Science as Art

Jan van Eijck

CWI

jve@cwi.nl

SNB Lecture, 16 september 2009

Abstract

Creating How do I create something new. How do I think up something?

Research How do I check my thoughts? Reality checking? Checking by reasoning? Other methods?

Communication How do I communicate the results of my work to others?

What is: ‘creating something’?

- Causing a jump from *nothingness* to *being*.
- The description conveys that there is *tension*. something is not there, and I want it to be there. I struggle to force it into existence.
- Or sometimes: there is something I cannot do, and I want to learn how to do it.
- Learning to create starts with getting to love the feeling of tension caused by the lack of something. If everything is as it should be, there is no need to create anything.

The creative process

- What do I want to do/construct/create (small things, large things)?
- Which direction do I choose?
- What is important? What is less important? You can create *clarity* for yourself by ranking things in order of importance.
- Where am I? Where do I want to go? What is the movement? Is there enough momentum? How can I create more momentum? Momentum is the crucial thing.

The creative process and creativity

- What are the characteristics of creative people?
- Can creativity be learned or acquired?
- How do I become a creative thinker or doer?
- Is there a principled difference between creating in the small and creating in the large?
- Is there a role for creativity in Software Engineering or System Administration?
- How does creativity relate to productivity?

Literature

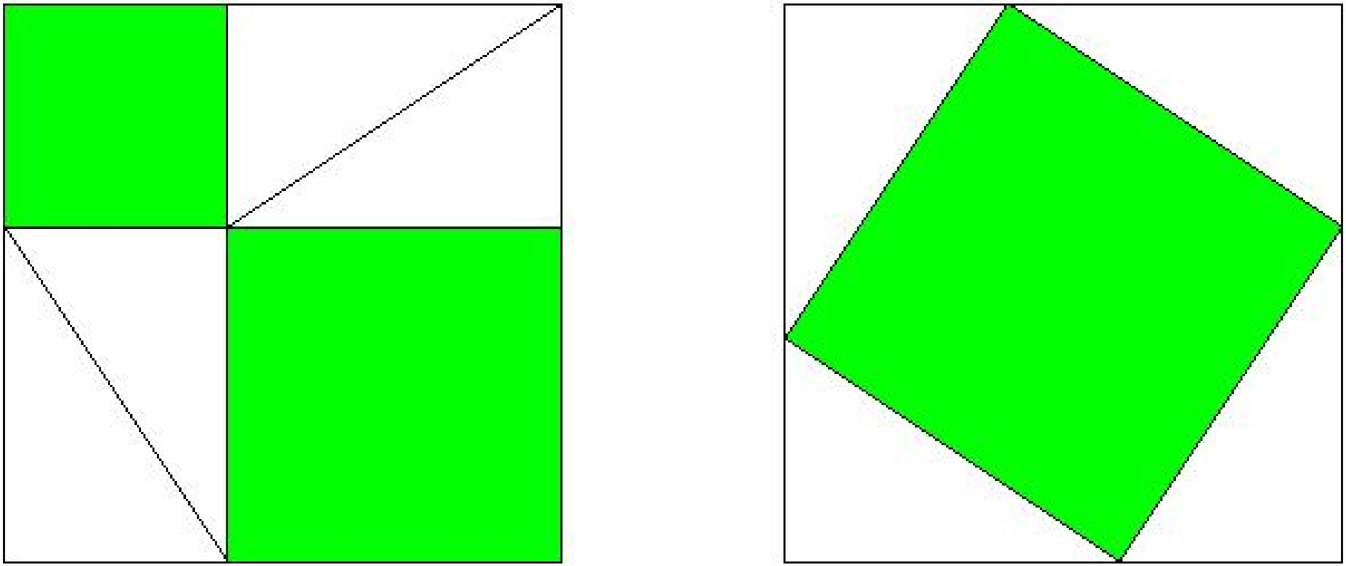
about creating Robert Fritz, *Creating*, Ballantine Books, 1991 [5].

