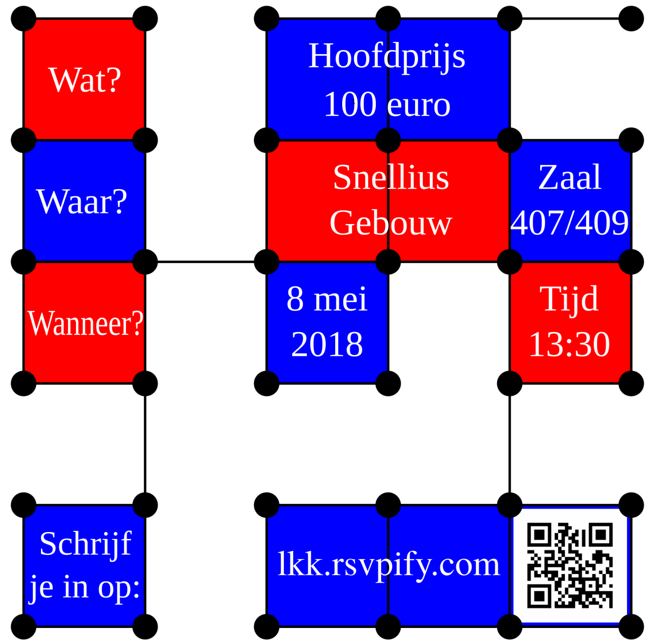
Kamertje verhuren = Dots and boxes

Walter Kosters

maandag 16 april 2018

www.liacs.leidenuniv.nl/~kosterswa/db.pdf

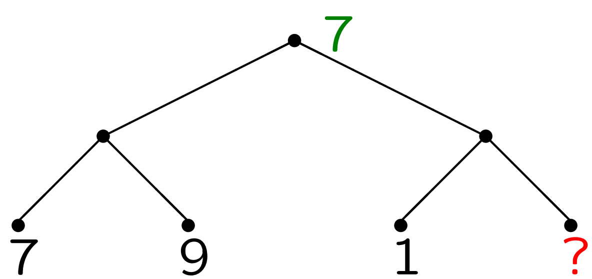
Leidsch Kampioenschap
Kamertje verhuren 2018



Organisatie:
Roland van der Veen en Jelle van der Voort,
Mathematisch Instituut Leiden.
www.rolandvdv.nl/LKK
LKKamertjeverhuren@gmail.com

