio recordings rather than the symbolic transcriptions. The melodies are segmented by detecting breath and pauses. Using the resulting audio segments, it proves possible for most of the tune families in the Annotated Corpus to find a characteristic segment that has approximate local matches in other melodies from the same tune family. Retrieval results using these musically meaningful segments are better than using fixed-length segments. The overall retrieval results are not good enough to use in an end user retrieval system yet. However, there is much room for improvement.

This thesis contributes both to Folk Song Research and Music Information Retrieval by incorporating musical knowledge in computational models. The process of developing such models of similarity relations between folk song melodies leads to better understanding of melodic similarity and, thus, of the concept of tune family, which is relevant for Folk Song Research. The models that have been developed, have successfully been used in the retrieval experiments. From the research that is presented in this thesis, it is clear that computational methods have a rich potential for the study of music; not as a replacement of 'traditional' methods, but as an extension of the research methods that are available to the musicologist.

Allograph Based Writer Identification, Handwriting Analysis and Character Recognition

Ph.D. thesis abstract
Ralph Niels

Promotores: Prof.dr. L.R.B. Schomaker, Prof.dr.ir. P.W.M. Desain
Copromotor: Dr. L.G. Vuurpijl
Date of defense: October 4, 2010



In this thesis, techniques and features are described that were developed for the automatic comparison of handwritten characters by first matching them to prototypical character shapes (allographs). These techniques and features were evaluated in experiments simulating different real-world applications.

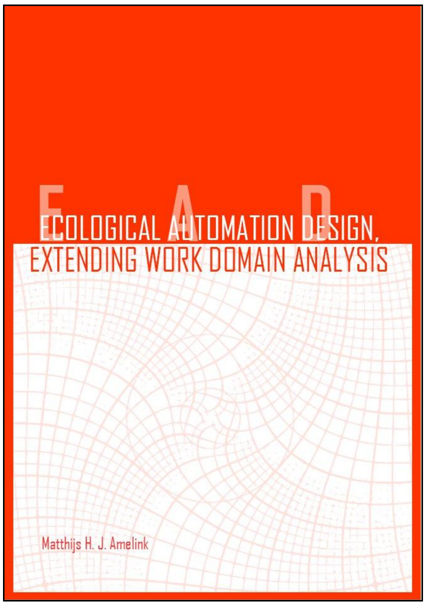
The majority of the experiments regard forensic writer identification, where the objective is to find the writer of a piece of handwriting by comparing it to a large set of handwritten documents of which the writer is already known. The assumption is that if two documents contain many similar allographs, they may have been produced by the same writer. In the experiment described, it is demonstrated that using the techniques and features, it is indeed possible to match the correct writer with a piece of unknown handwriting.

Other experiments were performed to evaluate the usefulness of the techniques and features for the classification of hand-drawn symbols and characters in different scripts (where the objective is not to find out who produced the writing, but what it represents) and the analysis of children's handwriting to diagnose Developmental Coordination Disorder.

Ecological Automation Design, Extending Work Domain Analysis

Ph.D. thesis abstract
Matthijs Amelink

Promotores: Prof.dr.ir. M. Mulder, Prof.dr.ir. B. Mulder
Copromotor: Dr.ir. R. van Paassen
Date of defense: October 18, 2010



In high-risk domains like aviation, medicine and nuclear power plant control, automation has enabled new capabilities, increased the economy of operation and has greatly contributed to safety. However, automation increases the number of couplings in a system, which can inadvertently lead to more complexity from the perspective of the operator. The automation of a system transforms the work domain of the human operator, and his role changes from controlling the core processes to managing the automated processes. The complexity of the automation and the lack of proper support can make the control task's overall difficulty larger than it needs to be, restricting safety, productivity, and efficiency.