about creative writing Julia Cameron, *The Artist's Way*, G.P. Putnam Books, 1992 [**?**]

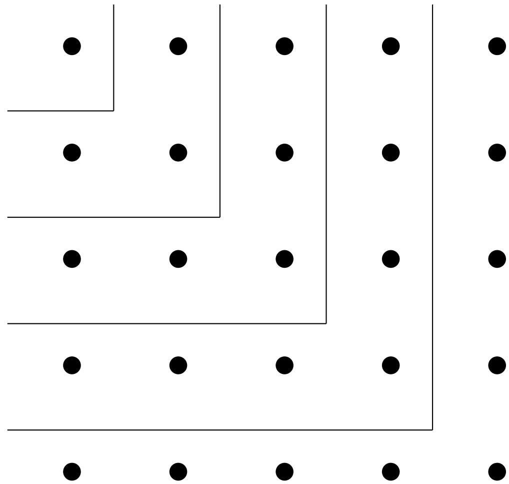
about problem solving J.D. Bransford and B.S. Stein, *The Ideal Problem Solver: A Guide or Improving Thinking, Learning, and Creativity*, W.H. Freeman, 1984. A Scientific American Special [2].

about realizing optimal learning experience Mihaly Csikszentmihaly, *Flow: the psychology of optimal experience*, Harper and Row, 1990 [4].

Pythagoras ‘direct’

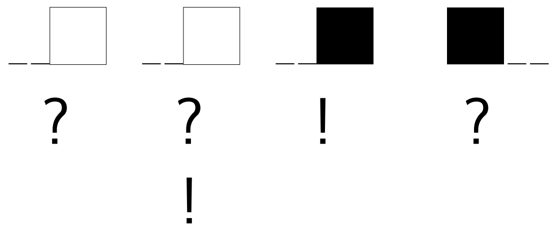


The Gnomon: the som of the first n odd numbers

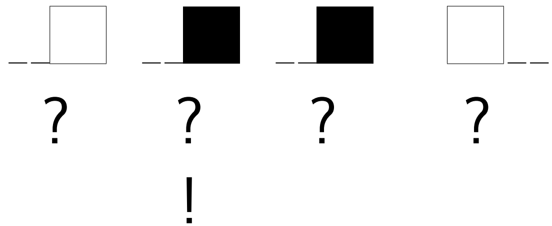


Example: The Riddle of the Caps

“Er zijn twee witte en twee zwarte petten.”



"Er zijn twee witte en twee zwarte petten."



Example: Muddy Children

"Minstens een van jullie heeft een stip op het voorhoofd."

a	b	c
○	○	●
?	?	!
!	!	

"Minstens een van jullie heeft een stip op het voorhoofd."

a	b	c	d
○	○	●	●
?	?	?	?
?	?	!	!
!	!		

“Minstens een van jullie heeft een stip op het voorhoofd.”

a	b	c	d
○	●	●	●
?	?	?	?
?	?	?	?
?	!	!	!
!			

Example: pebble puzzle

In een vaas zitten 35 witte en 35 zwarte steentjes. Je gaat, zolang dat mogelijk is, als volgt te werk. Je haalt steeds twee steentjes uit de vaas.

- Als ze dezelfde kleur hebben stop je een zwart steentje terug in de vaas (er zijn voldoende extra zwarte steentjes),
- als ze verschillende kleur hebben stop je het witte steentje terug in de vaas.

Omdat er bij elke stap een steentje verwijderd wordt is er na 69 stappen nog maar één steentje over. Welke kleur heeft dat steentje? Waarom?