Schaken

Deep Blue (met minimax/α-β) vs. Garry Kasparov, 1997

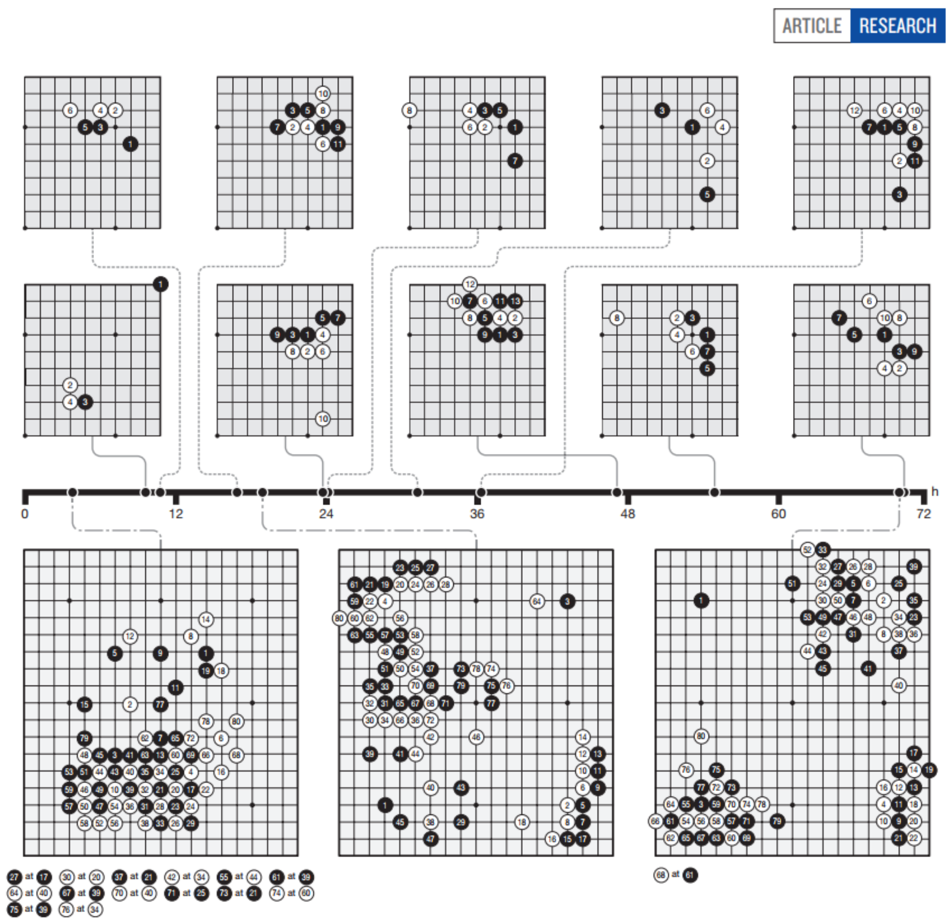


$\leftarrow$ MAX aan zet

$\leftarrow$ MIN aan zet

Dots and boxes

AlphaZero



breaking news

↓

Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm

David Silver,[1*] Thomas Hubert,[1*] Julian Schrittwieser,[1*]
Ioannis Antonoglou,[1] Matthew Lai,[1] Arthur Guez,[1] Marc Lanctot,[1]
Laurent Sifre,[1] Dharshan Kumaran,[1] Thore Graepel,[1]
Timothy Lillicrap,[1] Karen Simonyan,[1] Demis Hassabis[1]

[1]DeepMind, 6 Pancras Square, London N1C 4AG.
*These authors contributed equally to this work.

arXiv:1712.01815v1 [cs.AI] 5 Dec 2017

Abstract

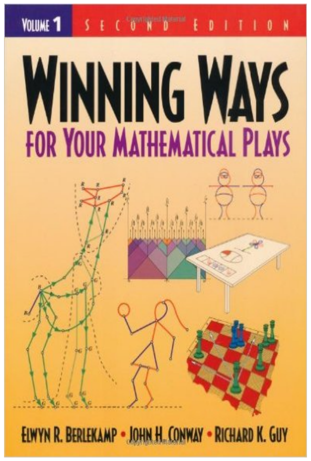
The game of chess is the most widely-studied domain in the history of artificial intelligence. The strongest programs are based on a combination of sophisticated search techniques, domain-specific adaptations, and handcrafted evaluation functions that have been refined by human experts over several decades. In contrast, the *AlphaGo Zero* program recently achieved superhuman performance in the game of Go, by *tabula rasa* reinforcement learning from games of self-play. In this paper, we generalise this approach into a single *AlphaZero* algorithm that can achieve, *tabula rasa*, superhuman performance in many challenging domains. Starting from random play, and given no domain knowledge except the game rules, *AlphaZero* achieved within 24 hours a superhuman level of play in the games of chess and shogi (Japanese chess) as well as Go, and convincingly defeated a world-champion program in each case.

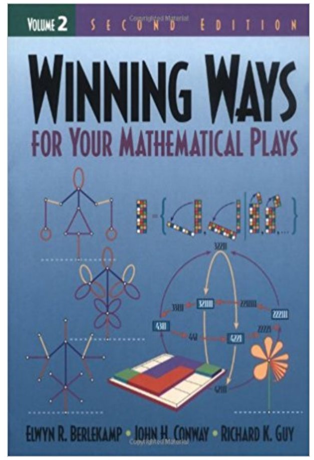
The study of computer chess is as old as computer science itself. Babbage, Turing, Shannon, and von Neumann devised hardware, algorithms and theory to analyse and play the game of chess. Chess subsequently became the grand challenge task for a generation of artificial intelligence researchers, culminating in high-performance computer chess programs that perform at superhuman level (*9, 13*). However, these systems are highly tuned to their domain, and cannot be generalised to other problems without significant human effort.