To address and limit the automation introduced complexity in the operator's work domain, and to find representations to support him, the *ecological approach to automation design* was taken. The ecological approach focuses on the relationship between the human operator and his work domain including the system he is controlling. The main research goals were to find how the ecological approach could be used to help limit the automation introduced complexity, and how the ecological approach could be used to support the human operator in controlling automated processes.

The formulation of Ecological Automation Design (EAD) was based on the Ecological Interface Design (EID) paradigm. One of the main underlying questions asked about the interface between the work domain and the human operator is: "how to represent work domain complexity?". The interface design paradigm was transformed into an automation design paradigm by first separating the automation component from the work domain and asking the same underlying question about the interfaces between the work domain, the human operator, and the automation. Then, the conceptual *shared domain representation* was defined to visualize that the apparent complexity of the system could be reduced when both the human operator and the automation view the same representation of constraints that the work domain imposes on control. As part of the ecological approach, Work Domain Analysis (WDA) was used to analyze and represent the constraints in a work domain. However, WDA is not yet fully developed and suffers from some methodological and conceptual issues. The research therefore focused on the further development and extension of WDA to include the representation of automated processes. Four case studies were conducted, and each case study generated new insights into the application of and extension of WDA.

In the first case study, EID was applied to the design of the Energy Augmented Tunnel In the Sky display. This display was designed to aid a pilot to fly the approach to landing by presenting energy management information. The WDA revealed the significance of the energy coupling between vertical flight path and speed control as an intermediate control goal. Based on the analysis, a creative design process resulted in a novel display that has the energy representations fully, and graphically integrated in the tunnel in the sky display. A preliminary evaluation indicated that the additional energy management information shown in relation to the control actions and control goals helped pilots to fly the approaches. The display is not expected to give a performance increase but to change the way in which pilots control the throttle and elevator to fly approaches.

The second case study was the analysis of the already existing Total Energy Control System (TECS). TECS is an unconventional automated flight control system that was based on the same energy management constraints as that were represented in the energy augmented display of the first case study. The design of TECS was mapped onto the abstraction hierarchy to represent the energy management principles as part of the whole automated system. The analysis and useful representation of TECS using the abstraction hierarchy was not straightforward. It involved a search for the interpretation of the levels of the abstraction hierarchy and the use of the means-ends relationship in conjunction with the aggregation relationship. The resulting WDA showed that the abstraction hierarchy could be used to map out the reasons for TECS's design features. Many constraints were represented in the same space, which cluttered the energy management principles. The focus was put on the energy management principles through selective aggregation of the represented functions, but other design principles were omitted. To provide a complete representation of the system but without the clutter, the levels of control sophistication were introduced to represented nested control problems separately. At each level of control sophistication the abstraction hierarchy was applied, resulting in the Abstraction-Sophistication Analysis (ASA).

In the third case study, the ASA framework was used to guide the design of *SmartUAV*. SmartUAV is a newly designed mini–UAV system that is capable of controlling multiple small UAVs from a laptop computer. By designing and developing SmartUAV we gained hands-on experience with how WDA, and especially the ASA, helped to keep track of and deal with the automation introduced constraints in the design phase. The levels of control sophistication were used from the beginning to separate the different control problems in the domain. They ranged from flying the platform to the achievement of missions. Starting at the lowest level of control sophistication, each higher level allowed the designer to include a larger part of the complete work domain incrementally, and to focus on more sophisticated control of the UAV. Furthermore, the ASA supported the visualization of how automation transformed the work domain, thus how automated functionalities that were created at lower levels of control sophistication affected the (automated) functions at higher levels of control sophistication. This study showed that the ASA could span a much larger problem space than the original WDA through the nesting of abstraction hierarchies. The ASA provided a systematic way to address the abstraction of the control problems (levels of control sophistication) and the abstraction of functions per control problem (abstraction hierarchy).