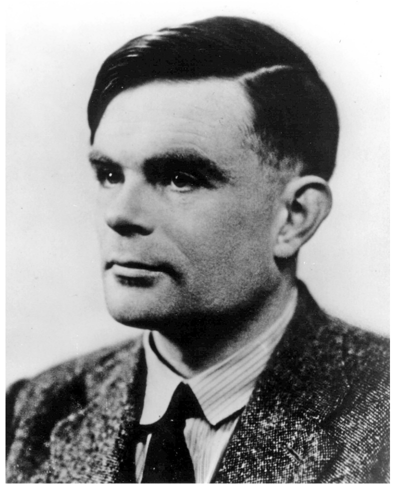
Example: MU puzzle

Kun je in een eindig aantal stappen het 'woord' MI transformeren in het woord MU, door uitsluitend gebruik te maken van de volgende regels:

1. Als een rijtje in een I eindigt mag je er een U achter schrijven (bij voorbeeld: van MI naar MIU).
2. Een rijtje dat volgt op M mag worden verdubbeld: je mag Mx veranderen in Mxx (bij voorbeeld: van MIU naar MIUIU).
3. Als III voorkomt in een rijtje, dan mag je dit vervangen door U (bij voorbeeld: van MUIIIU naar MUUU).
4. Als UU voorkomt in een rijtje, dan mag je dit weghalen (bij voorbeeld: van MUUU naar MU).

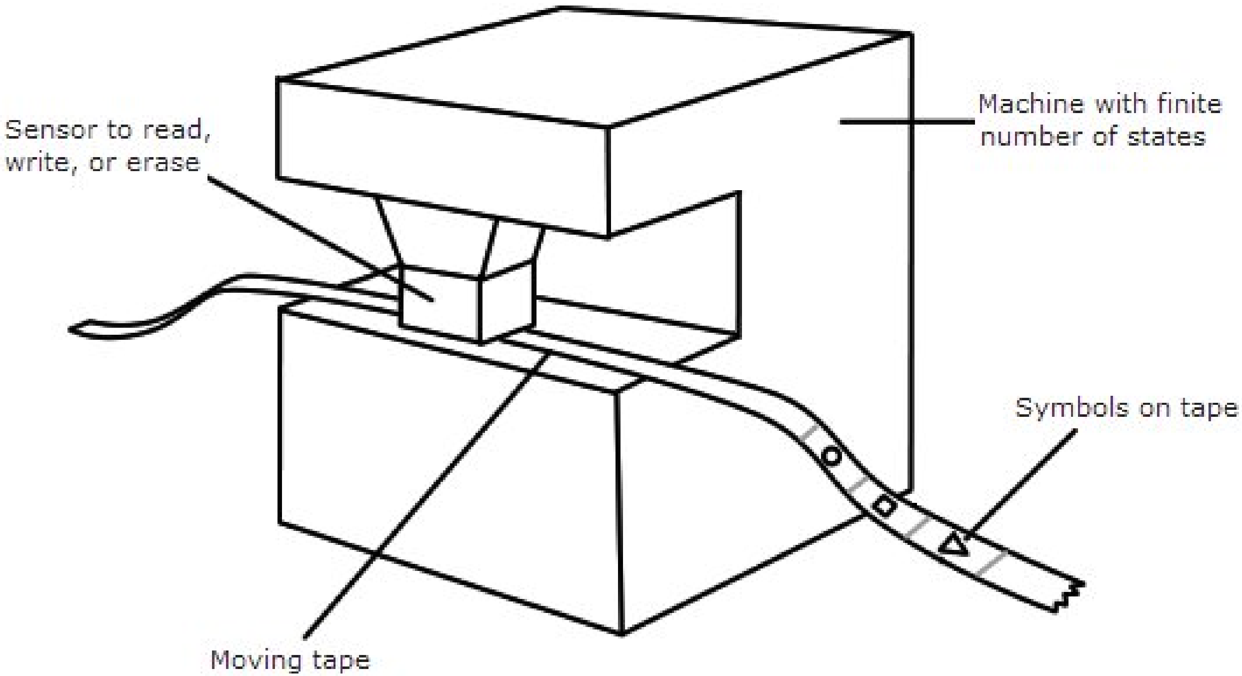
Als je denkt dat het kan: laat zien hoe. Als je denkt dat het *niet* kan: laat zien waarom niet.