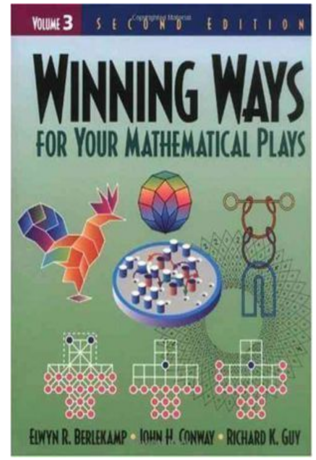
A long-standing ambition of artificial intelligence has been to create programs that can instead learn for themselves from first principles (*26*). Recently, the *AlphaGo Zero* algorithm achieved superhuman performance in the game of Go, by representing Go knowledge using deep convolutional neural networks (*22, 28*), trained solely by reinforcement learning from games of self-play (*29*). In this paper, we apply a similar but fully generic algorithm, which we

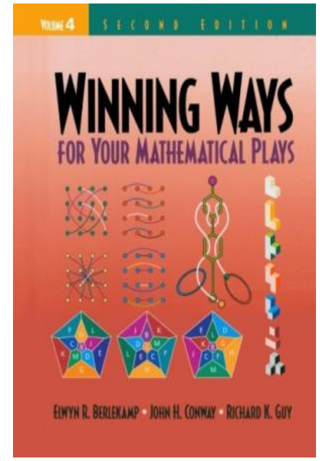
1

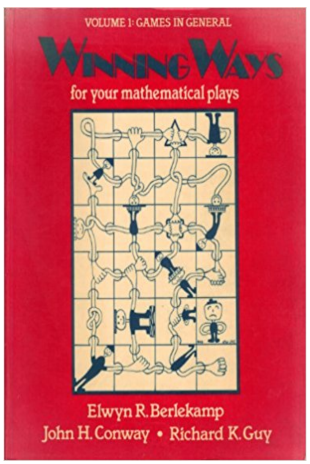
Silver et al., Nature 550 (2017) 354–359: Deep learning, MCTS, TPU's, . . .

