“ Quien sabe por Algebra, sabe scientificamente”

J.N. Oliveira

Dept. Informática,
Univ. Minho
Braga, Portugal

1st MATISSE meeting, 29th March 2008, Braga

Context

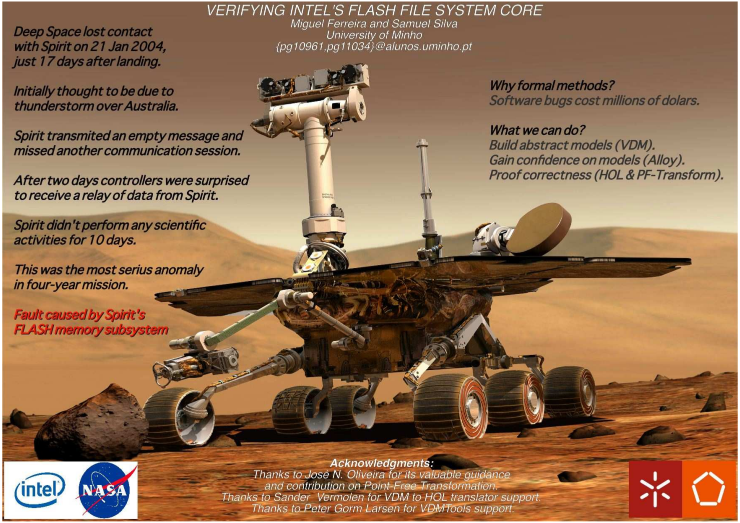
- Globalization: need for safe software systems and communications
- Dependence on **information technology**
- Safety and security
- **Why** is what matters in the {What, how, why} triad
- Demand for correctness **proofs**
- Maths (il)literacy — the **issue**!
- **High school** is too late...
- Opportunity for both mathematicians and computer scientists

A real-life example

NASA call for Verified File System (VFS) software:

- **VFS** (Verified File System) on Flash Memory — challenge put forward by Rajeev Joshi and Gerard Holzmann (NASA JPL) [1]
- NASA calls for 100% correct FS software layer for NAND flash memories to be installed in spacecrafts.

Work on the VFS challenge at U.Minho



Mastering complexity

- Real life software systems can be extremely **complex**
- Thousands/millions of lines of **code**, hundreds of **proofs** to discharge
- Mechanical **theorem proving** has limited application
- Need for **abstraction** skills
- Abstract **models** are the key to reliable system design
- Going "abstract" to obtain "concrete" success — not a contradiction!
- Abstract models bound to be written in **maths** notation.

Scientific? Pre-scientific?

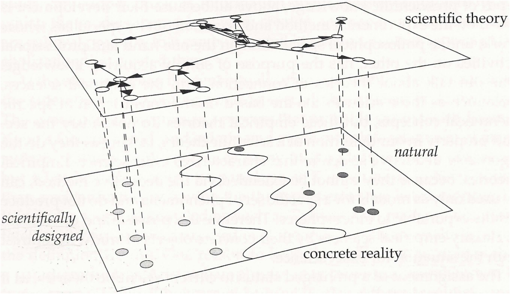
In an excellent book on the history of scientific technology (*"How Science Was Born in 300BC and Why It Had to Be Reborn*), Lucio Russo [4] writes:

> *The immense usefulness of* **exact** *science consists in providing* **models** *of the real world within which there is a guaranteed method for telling false statements from true. (...) Such models, of course, allow one to describe and* **predict** *natural phenomena, by translating them to the theoretical level via* **correspondence rules**, *then solving the "***exercises***" thus obtained and translating the solutions obtained back to the real world.*

Disciplines unable to build themselves around *"exercises"* are regarded as **pre-scientific**.

$e = m + c$

Also quoted from Russo's book :



Vertical lines mean **abstraction**, horizontal ones mean **calculation**:

$$\textit{engineering} = \underline{\textit{model}}\ \textit{first, then}\ \underline{\textit{calculate}}\ (e = m + c)$$

We (should) know how to **calculate** since the school desk...

School maths example

The problem

My three children were born at a 3 year interval rate. Altogether, they are as old as me. I am 48. How old are they?

The model

$$x + (x+3) + (x+6) = 48$$

The calculation

$$\begin{aligned} & 3x + 9 = 48 \\ \Leftrightarrow \quad & \{ \text{ ``al-djabr'' rule } \} \\ & 3x = 48 - 9 \\ \Leftrightarrow \quad & \{ \text{ ``al-hatt'' rule } \} \\ & x = 16 - 3 \end{aligned}$$

School maths example

The problem

My three children were born at a 3 year interval rate. Altogether, they are as old as me. I am 48. How old are they?