Alan Turing's Insight

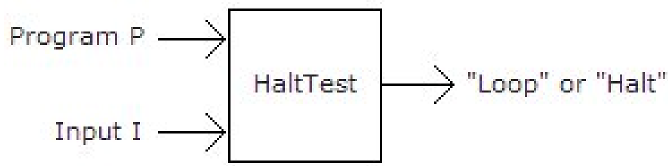


A language that allows the specification of 'universal procedures' cannot be decidable.

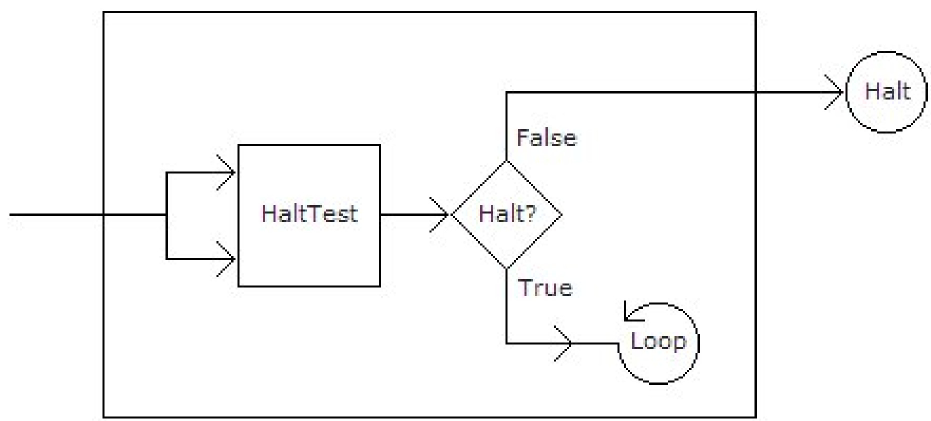
But first order predicate logic *is* such a language . . .