The fourth case study dealt with the analysis of a subset of a well-structured domain that lacks automation; sailboat racing. This study generated a clearer view on the nested structure that is inherent in a work domain, as opposed to the nested structure of the automation as found in TECS and SmartUAV. The nested structure inherent to this work domain was found to be the result of how sailboat racing has evolved over time, based on the capabilities of equipment, human performance and the racing rules. Due to the lack of automation, it became clear that human performance is in fact part of the work domain, in contrast to the original formulations of WDA. The crew's performance formed the basis for achieving the more sophisticated control of boat speed, tactics and strategy, thus was essential in the analysis. It was shown that the performance of the human crew could be represented in the ASA at a level of control sophistication, while this could not be supported in a non-nested WDA based on a single abstraction hierarchy.

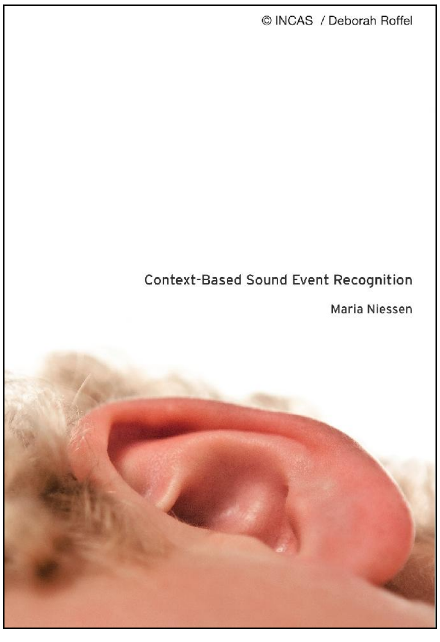
The four case studies exemplified WDA and led to its extension with a structure to explicitly nest abstraction hierarchies that map out different control problems: the ASA. Through generating the analyses, extensive modeling experience with the abstraction hierarchy was obtained, reducing its ambiguity and potential methodological and conceptual problems. We found that the abstraction hierarchy could be used to model the structure of the knowledge about a work domain but could not model the knowledge itself. Therefore, the abstraction hierarchy is a framework for structuring knowledge, linking different representations of a control problem, and explaining the reasons for design features of a system.

The abstraction hierarchy addressed the abstraction of elements belonging to a control problem, and the levels of control sophistication addressed the abstraction of the control problem itself. Representations in the ASA framework ranged from physical at the lower levels of control sophistication to non-physical at the higher levels of control sophistication. It allowed the structuring of, for example, the sailboat racing rules at the higher levels, and the law of conservation of energy at the lower levels. Although the application of the ASA did not inherently reduce the complexity of the design of SmartUAV, it enabled us to better understand the elements of the work domain that contribute to complexity of the system prior to and during its design. The extension of work domain analysis with the levels of control sophistication has led to a richer representation of the studied work domains than a single abstraction hierarchy or the abstraction-decomposition space.

Context-Based Sound Event Recognition

Ph.D. thesis abstract
Maria Niessen

Promotor: Prof.dr. L.R.B. Schomaker
Copromotor: Dr. T.C. Andringa
Date of defense: October 22, 2010



The research domain of automatic sound event recognition aims to describe an audio signal in terms of the sound events that compose a sonic environment. The ability to recognize events in a real-world environment requires a listener or system to separate the sound events from each other and the background. Furthermore, these separated events need to be recognized. To recognize a sound event implies that some representation of the event is already known to the receiver, and can be identified when it is encountered again. The ability of sound event recognition depends not only on the audio pattern specific to the event, but also on the semantics of the event. For example, a sound of a purring cat may seem unique, but without any other information than represented by the audio signal, it can sound like an engine as well. In this thesis we show that the task of recognizing a sound event can be alleviated with the semantics of the event, which is inferred from a model of the context in which the event occurs.