The model

$$x + (x + 3) + (x + 6) = 48$$

The calculation

$$\begin{aligned} & 3x + 9 = 48 \\ \Leftrightarrow \quad & \{ \text{ “al-djabr” rule } \} \\ & 3x = 48 - 9 \\ \Leftrightarrow \quad & \{ \text{ “al-hatt” rule } \} \\ & x = 16 - 3 \end{aligned}$$

School maths example

The problem

My three children were born at a 3 year interval rate. Altogether, they are as old as me. I am 48. How old are they?

The model

$$x + (x + 3) + (x + 6) = 48$$

The calculation

$$\begin{array}{ll} & 3x + 9 = 48 \\ \Leftrightarrow & \qquad \{ \text{ “al-djabr” rule } \} \\ & 3x = 48 - 9 \\ \Leftrightarrow & \qquad \{ \text{ “al-hatt” rule } \} \\ & x = 16 - 3 \end{array}$$

School maths example

The solution

$$\begin{aligned} x &= 13 \\ x+3 &= 16 \\ x+6 &= 19 \end{aligned}$$

Comments, questions....

- proof layout (maths + "why" lines)
- "al-djabr" rule ?
- "al-hatt" rule ?

Have a look at Pedro Nunes (1502-1578) *Libro de Algebra en Arithmetica y Geometria* dated 1567 [3] . . .

School maths example

The solution

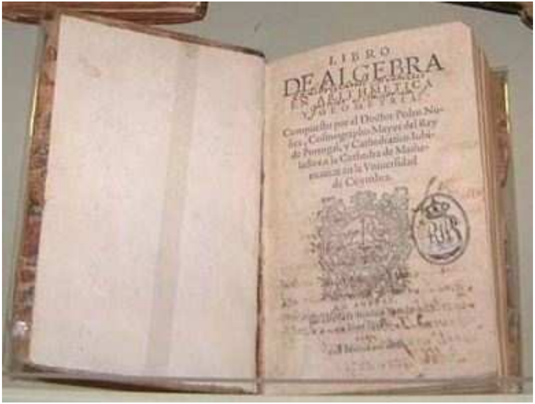
$$\begin{aligned} x &= 13 \\ x + 3 &= 16 \\ x + 6 &= 19 \end{aligned}$$

Comments, questions....

- proof layout (maths + "why" lines)
- "al-djabr" rule ?
- "al-hatt" rule ?

Have a look at Pedro Nunes (1502-1578) *Libro de Algebra en Arithmetica y Geometria* dated 1567 [3] . . .

Libro de Algebra en Arithmetica y Geometria (1567)



(...) ho inuẽtor desta arte foy hum Mathematico Mouro, cujo nome era Gebre, & ha em alguãs Liuarias hum pequeno tractado Arauigo, que contem os capitulos de q̃ usamos
(fol. a ij r)

Reference to *On the calculus of al-gabr and al-muqâbala* [1] by Abû Abd Allâh Muhamad B. Mûsâ Al-Huwârizmî, a famous 9c Persian mathematician.

[1] Original title: *Kitâb al-muhtasar fi hisab al-gabr wa-almuqâbala*.

Calculus of al-gabr and al-muqâbala

al-djabr:

$$x + z \leq y \quad \Leftrightarrow \quad x \leq y - z$$

al-hatt:

$$x * z \leq y \quad \Leftrightarrow \quad x \leq y \div z \qquad\qquad (z > 0)$$

al-muqâbala, ex:

$$4x^2 + 3 = 2x^2 + 2x + 6 \quad \Leftrightarrow \quad 2x^2 = 2x + 3$$

Geometry in the Renaissance

Hot topic in the 16c: revisit old geometrical problems, inc. Euclid's Elements.

Problem 12 in Johan Müller's (1436-1476) *"De Triangulis"*, vol.II:

Given

$$\begin{aligned} bc &= 20 \\ ad &= 5 \\ \frac{ab}{ac} &= \sqrt{5} \end{aligned}$$

find ab, ac and bd.

This is Question 46 in Nunes book (fol. 270r), given as example of problem which Müller could not solve on pure geometric grounds. . .