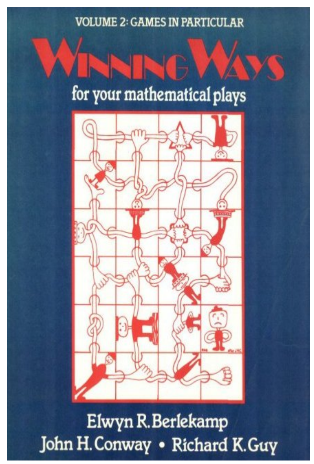


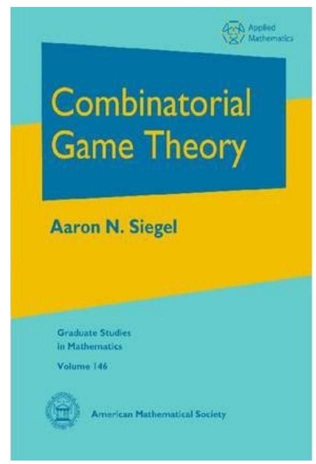


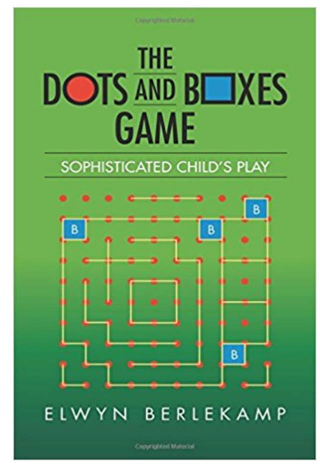








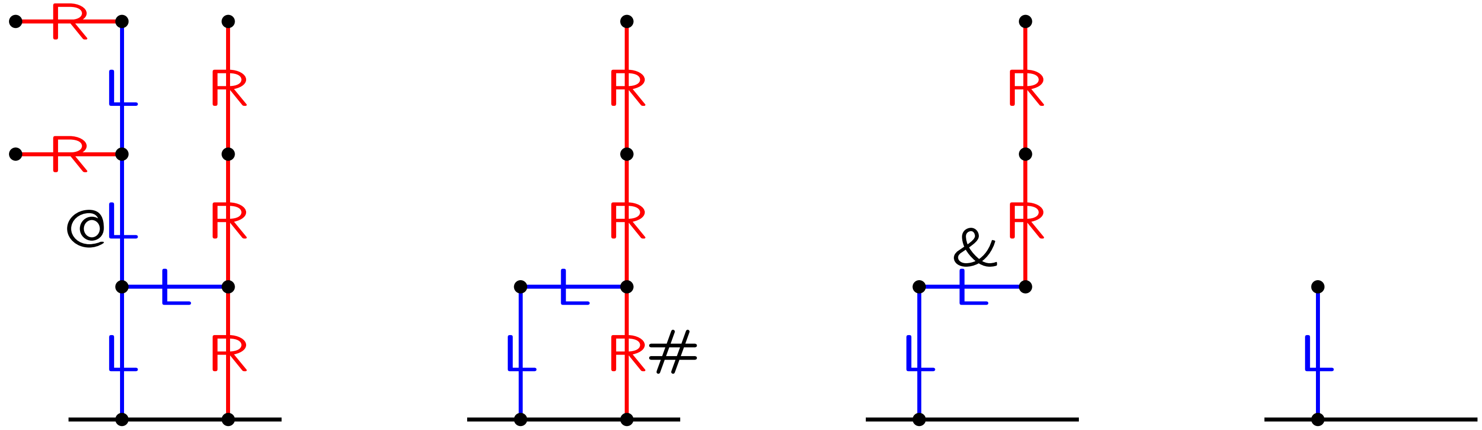




Combinatorial Game Theory

Berlekamp, Conway & Guy: Winning ways

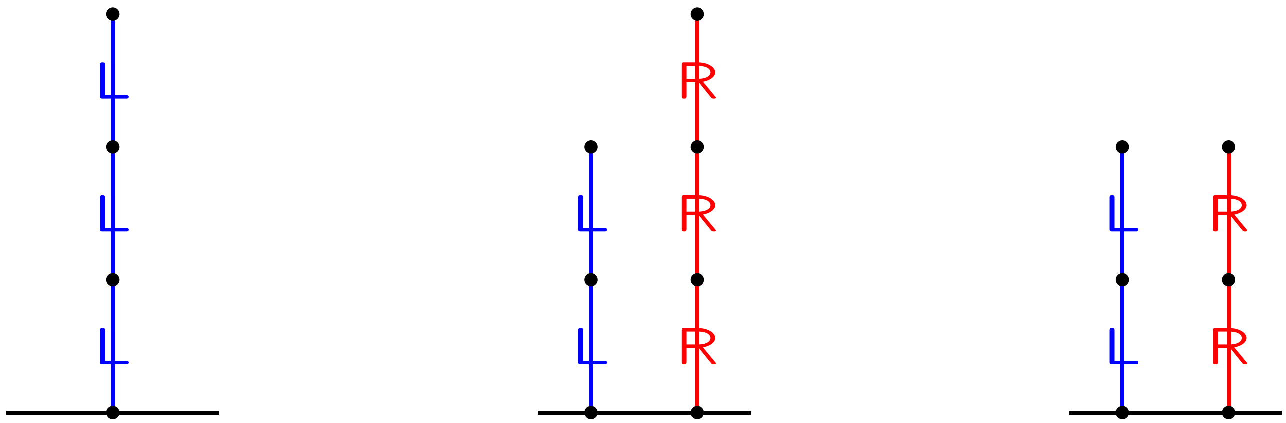
Bij het spel Hackenbush verwijderen Links en Rechts om de beurt respectievelijk een bLauw of een Rood streepje, waarna alle streepjes die niet meer met de grond verbonden zijn ook worden verwijderd. *Wie niet kan, heeft verloren!*



Links kiest @, Rechts kiest #, Links kiest & en wint

Overigens: hier kan Rechts altijd winnen, wie er ook begint!