A possible strategy to provide an automatic sound recognition system with the semantics of a sound event, is to develop it for a specific application, and hence a specific type of context. For example, automatic speech recognition systems expect a speech signal as input, which is ensured by a user. Therefore, particular assumptions about the audio signal can be made, and context information in the form of grammar rules can be applied to recognize a word sequence (section 2.2). However, if a system for automatic sound event recognition is not designed for a specific application, and should work in variable and uncontrolled real-world environments, no assumptions about the environmental conditions can be made. Therefore, additional analysis of the audio signal is required. First, the sound events to be recognized have to be separated from the background, because the input signal consists of more than one type of sound (section 2.3). Second, the operating environment cannot be controlled, hence the system has to deal with transmission effects, such as reverberation (section 2.4).

In addition to handling the challenges of real-world environments with signal-driven methods, the semantics of the event and its environment are essential to recognize sound events in an unreliable or ambiguous audio signal. People have no difficulty in recognizing sound events in many different and noisy situations. Hence, we developed a model for automatic sound event recognition that is inspired by the strategies of human listeners as investigated by (psycho-)acoustics and cognitive psychology. The human percept of a sound event, referred to as an auditory event (section 3.2), has properties that a representation of a sound event in an automatic system can benefit from as well. People can generalize auditory events over different experiences, environments, and senses. Therefore, our model should store invariant representations of sound events, so that it is robust to variable environmental conditions, similar to human perception. Moreover, people benefit from information about the environmental context to recognize sound events (section 3.3). This facilitatory effect of contextual knowledge is an important design objective of our model.

Some other studies have been aimed at modeling context awareness in acoustics (section 4.1). However, in these studies the goal has been to estimate the context in itself, rather than to use context for the improvement of sound event recognition. In other research domains, such as information retrieval and handwriting recognition, context has been used to improve recognition or retrieval of objects. Often methods in these research domains use spreading activation networks, which are based on a model of human memory (Collins and Loftus, 1975). In this thesis we show that spreading activation networks can also be applied to estimate the most likely interpretation of an audio pattern. We introduce a context model in which the semantics of the events and the context are represented as nodes in the network (section 4.2). The activation of the nodes spreads through the network to determine the confidence of possible interpretations of an audio pattern (section 4.3).

The advantage of modeling context in automatic sound event recognition is demonstrated by applying an integrated system to audio recorded in real-world environments. This integrated system is a combination of a signal-driven analysis of the audio signal, which provides hypotheses of sound events, and the context model, which interprets these hypotheses. Knowledge about the environmental context is learned in a training phase. This knowledge is represented as a network of nodes, in which the nodes represent the semantics of the sound events and the different contexts. Furthermore, the connections between the nodes carry weights that indicate the probabilities that these sound events and contexts are encountered subsequently or concurrently. The values of the weights between nodes are learned from annotated training data (section 5.2). The type of context that is learned depends on the application domain and the data set. In a stable environment, information of co-occurring events can help to form expectancies of future events. For example, at a train station the beeping of a closing door is likely to be followed by a departing train. Alternatively, if the data set is recorded at qualitatively different types of locations, the estimated location can help to predict the types of sound events that may be heard. For example, birds are more commonly heard in parks than near busy roads. We explored the benefit of both types of context on sound event recognition in two experiments (sections 5.3 and 5.4). The results of these experiments show that the evaluation of contextual knowledge improves the recognition of sound events compared to an exclusively signal-driven method.

Contextual knowledge is not restricted to knowledge that can be derived from the audio signal and annotations of the audio signal. Other types of knowledge can be beneficial for sound event recognition as well. Knowledge inferred from different types of input, such as sound, image, and location, can reinforce each other to obtain an increased awareness of events in the environment. Ideally, these different informational resources can be combined in a single system. Because the nodes in the context model are not described by modality specific knowledge, they can be used for other types of information. We tested the applicability of the context model on the recognition of ambiguous visual information, in the domain of robot localization. Visual information received by a robot is often ambiguous (similar to acoustic information), because similar observations, such as (parts of) chairs or windows, can be made at distinct places in an environment. Learned knowledge about the environment, and the robot's hypothesized position in the environment, can help to disambiguate these observations. As a result, the position prediction improves compared to a signal-driven approach (chapter 6).