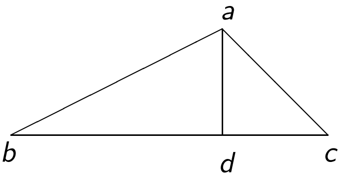
Geometry in the Renaissance

Hot topic in the 16c: revisit old geometrical problems, inc. Euclid's Elements.

Problem 12 in Johan Müller's (1436-1476) *"De Triangulis"*, vol.II:

Given

$$\begin{aligned} bc &= 20 \\ ad &= 5 \\ \frac{ab}{ac} &= \sqrt{5} \end{aligned}$$



find ab, ac and bd.

This is Question 46 in Nunes book (fol. 270r), given as example of problem which Müller could not solve on pure geometric grounds. . .

... solved "by algebra"

Nunes model is based on the *inuento Pithagorico* [2]:

Model

(...) Queriendo pos conoscer los lados (...) pornemos .d.c. parte menor ser .1.co. [read: $x = dc$, where *co* is "cousa" = "the thing" we are looking for] *(...) Y porque .bd. es .20.m̃.1.co (...) sera el su quadrado 400.p̃.1.ce.m̃.40.co* [read: $20^2 + x^2 - 40x$] *(...)*

Thus he reaches model

$$\frac{ab^2}{ac^2} = \frac{425 - 40x + x^2}{x^2 + 25} = 5$$

[2] "Pythagoras invention", ie. Prop. 47 of Euclid's Elements — see eg. `http://aleph0.clarku.edu/ djoyce/java/elements/bookI/propI47.html`

...solved "by algebra"

Nunes algebraic calculation

$$\frac{425 - 40x + x^2}{x^2 + 25} = 5$$

$$\Leftrightarrow \qquad \{ \text{ rule } \tfrac{a}{b} = \tfrac{c}{d} \Leftrightarrow ad = bc \text{ etc } \}$$

$$425 - 40x + x^2 = 5x^2 + 125$$

$$\Leftrightarrow \qquad \{ \text{ “calculus of al-gabr and al-muqâbala” (...) } \}$$

$$75 = x^2 + 10x$$

This leads to the expected

Solution

(...) sera luego *.a.b. R.250.* e *.a.c. R.50* *[read: $ab = \sqrt{250}$ and $ac = \sqrt{50}$]*

Nunes comments

Algebra (...) is thing causing admiration

(...) *Principalmente que vemos algumas vezes, no poder vn gran Mathematico resoluer vna question por medios Geometricos, y resolverla por Algebra, siendo la misma Algebra sacada de la Geometria, q̃ es cosa de admiraciõ.*

that is (literal — not literary — translation):

> *(...) Mainly because we see often a great Mathematician unable to resolve a question by Geometrical means, and solve it by Algebra, being that same Algebra taken from Geometry, which is thing causing admiration.*

[Pedro Nunes (1502-1578) in Libro de Algebra en Arithmetica y Geometria, 1567, fols. 270–270v.]

Letting "the symbols do the work" in the 16c

fol. 269r-269v:

Going algebraic

Mas la causa porq̃ obro por Algebra quasi siempre, es que este tratado es hecho para que en el practiquen las Reglas de Algebra en los casos de Geometria.

ie.

> *But the reason why I work by Algebra almost always, is that this treaty is written so that in it you practise the rules of Algebra in case studies of Geometry*

Letting "the symbols do the work" in the 16c

fol. 269r-269v:

Deduction first

Y tambien porque quien obra por Algebra va entendiendo la razon de la obra que haze, hasta la yqualacion ser acabada. (...) De suerte que, quien obra por Algebra, va haziendo discursos demonstrativos.

ie.

> *And also because one performing by Algebra is understanding the reason of the work one does, until the equality is finished. (...) So much so that, who works by Algebra is doing a demonstrative discourse.*

Verdict

(...) De manera, que quien sabe por Algebra, sabe <u>scientificamente</u>.

(in this way, who knows by Algebra knows scientifically)

From a textbook (vectorial calculus)

Moving on:

- Enough about the past
- What about today?

Problems:

- Proof skills acquired at middle school level don't **scale up** to high school
- **Proofs** virtually absent from middle school curricula
- Geometry still an exception

Challenges:

- Train students to do **constructive** proofs (eg. in geometry)

What do we mean by *"constructive"* ? Let us see an example.

From a textbook (vectorial calculus)

The problem

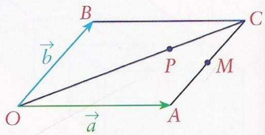
Uma demonstração

$[OACB]$ é um paralelogramo.