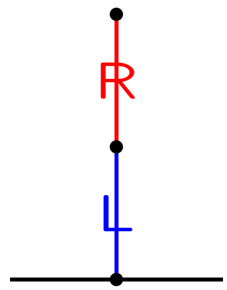
Wat is bij Hackenbush de waarde van een positie?



waarde 3 waarde $2-3=-1$ waarde $2-2=0$

Als de waarde positief (>0) is, *kan* Links altijd winnen (wie er ook begint; in het linker voorbeeld met voorsprong 3), als de waarde negatief (<0) is *kan* Rechts altijd winnen, en als de waarde 0 is verliest de beginspeler.

Maar wat is de waarde van deze positie?



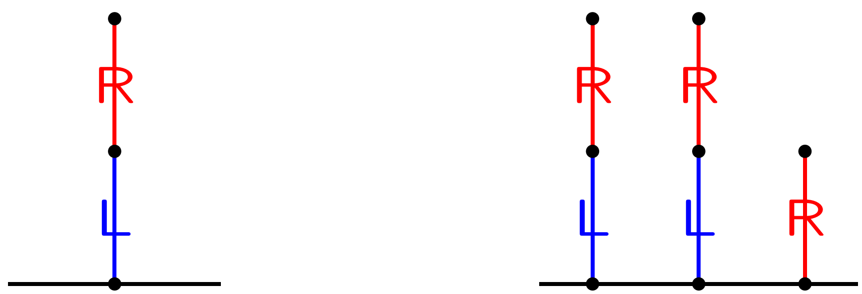
Als Links begint, wint hij meteen; als Rechts begint, kan Links nog een keer, en wint hij ook. Dus Links wint altijd. De waarde is daarom > 0.

Vraag: is de waarde gelijk aan 1?

Als de waarde links 1 zou zijn, zou de waarde van de rechter positie $1 + (-1) = 0$ moeten zijn, en zou de beginspeler hier moeten verliezen. Is dat zo?

Nee: het is zo dat als Links begint, Links verliest, en als Rechts begint Rechts ook kan winnen. Dus Rechts wint altijd (= kan altijd winnen), en daarom is de rechter positie < 0, en de linker tussen 0 en 1.

We noteren de waarde van de linker positie met $\{\,0 \mid 1\,\}$.



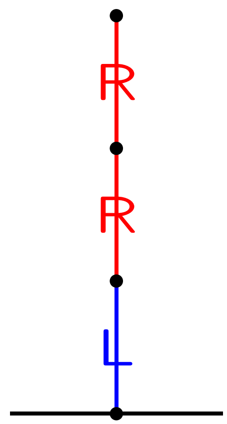
We merken op dat de rechter positie wél waarde 0 heeft: de beginspeler verliest. En dus geldt:

$$\{\,0 \mid 1\,\} + \{\,0 \mid 1\,\} + (-1) = 0,$$

en blijkbaar $\{\,0 \mid 1\,\} = 1/2$.

En: $0 = \{\,\mid\,\}$: beginspeler verliest (hij/zij kan niets).

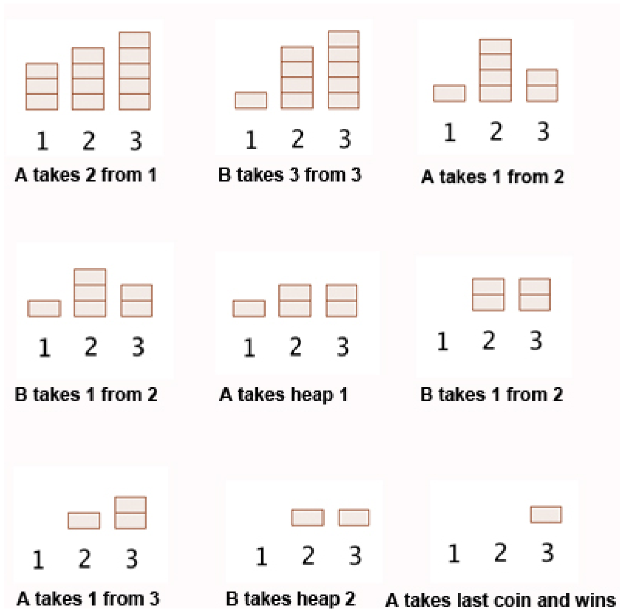
We noteren de waarde van een positie waarin Links kan spelen naar (waardes van) posities uit de verzameling L en Rechts kan spelen naar (waardes van) posities uit de verzameling R met $\{\ L \mid R\ \}$. Een voorbeeld:



De waarde is hier $\{\ 0 \mid \frac{1}{2}, 1\ \} = \frac{1}{4}$.