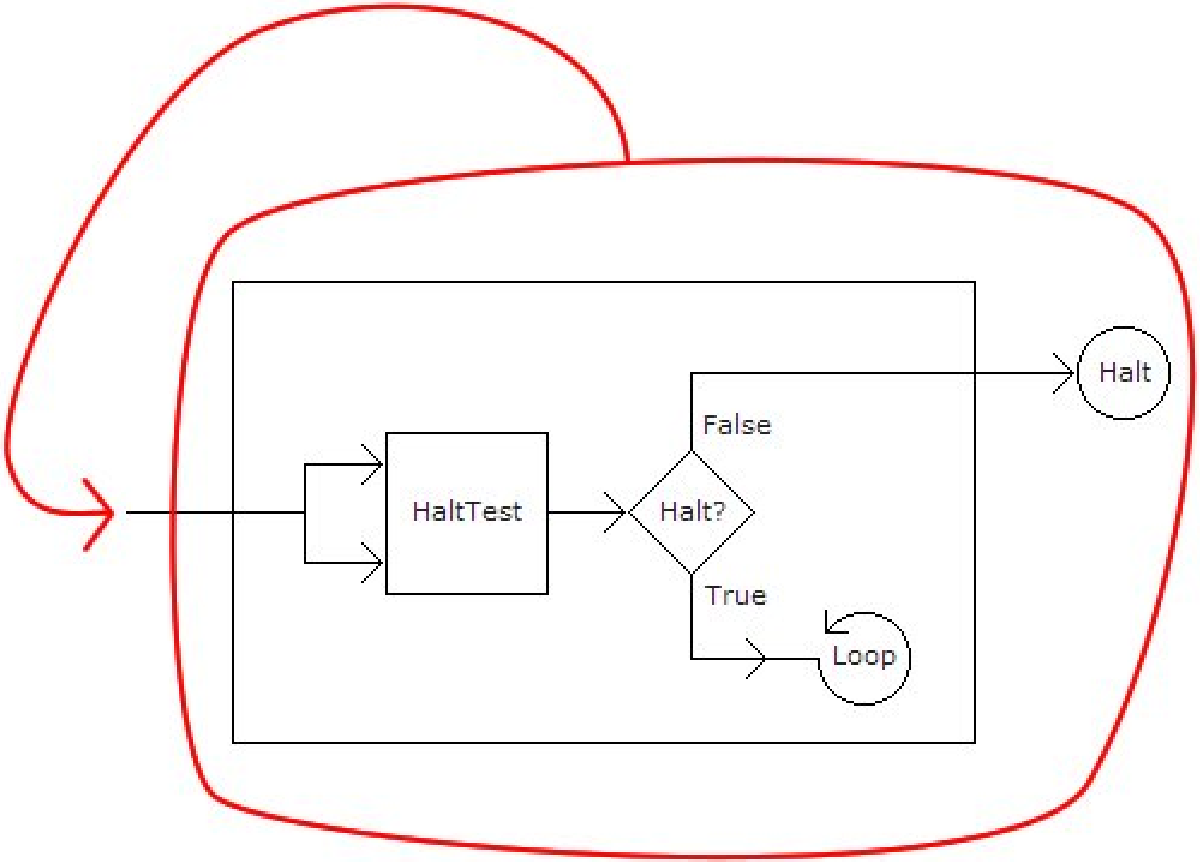


A Turing Machine


Program P
Input I
HaltTest
"Loop" or "Halt"


Program P
HaltTest
"Loop" or "Halt"


HaltTest
Halt?
False
True
Halt
Loop


Halt
False
HaltTest
Halt?
True
Loop

Proof of Undecidability of First Order Logic

- The formal proof of the undecidability of first order logic consists of
  - A **very general** definition of *computational procedures*.
  - A demonstration of the fact that such computational procedures can be expressed in first order logic.
  - A demonstration of the fact that the halting problem for computational procedures is undecidable (see the above sketch).
  - A formulation of the halting problem in first order logic.
- This formal proof was provided by Alan Turing in [8].

This Ruby will always halt

```
print "Hoe heet je? "
input = STDIN.gets.chop
if input == 'Alan Turing'
    print "Fantastisch\n"
else
    print "Jammer\n"
end
```

This program can get into an infinite loop

```
print "Hoe heet je? "
input = STDIN.gets.chop
if input == 'Alan Turing'
    print "Fantastisch\n"
else
   while not false
     print "Jammer\n"
   end
end
```

Next: poetic proof of the undecidability of the halting problem.

What does this program do?

```
main = putStrLn (s ++ show s)
  where s = "main = putStrLn (s ++ show s) \n  where s = "
```

Experiments with the Computer

```
Main> run 0
Program error: argument < 1
Main> run 1
[1]
Main> run 5
[5,16,8,4,2,1]
Main> run 7
[7,22,11,34,17,52,26,13,40,20,10,5,16,8,4,2,1]
Main> run 4
[4,2,1]
Main> run 21
[21,64,32,16,8,4,2,1]
```

What does this program do?

Haskell version:

```haskell
run :: Integer -> [Integer]
run n | n < 1 = error "argument < 1"
      | n == 1 = [1]
      | even n = n: run (div n 2)
      | odd n  = n: run (3*n+1)
```

Ruby version:

```ruby
def run k
  raise "argument < 1" if k < 1
  return [k] if k == 1
  return [k] + run(k/2) if k.modulo(2) == 0
  return [k] + run(3*k + 1)
end
```

Questions

- How can we test this program?
- Is there an experimental method for checking that the program always terminates?
- What does this have to do with the halting problem?

Simulation

> With this now famous, simple, desktop demonstration Feynman showed that responsibility for the disaster lay largely with NASA managers who ignored their engineers' warnings to abort a launch because of the unusually cold temperature that morning, twenty-nine degrees Fahrenheit (the lowest temperature at any previous lauch had been fifty-three degrees Fahrenheit). [7]

(29°F = -2°C, 53°F = 11°C.)