The Value of a Thesis

H. Jaap van den Herik
TiCC, UvT, Tilburg

The economic decrease encourages researchers to put emphasis on collecting more knowledge, so that they have many things to ponder and to deepen in times of scarcity. The value of a thesis can be seen as an investment in the future. However, it is not clear whether all ideas can be implemented and realised in the real world. So, next to the value of a thesis, we have the value of an application. In Science, we see many times a discussion in the procedure of approving a thesis. A first prerequisite is, of course, whether the thesis contains new ideas. Yet, ideas that cannot be followed by a process of realisation lose some of their value.

Currently, at many universities we see that the "Breimer" norm has been accepted. This means that a thesis should contain roughly 100,000 words. The norm is introduced to protect the promovendus, the supervisor, and the assessment committee against an overload of effort when reading (for the promovendus when writing) the thesis.

At this moment, the *constructive sciences* are a topic of study by NWO committees and KNAW committees. The prevailing question here is whether a "construction" (airplane, boat, intelligent bridge, autonomous train) can be seen as a new idea, a contribution to science and a piece of work that is worth the Doctor's title. If so, then the following question is whether the construction of a piece of art is also a contribution to cultural sciences and thus worth the Doctor's title.

In the list of announcements we see the thesis by B. Thomee titled *A Picture is Worth a Thousand Words*. A simple calculation based on the above estimations brings us the following. The "Breimer" norm may lead to dissertations containing some one hundred pictures. In the extreme case, we may thus expect a thesis of hundred pictures. If they have the quality of Rembrandt van Rijn, Jacob Maris, or Jacob Zekveld it may lead to a Doctor's title. We will see what the future has in store.

The current list of announcements is long and divers. Many titles are long (too long, in my opinion), but I also see titles with five or fewer words (well done). For all promovendi, a milestone is announced, their thesis. Such an event deserves our congratulations. We wish all defenders a bright future.

Mark van Assem (VU) (October 1, 2010). *Converting and Integrating Vocabularies for the Semantic Web*. Vrije Universiteit Amsterdam. Promotores: Prof.dr. A.Th. Schreiber (VU). Copromotor: Dr. J.R. van Ossenbruggen (VU/CWI).

Vasilios Andrikopoulos (UvT) (October 1, 2010). *A Theory and Model for the Evolution of Software Services*. Tilburg University. Promotor: Prof.dr. M. Papazoglou (UvT).

Peter van Kranenburg (UU) (October 4, 2010). *A Computational Approach to Content-Based Retrieval of Folk Song Melodies*. Utrecht University. Promotores: Prof.dr. R.C. Veltkamp (UU), Prof.dr. L.P. Grijp (UU). Copromotor: Dr. F. Wiering (UU).

Ralph Niels (RUG) (October 4, 2010). *Allograph Based Writer Identification, Handwriting Analysis and Character Recognition.* University of Groningen. Promotor: Prof.dr. L.R.B. Schomaker (RUG). Copromotor: Dr. L.G. Vuurpijl (RUN).

Mark van Staalduinen (TUD) (October 7, 2010). *Content-Based Paper Retrieval. Towards reconstruction of art history.* Delft University of Technology. Promotor: Prof.dr.ir. J. Biemond (TUD). Copromotor: Dr.ir. J.C.A. van den Lubbe (TUD).

Pieter Bellekens (TU/e) (October 7, 2010). *An Approach towards Context-Sensitive and User-Adapted Access to Heterogeneous Data Sources, Illustrated in the Television Domain.* Eindhoven University of Technology. Promotores: Prof.dr. P.M.E. De Bra (TU/e), Prof.dr.ir. G.J.P.M Houben (TUD). Copromotor: Dr. L.M. Aroyo (VU).