De waarde blijkt altijd het "eenvoudigste" getal dat tussen linker en rechter verzameling in zit.

Bij het spel Nim is er een aantal stapels met munten. Je moet van een stapel naar keuze een zelfgekozen aantal munten ($\neq 0$) pakken. Wie niet kan heeft verloren.



Het spel is impartial: beide spelers hebben dezelfde zetten. Je hebt ook een misère versie: wie niet kan heeft gewonnen.

Voor Nim geldt Bouton's analyse uit 1901.

We definiëren de nim-som $x \oplus y$ van twee gehele positieve getallen x en y als de bitsgewijze XOR van de binaire representaties van x en y: optellen zonder carry. Met twee stapels van gelijke grootte verliest de beginspeler ($x \oplus x = 0$): gebruik de "spiegel-strategie".

Een Nim-spel met stapels ter grootte $a_1, a_2, \ldots, a_k$ is verloren voor de beginspeler precies als $a_1 \oplus a_2 \oplus \ldots \oplus a_k = 0$. En deze som is de Sprague-Grundy waarde.

We noteren een spel met waarde m als $*m$. En $*1 = *$. Als $m \neq 0$ wint dus de beginspeler.

De Stelling van Sprague-Grundy zegt ruwweg dat elk impartial spel een Nim-spel is.

Als voorbeeld: stapels met grootte 29, 21 en 11.

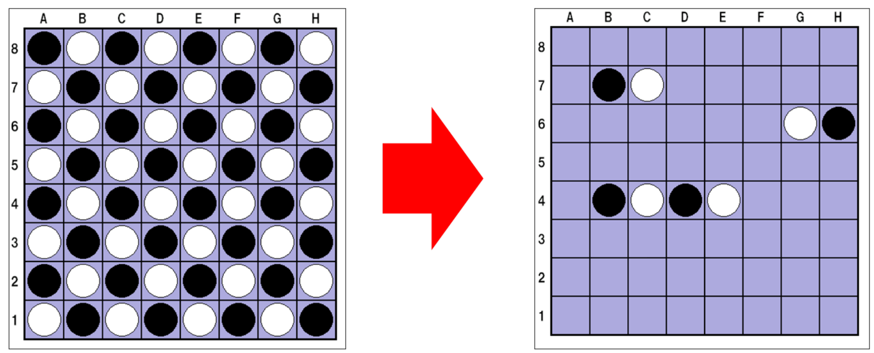
We berekenen $29 \oplus 21 \oplus 11 = 3$, via

```
11101
10101
 1011
-----
00011
```

. . . gewonnen voor de beginspeler, met unieke zet $11 \to 8$.

We spelen nu Clobber, op een m bij n bord, met witte (Rechts) en zwarte (Links) stenen. Een steen kan een direct horizontaal/verticaal aangrenzende steen van de andere kleur pakken = clobberen. Wie niet kan, verliest.

Enkele voorbeelden:



●○ $= \{\, 0 \mid 0 \,\} = *$

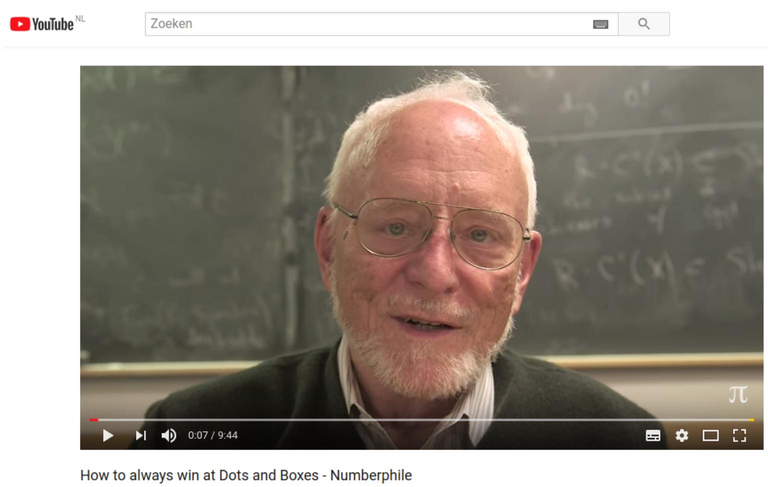
●●○ $= \{\, 0 \mid * \,\} = \uparrow > 0$