Evaluation

- When are the conclusions from a simulation correct?
- Does the model M resemble the real thing T closely enough in the relevant aspects?
- How can you find out that a simulation gives the wrong results?

Error detection in software using the scientific method

Empirical methods in computer science.

- Cheaper than in other empirical sciences. But is computer science an empirical science?
- Complete control over the set-up of the experiments.
- Reproducibility: computers are deterministic in principle. But random fault behaviour exists ...
- No need to sacrifice monkeys, mice or frogs.

The Empirical Cycle

1. Explore
2. Hypothesize
3. Test
4. Analyse results
5. Draw conclusions, and . . . back to 1.

Experimenten ontwerpen

- Wat is de hypothese die wordt getoetst?
- Wat is de experimentele procedure?
- Wat zijn de gegevens die je wilt verzamelen?
- Is bij elke testuitkomst duidelijk wat hij zegt over de hypothese?
- Hoeveel gegevens heb ik nodig? Heb ik genoeg? Heb ik niet teveel?
- Bedenk hoe de testresultaten moeten worden geïnterpreteerd.
- Wees bedacht op onvoorziene uitkomsten: die kunnen tot nieuwe hypothesen leiden.
- Elk experiment moet een poging zijn om een vooraf geformuleerde vraag te beantwoorden. Zonder vraagstelling vooraf heeft testen geen zin.

Manipulatie Experimenten, en hun rol bij Testen en Debuggen

- Onafhankelijke variabele x (b.v., $x =$ grootte van het woordenboek).
- Afhankelijke variabele y (b.v., $y =$ snelheid van het zoekalgoritme).
- Hypothese: y hangt (op een of andere manier) af van x.
- Manipulatie experiment: x veranderen, y meten of registreren.
- Toepassen hiervan bij Testen en Debuggen: zie Zeller [9].
- Verdere literatuur: Cohen [3].

Empirie in Software Engineering

- Computer programma's zijn formele objecten.
- Betekent dit dat we ze alleen met formele middelen moeten benaderen?
  - Nee, want formele analyse is vaak te moeilijk.
  - Nee, want sommige vragen die we erover kunnen stellen zijn 'proefondervindelijk' van aard.
- Het gedrag van complexe systemen kan worden beschouwd als 'empirisch gegeven'.

Communicatie

Taal Oefen je in het schrijven van correct Nederlands (of Engels).

Structuur Kies voor een bepaalde opbouw en houd je daaraan.
Gebruik alinea's als bouwstenen.

Volgorde van presenteren De volgorde van *presenteren* hoeft niet hetzelfde te zijn als die van *bedenken*.

Werkwijze Schrijven, laten bezinken, herschrijven.
Ga er niet van uit dat wat je schrijft in één keer goed is.

Hulpmiddelen

- Vraag anderen om commentaar
- Stijl-handboeken: zeer goed beknopt stijl-boekje: W. Strunk and E.B. White, The Elements of Style, Macmillan [6]

- Spelling checkers, gespecialiseerde woordenboeken, b.v. [1].
- Engels woordenboek: A.S. Hornby, Oxford Advanced Learner's Dictionary of Current English, Oxford
- Engelse grammatica online (met heel veel oefeningen): `http://www.ucl.ac.uk/internet-grammar/`
- Goed boek over veelgemaakte fouten bij het gebruik van het Engels: Michael Swan, *Practical English Usage*, Oxford University Press, Oxford (vele drukken).

Leren van hoe het niet moet

How to Write a Scientific Paper

E. Robert Schulman
Charlottesville, Virginia

Abstract

We (meaning I) present observations on the scientific publishing process which (meaning that) are important and timely in that unless I have more published papers soon, I will never get another job. These observations are consistent with the theory that it is difficult to do good science, write good scientific papers, and have enough publications to get future jobs.