$\overrightarrow{OA} = \vec{a}$ e $\overrightarrow{OB} = \vec{b}$

$\overrightarrow{OP} = \frac{2}{3}\overrightarrow{OC}$

M é o ponto médio de $[AC]$.



Prove que B , P e M pertencem à mesma recta.

Procure acompanhar os passos seguidos na seguinte demonstração.

From a textbook (vectorial calculus)

The proof given:

Demonstração:

Tem-se sucessivamente:

$$\overrightarrow{OC} = \overrightarrow{OA} + \overrightarrow{AC} = \vec{a} + \vec{b}$$

$$\overrightarrow{OP} = \frac{2}{3}\overrightarrow{OC} = \frac{2}{3}(\vec{a} + \vec{b})$$

$$\overrightarrow{OM} = \vec{a} + \frac{1}{2}\vec{b}$$

$$\overrightarrow{BP} = \overrightarrow{BO} + \overrightarrow{OP} = -\vec{b} + \frac{2}{3}(\vec{a} + \vec{b}) = \frac{2}{3}\vec{a} - \frac{1}{3}\vec{b}$$

$$\overrightarrow{BM} = \overrightarrow{BO} + \overrightarrow{OM} = -\vec{b} + \vec{a} + \frac{1}{2}\vec{b} = \vec{a} - \frac{1}{2}\vec{b} = \frac{3}{2}\left(\frac{2}{3}\vec{a} - \frac{1}{3}\vec{b}\right)$$

Portanto, $\overrightarrow{BM}$ é colinear com $\overrightarrow{BP}$ porque $\overrightarrow{BM} = \frac{3}{2}\overrightarrow{BP}$.

Como $[BP]$ e $[BM]$ são paralelos e têm em comum o ponto B, os pontos B, P e M pertencem à mesma recta.

É capaz de reproduzir esta demonstração?

From a textbook (vectorial calculus)

Comments:

- **verification**, not calculation
- not **goal** oriented
- not "constructive" enough
- tricky? cf. final question *"É capaz de reproduzir esta demonstração?"*...

In a constructive proof, the **starting** point will be the **goal** itself:

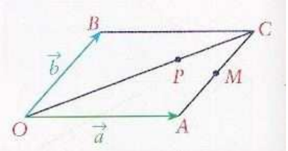
Goal: B, P and M on the same line

equivalent to

$\overrightarrow{BP} = k\overrightarrow{BM}$ **for some** k

Proof amounts to calculating such k, if any

"Constructive proof" (modeling)



$$\overrightarrow{BP} = k\overrightarrow{BM}$$

$$\Leftrightarrow \quad \{ \text{ vector sums } \}$$

$$\overrightarrow{BO} + \overrightarrow{OP} = k(\overrightarrow{BO} + \overrightarrow{OA} + \overrightarrow{AM})$$

$$\Leftrightarrow \quad \{ \text{ let } \overrightarrow{OP} = k_1\overrightarrow{OC},\ \overrightarrow{AM} = k_2\overrightarrow{AC} \ \}$$

$$\overrightarrow{BO} + k_1\overrightarrow{OC} = k(\overrightarrow{BO} + \overrightarrow{OA} + k_2\overrightarrow{AC})$$

$$\Leftrightarrow \quad \{ \text{ let } \overrightarrow{OA} = \overrightarrow{a},\ \overrightarrow{OB} = \overrightarrow{b} \ \}$$

$$-\overrightarrow{b} + k_1(\overrightarrow{a} + \overrightarrow{b}) = k(-\overrightarrow{b} + \overrightarrow{a} + k_2\overrightarrow{b})$$

"Constructive proof" (calculation)

$$-\overrightarrow{b} + k_1(\overrightarrow{a} + \overrightarrow{b}) = k(-\overrightarrow{b} + \overrightarrow{a} + k_2\overrightarrow{b})$$

$\Leftrightarrow$ { "al-muqâbala" (1,2,3) }

$$k_1\overrightarrow{a} + (k_1 - 1)\overrightarrow{b} = k\overrightarrow{a} + k(k_2 - 1)\overrightarrow{b}$$

$\Leftrightarrow$ { equality rule (4) }

$$k_1 = k \wedge k_1 - 1 = k(k_2 - 1)$$

$\Leftrightarrow$ { "calculus of al-gabr and al-muqâbala" }

$$k = k_1 = \frac{1}{2 - k_2}$$

Given problem corresponds to $k_2 = \frac{1}{2}$ and $k_1 = \frac{2}{3}$.
Other cases: $k_2 = k_1 = 1$ ($P = M = C$), $k_2 = 0 \wedge k_1 = \frac{1}{2}$ ($M = A$), etc