Matthijs Amelink (TUD) (October 18, 2010). *Ecological Interface Design, Extending Work Domain Analysis.* Delft University of Technology. Promotores: Prof.dr.ir. M. Mulder (TUD), Prof.dr.ir. B. Mulder (TUD), Copromotor: Dr.ir. R. van Paassen (TUD).

Maria Niessen (RUG) (October 22, 2010). *Context-Based Sound Event Recognition.* University of Groningen. Promotor: Prof.dr. L.R.B. Schomaker (RUG). Copromotor: Dr. T.C. Andringa (RUG).

Sybren de Kinderen (VUA) (October 25, 2010). *Needs-Driven Service Bundling in a Multi-Supplier Setting. The computational e^3-service approach.* Vrije Universiteit Amsterdam. Promotor: Prof.dr. J.M. Akkermans (VUA). Copromotor: dr. J. Gordijn (VUA).

Alin Gavril Chitu (TUD) (November 2, 2010). *Towards Robust Visual Speech Recognition.* Delft University of Technology. Promotores: Prof.dr. C.M. Jonker (TUD), Prof.drs.dr. L.J.M. Rothkrantz (TUD).

Marcel Heerink (UvA) (November 3, 2010). *Asessing Acceptance of Assistive Social Robots by Aging Adults*. University of Amsterdam. Promotor: Prof.dr. B.J. Wielinga (UvA). Copromotores: Dr.ir. B.J.A. Krose, Dr. V. Evers (UvA).

B. Thomee (UL) (November 4, 2010). *A Picture is Worth a Thousand Words.* Leiden University. Promotor: Prof.dr. J.N. Kok (UL).

Milan Lovric (EUR) (November 5, 2010). *Behavioral Finance and Agent-Based Artificial Markets.* Erasmus University Rotterdam. Promotores: Prof.dr. J. Spronk (EUR), Prof.dr.ir. U. Kaymak (EUR, TU/e).

Siska Fitrianie (TUD) (November 8, 2010). *Human Handheld-Device Interaction: An adaptive user interface*. Delft University of Technology. Promotores: Prof.dr. H. Koppelaar (TUD), Prof.drs.dr. L.J.M. Rothkrantz (TUD).

Chen Li (UT) (November 11, 2010). *Mining Process Model Variants: Challenges, Techniques, Examples***.** Twente University. Promotor: Prof.dr. R.J. Wieringa (UT). Co-promotor: Dr. A. Wombacher (UT).

J.S. de Bruin (UL) (November 23, 2010). *Service-Oriented Discovery of Knowledge*. Leiden University. Promotor: Prof.dr. J.N. Kok (UL).

Bouke Huurnink (UvA) (November 26, 2010) *Search in Audiovisual Broadcast Archives*. University of Amsterdam. Promotores: Prof.dr. M. de Rijke (UvA), Prof.dr.ir A.W.M. Smeulders (UvA).

Alia Khairia Amin (CWI) (December 8, 2010). *Understanding and Supporting Information Seeking Tasks in Multiple Sources***.** Eindhoven University of Technology. Promotor: Prof.dr. L. Hardman (CWI/TU/e). Copromotor: Dr. J.R. van Ossenbruggen (VU/CWI).

Peter-Paul van Maanen (VU) (December 9, 2010). *Adaptive Support for Human-Computer Teams: Exploring the use of cognitive models of trust and attention.* Vrije Universiteit Amsterdam. Promotor: Prof.dr. J. Treur (VU). Copromotor: Dr. T. Bosse (VU).

Edgar Meij (UVA) (December 10, 2010). *Combining Concepts and Language Models for Information Access.* University of Amsterdam. Promotor: Prof.dr. M. de Rijke (UvA).

Jahn-Takeshi Saito (UM) (December 15, 2010). *Solving Difficult Game Positions.* Maastricht University. Promotor: Prof.dr. G. Weiss (UM). Co-promotores: Dr.ir. J.W.H.M. Uiterwijk (UM), Dr. M.H.M. Winands (UM).