Voor de rest, zie:

http://members.bellatlantic.net/~vze3fs8i/air/airpaper.html

Structuur van een wetenschappelijk betoog

Titel

Inhoudsopgave

Samenvatting

Inleiding

Betoog

Belang, Beperkingen

Context

Hoe verder

Dank aan...

Literatuur

Titel

- De kortst mogelijke omschrijving
  Voorbeeld: *Leaf Trees*
- Titel + subtitel = Context + onderwerp
  Voorbeeld: *MOLECULAR STRUCTURE OF NUCLEIC ACIDS*
  —
  A Structure for Deoxyribose Nucleic Acid
- Slogan + onderwerp
  Voorbeeld: *Scrap Your Boilerplate:*
  A Practical Design Pattern for Generic Programming

Samenvatting

- Kernwoorden
- Kernzinnen die kernwoorden met elkaar verbinden
- Voorbeeld:

  We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

Inleiding

- Wat heb ik te zeggen?
- Tot welk publiek richt ik mij?
- Wat wordt de lezer al verondersteld te weten?
- Hoe is het nu volgende betoog opgebouwd?
- Wat kan de lezer eventueel overslaan?

Betoog

- Wat heb ik gedaan?
- Welke methode heb ik gebruikt en waarom?
- Wat waren de belangrijkste resultaten?
- Wat valt hieruit te leren?

Belang, Beperkingen

- Wat is het belang van wat ik heb gedaan?
- Wat heb ik verzuimd te doen, en waarom?
- Waar ben ik vastgelopen, en hoe kwam dat?
- Welke problemen voorzie ik bij verdere uitwerking?

Context

- Wie heeft zich eerder met dit onderwerp beziggehouden?
- Wat hebben die anderen over het onderwerp gezegd?
- Met wat daarvan ben ik het eens of oneens, en waarom?
- Hoe verhoudt mijn bijdrage zich tot wat anderen hebben gezegd?

Uitleiding

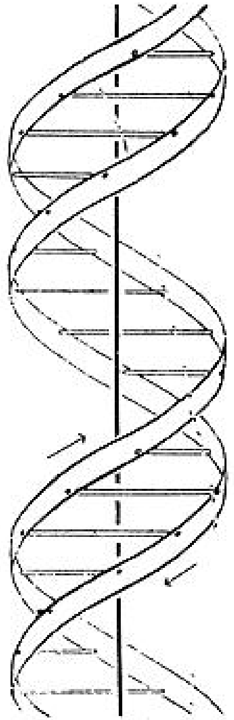
- Hoe nu verder?
- Wat betekent dit?
- Waar is meer informatie te vinden?