Thanks to

Vectorial calculus:

$$\begin{aligned} k(\overrightarrow{a} + \overrightarrow{b}) &= k\overrightarrow{a} + k\overrightarrow{b} &\quad (1)\\ (k+j)\overrightarrow{a} &= k\overrightarrow{a} + j\overrightarrow{a} &\quad (2)\\ k(j\overrightarrow{a}) &= (kj)\overrightarrow{a} &\quad (3) \end{aligned}$$

and, for non co-linear $\overrightarrow{a}$, $\overrightarrow{b}$, the equality rule:

$$k\overrightarrow{a} + j\overrightarrow{b} = m\overrightarrow{a} + n\overrightarrow{b} \quad \Leftrightarrow \quad k = m \ \wedge \ j = n \qquad (4)$$

Summary

Constructive proof means **calculating** a (often necessary and) sufficient condition for the **goal** to hold.

Back to the primary school desk

The **whole division** algorithm

$$\begin{array}{r|l} 7 & 2 \\ \hline 1 & 3 \end{array} \qquad 2 \times 3 + 1 = 7 \quad , \text{"ie."} \qquad 3 = 7 \div 2$$

However

$$\begin{array}{r|l} 7 & 2 \\ \hline 3 & 2 \end{array} \qquad 2 \times 2 + 3 = 7 \quad \wedge \quad 2 \neq 7 \div 2$$

$$\begin{array}{r|l} 7 & 2 \\ \hline 5 & 1 \end{array} \qquad 2 \times 1 + 5 = 7 \quad \wedge \quad 1 \neq 7 \div 2$$

In fact:

$$\begin{array}{r|l} n & d \\ \hline r & q \end{array} \qquad q = n \div d \;\Leftrightarrow\; d \times q + r = n$$

provided q is the largest such q (r smallest)

Back to the primary school desk

So, $n \div d$ is a a supremum:

$$\begin{aligned} n \div d &= \langle \bigvee q :: \langle \exists\, r \geq 0 :: d \times q + r = n \rangle\rangle \\ &= \langle \bigvee q :: d \times q \leq n \rangle \end{aligned}$$

It takes a while before children master the " $\bigvee$ semantics ". Even once they grasp it,

- how tractable is the above definition?
- challenge: ask your students to check

$$(n \div m) \div d = n \div (d \times m) \tag{5}$$

from such a definition. (It may take a while.)

Alternative

We know from mathematics that suprema such as $\langle \bigvee\ q\ ::\ d \times q \leq n \rangle$ satisfy their own **al-djabr** rules, in this case:

$$q \times d \leq n \;\Leftrightarrow\; q \leq n \div d \qquad (d > 0) \tag{6}$$

recall the “al-hatt” rule.

- This property **means the same** as the $\bigvee$ definition and has one further advantage: it is easy to calculate with.
- From it, we will both infer **properties** of $\div$ and derive an **algorithm** which will compute whole division.
- We only need one more ingredient: **indirect equality**

Indirect equality principle

Well-known in set theory to define set equality:

$$A = B \quad \Leftrightarrow \quad \langle \forall\, x \;::\; x \in A \Leftrightarrow x \in B \rangle$$

Another form is:

$$A = B \quad \Leftrightarrow \quad \langle \forall\, X \;::\; X \subseteq A \Leftrightarrow X \subseteq B \rangle$$

In fact, any **partial order** can be used to establish equality by indirection. In case of numbers:

$$n = m \quad \Leftrightarrow \quad \langle \forall\, x \;::\; x \leq n \Leftrightarrow x \leq m \rangle \tag{7}$$

"Al-djabr" calculation

Back to (5), our goal

$$(n \div m) \div d = n \div (d \times m)$$

is (by indirect equality) the same as

$$\langle \forall\, x \;::\; x \leq (n \div m) \div d \;\Leftrightarrow\; x \leq n \div (d \times m) \rangle$$

We calculate (for $m, d > 0$):

$$
\begin{array}{ll}
 & x \leq (n \div m) \div d \\
\Leftrightarrow & \quad \{ \text{ “al-djabr” (6) } \} \\
 & x \times d \leq n \div m \\
\Leftrightarrow & \quad \{ \text{ “al-djabr” (6) } \} \\
 & (x \times d) \times m \leq n \\
\Leftrightarrow & \quad \{ \times \text{ is associative } \} \\
 & x \times (d \times m) \leq n \\
\Leftrightarrow & \quad \{ \text{ “al-djabr” (6) } \} \\
 & x \leq n \div (d \times m)
\end{array}
$$