Vincent Pijpers (VU) (December 17, 2010). *e3-Alignment: Exploring Inter-Organizational Business-ICT Alignment.* Vrije Universiteit Amsterdam. Promotor: Prof.dr. J.M. Akkermans (VU). Copromotor: Dr. J. Gordijn (VU).

Section Editor
Richard Starmans

SIKS Basic Course "Research Methods and Methodology for IKS"

INTRODUCTION

On November 24, 25, and 26, 2010, the School for Information and Knowledge Systems (SIKS) organizes the annual three-day course "Research Methods and Methodology for IKS".

The location will be Conference center Woudschoten in Zeist.

The course will be given in English and is part of the educational Program for SIKS-Ph.D. students. Although the course is primarily intended for SIKS-Ph.D. students, other participants are not excluded. However, their number of passes will be restricted and depends on the number of SIKS-Ph.D. students taking the course.

"Research Methods and Methodology for IKS" is relevant for all SIKS-Ph.D. students (whether working in computer science or in information science). The primary goal of this hands-on course is to enable these Ph.D. students to make a good research design for their own research project. To this end, it provides an interactive training in various elements of research design, such as the conceptual design and the research planning. But the course also contains a general introduction to the philosophy of science (and particularly to the philosophy of mathematics, computer science and AI). And, it addresses such divergent topics as "the case-study method", "elementary research methodology for the empirical sciences" and "empirical methods for computer science".

"Research Methods and Methodology for IKS" is an intense and interactive course. First, all students enrolling for this course are asked to read some pre-course reading material, comprising some papers that address key problems in IKS-methodology. These papers will be sent to the participants immediately after registration. Secondly, all participants are expected to give a brief characterization of their own research project/ proposal, by answering a set of questions, formulated by the course directors, and based on the aforementioned literature. We believe that this approach results in a more efficient and effective course; it will help you to prepare yourself for the course and this will increase the value that you will get from it.

COURSE COORDINATORS

- Hans Weigand (UvT)
- Roel Wieringa (UT)
- John-Jules Meyer (UU)
- Hans Akkermans (VU)
- Richard Starmans (UU)

PROGRAM

More details on the program will be made available in due course.

REGISTRATION

In the conference center there is a limited number of places and there is interest from other groups in the topic as well. Therefore, an early registration is required. For registration you are kindly requested to fill in the registration form at the SIKS website.

Arrangement 1 includes single room, all meals, and course material.

Arrangement 2 includes only lunch, dinner and course material. So no stay in the hotel and no breakfast.

Deadline for registration for SIKS-Ph.D. students: November 1, 2010.

After that date, applications to participate will be honoured in a first-come first-serve manner. Of course, applications to participate from other interested groups are welcome already. They will receive a notification whether they can participate as soon as possible.

Information for non-SIKS-Ph.D. students: SIKS needs a confirmation from your supervisor/office that they agree with the arrangement and paying conditions.

SIKS Basic Courses
"Mathematical Methods for IKS" and "Knowledge Modeling"

INTRODUCTION

On December 7-10, 2010, the School for Information and Knowledge Systems (SIKS) organizes two basic courses "Mathematical Methods for IKS" and "Knowledge Modeling". Both courses will be given in English and are part of the obligatory Basic Course Program for SIKS-Ph.D. students. Although these courses are primarily intended for SIKS-Ph.D. students, other participants are not excluded. However, their number of passes will be restricted and depends on the number of SIKS-Ph.D. students taking the course.