●○●○●○ $= 0$

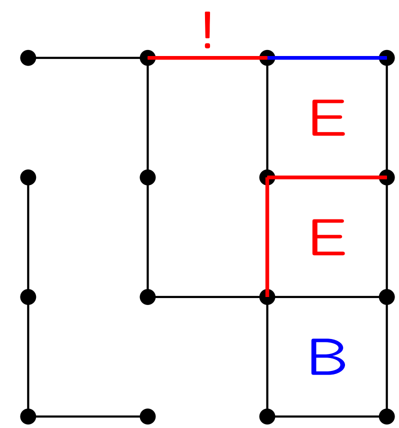
$* + * = 0$ en . . .

AI

En tot slot Kamertje verhuren = Dots and boxes. Als je aan de beurt bent, moet je een muurtje tekenen. Als je een of twee vierkanten volmaakt, mag = moet je weer. Wie de meeste vierkanten heeft wint.



Elwyn Berlekamp: [YouTube]



Nu moet B een “doublecross” doen

“Try to get control and keep it”

Als je begint: probeer aantal beginstippen plus aantal long chains (≥ 3 vierkanten) even te maken, en anders: oneven.

Bewijs: het aantal beginstippen plus het aantal “double-crosses” is gelijk aan het totaal aantal beurten. Zie boek!

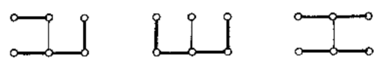


Figure 8 (a). Hard-Hearted Handouts.

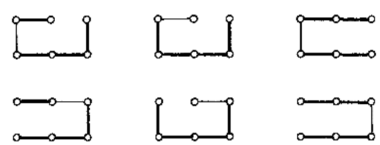


Figure 8 (b). Half-Hearted Handouts.

When you think you are winning, but are forced to give away a pair of boxes, you should always make a hard-hearted handout, so that your opponent has no option but to accept. If you use a **half-hearted** one (Fig. 8(b)) he might reply with a double-dealing move and regain control. But if you're losing, you might try a half-hearted handout on the Enough Rope Principle (Chapter 1 Extras). Officially this is a bad move, since your opponent, if he has any sense, will grab both squares. But boys by billions, being bemused by Bertha's brilliance, blindly blunder both boxes back.

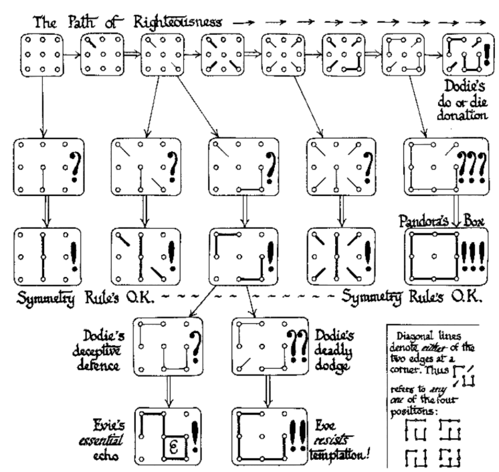
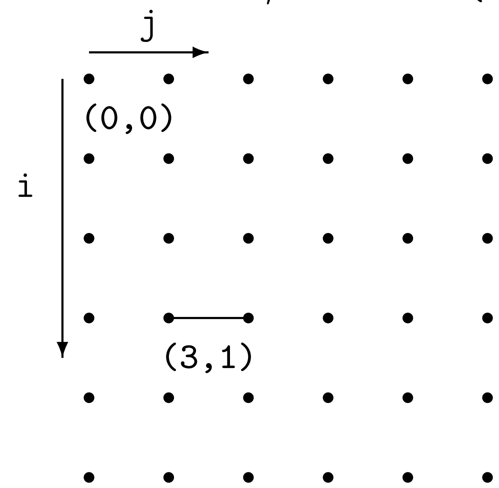