Good news: "al-djabr" rules scale-up

From the "al-djabr" rules studied so far,

$$a \times b \ \leq \ c \quad \Leftrightarrow \quad a \ \leq \ c \div b \qquad (b > 0) \tag{8}$$

$$a + b \leq c \quad \Leftrightarrow \quad a \leq c - b \tag{9}$$

we easily derive the **composite** rules ($c > 0$)

$$ax + b \leq c \quad \Leftrightarrow \quad a \leq (c - b) \div x$$

$$(a + b)c \leq x \quad \Leftrightarrow \quad a \leq (x \div c) - b$$

from which we calculate eg.

$$a \div c - b \quad = \quad (a - bc) \div c$$

and so and so on.

Follow up

- As computer scientists, our main job is to design **programs** which are **correct** with respect to their specifications
- "Al-djabr" can be regarded as specifications
- João Ferreira will show how to derive a computer program for whole division from its "al-djabr" rule (see also these slides' addendum)
- This closes the cycle and illustrates the intimate interplay between **maths** and **computing**.

Epilogue

- "Al-djabr" method scales up to complex problem domains — cf. Galois connections
- Logic can be "al-djabr'ed" — cf. binary relation calculus
- Teaching maths in this way would ensure **smooth** transition from middle to high school — while **problem domains** become more complex the **principles** remain the same
- The current situation: 1st year students feel often completely uneasy when faced with computer science problems
- They lack in **abstraction** and proof skills
- **Elegance** and conciseness are a must: hundreds of proofs!

Epilogue

Quoting Jeff Kramer [2]:

> *Why is it that some software engineers and computer scientists are able to produce clear,* **elegant** *designs and programs, while others cannot? Is it possible to improve these skills through* **education** *and training? Critical to these questions is the notion of* **abstraction**.

Epilogue

Still Jeff Kramer [2]:

Abstraction *is widely used in other disciplines such as* **art** *and* **music**. *For instance (...) Henri Matisse manages to clearly represent the* **essence** *of his subject, a naked woman, using only simple lines or cutouts. His representation* **removes** *all detail yet* **conveys** *much.*

Addendum

Calculation of recursive version of $x \div y$:

$$
\begin{array}{ll}
 & z \leq x \div y \\
\Leftrightarrow & \qquad \{ \text{ "al-djabr" (6) } \} \\
 & zy \leq x \\
\Leftrightarrow & \qquad \{ \text{ cancellation } \} \\
 & zy - y \leq x - y \\
\Leftrightarrow & \qquad \{ \text{ "al-djabr" (6) provided } x \geq y \ \} \\
 & z - 1 \leq (x - y) \div y \\
\Leftrightarrow & \qquad \{ \text{ "al-djabr" (9) } \} \\
 & z \leq (x - y) \div y + 1
\end{array}
$$

By indirection, the general case $(x \geq y)$ is thus:

$$x \div y = (x - y) \div y + 1 \tag{10}$$

Addendum

Finally the case in which $x < y$:

$$
\begin{array}{ll}
 & z \leq x \div y \\
\Leftrightarrow & \{ \text{ "al-djabr" (6) } \} \\
 & zy \leq x \\
\Leftrightarrow & \{ \ zy < y \text{ since } x < y \ \} \\
 & z \leq 0
\end{array}
$$

Again by indirection we get $x \div y = 0$. Altogether, in Haskell:

```
x -:- y |   x <  y    = 0
x -:- y |   x >= y    = 1 + (x - y) -:- y
```

Rajeev Joshi and Gerard J. Holzmann.
A mini challenge: build a verifiable filesystem.
Formal Asp. Comput., 19(2):269–272, 2007.

Jeff Kramer.
Is abstraction the key to computing?
Commun. ACM, 50(4):37–42, April 2007.

Pedro Nunes.
Libro de Algebra en Arithmetica y Geometria.
Original edition by Arnoldo Birckman (Anvers), 1567.
Re-edited by the Lisbon Academy of Sciences in 1950 as Vol. VI of Pedro Nunes' Works (Imprensa Nacional de Lisboa).

Lucio Russo.
The Forgotten Revolution: How Science Was Born in 300BC and Why It Had to Be Reborn.
Springer-Verlag, September 2003.