Location: Landgoed Huize Bergen, Vught
Date: December 7-10, 2010

SCIENTIFIC DIRECTORS

- Prof.dr. Eric Postma (UvT), Mathematical Methods for IKS
- Prof.dr. Tom Heskes (RUN), Mathematical Methods for IKS
- Dr. Bert Bredeweg (UVA), Knowledge Modeling

PROGRAM

The program is not available yet, but may include the following topics:

Mathematical Methods for IKS

- Basic formalisms relevant to modern intelligent knowledge systems
- Automatically acquired knowledge representations
- Inductive Learning
- Bayesian Statistics
- Entropy and Information Theory
- Computer intensive techniques
- Minimum Description Length Methods

Knowledge Modeling

- Ontologies, Epistemology and Models
- Modeling with Description Logics
- Methodology for Ontology Engineering
- KADS
- OWL: Ontology Language for the Web
- Ontology Patterns, Re-use of information

REGISTRATION

In the conference center there is a limited number of places and there is interest from other groups in the topic as well. Therefore, an early registration is required.

Deadline for registration for SIKS-Ph.D. students: **November 30, 2010.** After that date, applications to participate will be honoured in a first-come first-serve manner. Of course, applications to participate from other interested groups are welcome already. They will receive a notification whether they can participate as soon as possible.

For registration you are kindly requested to fill in the registration form at the SIKS website.

Arrangement 1 includes single room, all meals, and course material.

Arrangement 2 includes two lunches, one dinner and course material. So no stay in the hotel and no breakfast.

Advertisements in the BNVKI Newsletter

Do you want to place a (job) advertisement in the Newsletter of the BNVKI?

- Whole page: € 400 for 1 issue; € 600 for 2 subsequent issues; € 900 for 6 subsequent issues.
- Half page: € 300 for 1 issue; € 450 for 2 subsequent issues; € 675 for 6 subsequent issues.

You reach an audience of AI professionals, academics and students. Your logo (with link to your company) will also be shown on the BNVKI/AIABN website during the period of advertisement.

Contact sien.moens@cs.kuleuven.be for additional information.

HOW TO SUBSCRIBE

The BNVKI-AIABN Newsletter is a direct benefit of membership of the BNVKI-AIABN: Benelux Association for Artificial Intelligence. Membership dues are € 40 for regular members; € 25 for doctoral students (AIO's); and € 20 for students. In addition, members will receive access to the electronic version of the European journal *AI Communications*. The Newsletter appears bimonthly and contains information about conferences, research projects, job opportunities, funding opportunities, etc., provided enough information is supplied. Therefore, all members are encouraged to send news and items they consider worthwhile to the editorial office of the BNVKI/AIABN Newsletter. Subscription is done by payment of the membership due to Postbank no. 3102697 in The Netherlands (IBAN: NL 74 PSTB 0003 1026 97; BIC: PSTBNL21). Specify *BNVKI/AIABN* as the recipient, and please do not forget to mention your name and address. Sending of the BNVKI/AIABN Newsletter will only commence after your payment has been received. If you wish to conclude your membership, please send a written notification to the editorial office before December 1, 2010.

COPY

The editorial board welcomes product announcements, book reviews, product reviews, overviews of AI education, AI research in business, and interviews. Contributions stating controversial opinions or otherwise stimulating discussions are highly encouraged. Please send your submission by E-mail (MS Word or text) to newsletter@maastrichtuniversity.nl.

ADVERTISING

It is possible to have your advertisement included in the BNVKI/AIABN Newsletter. For further information about pricing etc., see elsewhere in the Newsletter or contact the editorial office.

CHANGE OF ADDRESS

The BNVKI/AIABN Newsletter is sent from Maastricht. The BNVKI/AIABN board has decided that the BNVKI/AIABN membership administration takes place at the editorial office of the Newsletter. Therefore, please send address changes to:

Editorial Office BNVKI/AIABN Newsletter
Marijke Verheij
Department of Knowledge Engineering (DKE),
Maastricht University
P.O. Box 616, 6200 MD Maastricht, the Netherlands
E-mail: newsletter@maastrichtuniversity.nl
http://www.bnvki.org