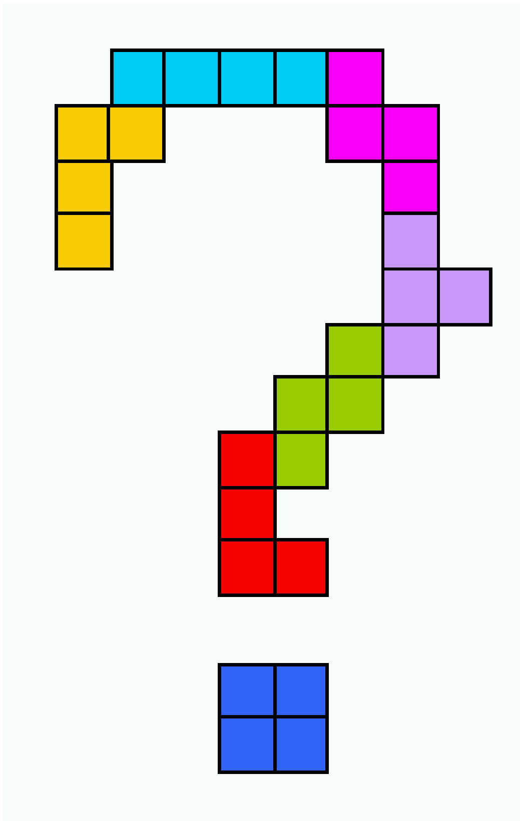
Figure 9. Evie Envisaging every Eventuality.

6×6, `adminhor (3,1)`

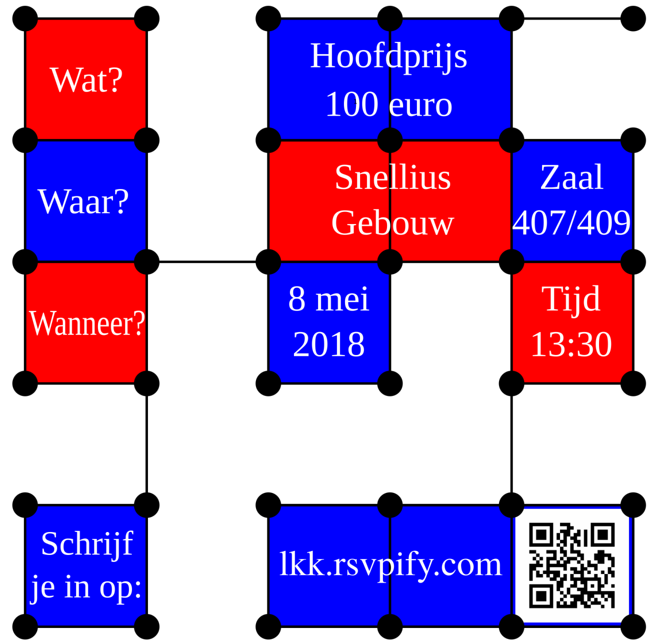



```
class DB {
    int h, w; bool turn; // h,w < THEMAX = 10
    char player1, player2;
    bool hor[THEMAX][THEMAX];
    bool ver[THEMAX][THEMAX];
    char occupation[THEMAX][THEMAX];
    int occ1, occ2; int moves;

    void print ( );
    void adminhor (int i, int j); //do horizontal move *above* (i,j)
    void adminver (int i, int j); //do vertical move *left of* (i,j)
    void domove (int mymove); //do i-th move from 0,1,...,moves-1
    void playrandommove ( );
    void playeagermove ( );
    void playrandomgame (int depth);
    void playeagergame (int depth); //play full game <= depth moves
    void playgame ( );
    ...
};//DB
```



Leidsch Kampioenschap Kamertje verhuren 2018



Organisatie:
Roland van der Veen en Jelle van der Voort,
Mathematisch Instituut Leiden.
www.rolandvdv.nl/LKK
LKKamertjeverhuren@gmail.com