Voorbeeld

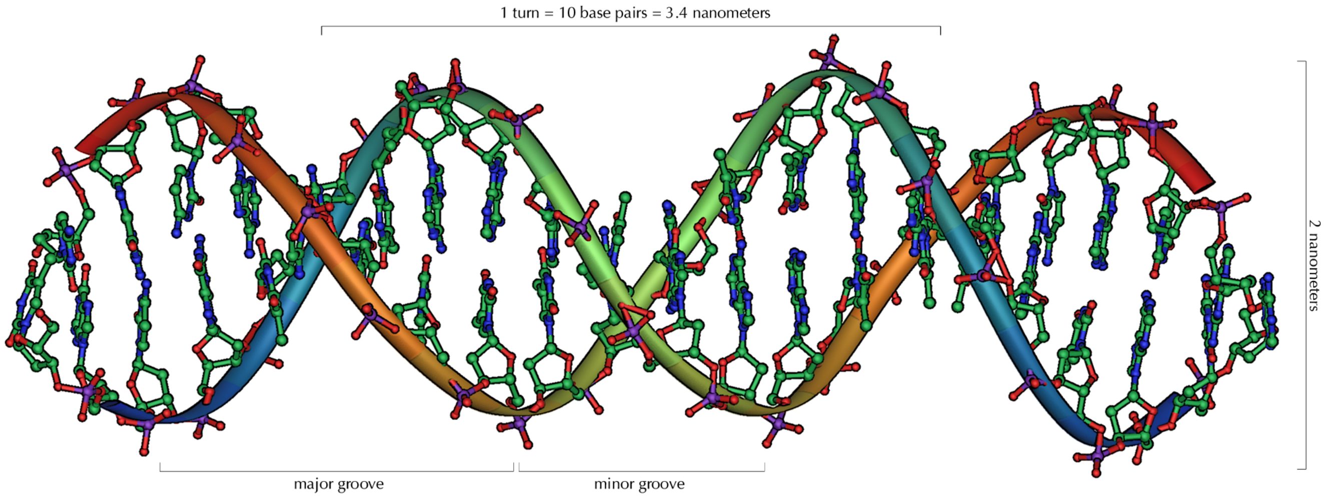
2 April 1953

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid



This figure is purely diagrammatic. The two ribbons symbolize the two phosphate—sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis


1 turn = 10 base pairs = 3.4 nanometers
2 nanometers
major groove
minor groove

[intro]

We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

[context]

A structure for nucleic acid has already been proposed by Pauling and Corey (1). They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

[betoog]

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining -D-deoxyribofuranose residues with 3',5' linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right- handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's2 model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each every 3.4 A. in the z-direction. We have assumed an angle of 36 between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. The are joined together in pairs, a single base from the other chain, so that the two lie side by side with identical z-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows : purine position 1 to pyrimidine position 1 ; purine position 6 to pyrimidine position 6.

[gevolgen]

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are : adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine ; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

[empirische ondersteuning]

It has been found experimentally (3,4) that the ratio of the amounts of adenine to thymine, and the ration of guanine to cytosine, are always bery close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact. The previously published X-ray data (5,6) on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

[belang]

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

[verwijzingen, dank]

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at King's College, London. One of us (J. D. W.) has been aided by a fellowship from the National Foundation for Infantile Paralysis.

J. D. WATSON F. H. C. CRICK

Medical Research Council Unit for the Study of Molecular Structure of Biological Systems, Cavendish Laboratory, Cambridge. April 2.

[literatuur]

1. Pauling, L., and Corey, R. B., Nature, 171, 346 (1953); Proc. U.S. Nat. Acad. Sci., 39, 84 (1953).

2. Furberg, S., Acta Chem. Scand., 6, 634 (1952).

3. Chargaff, E., for references see Zamenhof, S., Brawerman, G., and Chargaff, E., Biochim. et Biophys. Acta, 9, 402 (1952).

4. Wyatt, G. R., J. Gen. Physiol., 36, 201 (1952).

5. Astbury, W. T., Symp. Soc. Exp. Biol. 1, Nucleic Acid, 66 (Camb. Univ. Press, 1947).

6. Wilkins, M. H. F., and Randall, J. T., Biochim. et Biophys. Acta, 10, 192 (1953).

VOL 171, page737, 1953

References

[1] H. Biemond. *Woordenboek Automatisering*. Koninklijke PBNA, 1985 (tweede druk).

[2] J.D. Bransford and B.S. Stein. *The Ideal Problem Solver: A Guide for Improving Thinking, Learning, and Creativity*. W.H. Freeman, 1984. A Scientific American Special.

[3] P. Cohen. *Empirical Methods for Artificial Intelligence*. MIT Press, 1995.

[4] Mihaly Csikszentmihaly. *Flow: the psychology of optimal experience*. Harper and Row, 1990.

[5] R. Fritz. *Creating*. Balantine Books, 1991.

[6] W. Strunk Jr. and E.B. White. *The Elements of Style*. Macmillan, 1979 (third edition).

[7] L. Mlodinow. *Some time with Feynman*. Penguin, 2003.

[8] A.M. Turing. On computable real numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 2(42):230–265, 1936.

[9] Andreas Zeller. *Why Programs Fail: A Guide to Systematic Debugging*. Morgan Kaufmann, 2005.