

Embedding

Multi-purpose Device for Understanding Mathematics and its Development, or Empty Generalization?

Høyrup, Jens

Published in:
Ganita Bharati

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Høyrup, J. (2016). Embedding: Multi-purpose Device for Understanding Mathematics and its Development, or Empty Generalization? *Ganita Bharati*, 38(1), 1-29.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact rucforsk@ruc.dk providing details, and we will remove access to the work immediately and investigate your claim.

Embedding: Multipurpose Device for Understanding Mathematics and its Development, or Empty Generalization?

JENS HØYRUP

Yvonne Dold-Samplonius and Paulus Gerdes
in memoriam

ABSTRACT

“Embedding” as a technical concept comes from linguistics, more precisely from grammar. The present paper investigates whether it can be applied fruitfully to certain questions that have been investigated by historians (and sometimes philosophers) of mathematics:

1. The construction of numeral systems, in particular place-value and quasi place-value systems.
2. The development of algebraic symbolisms.
3. The discussion whether “scientific revolutions” ever take place in mathematics, or new conceptualizations always include what preceded them.

A final section investigates the relation between spatial and linguistic embedding and concludes that the spatio-linguistic notion of embedding can be meaningfully applied to the former two discussions, whereas the apparent embedding of older within new theories is rather an ideological mirage.

Key words: Embedding; Generative Grammar; Place-value systems; algebraic symbolism; Revolutions in mathematics; Spatiality and language; Spatiality and mathematical thought.

The following reflections were spurred by an invitation to present something of my own choice at the IX Congreso de la Asociación Española de Semiótica “Humanidades, ciencia y tecnología” in Valencia in December 2000.¹ Semiotics as such was never my field, but I decided to take advantage of the occasion and explore

¹ The paper was accepted for the proceedings, which however never appeared. I am grateful for the possibility to publish a revised version in the present context.

an idea located somewhere in the boundary region between the history of mathematics, philosophy of mathematics, psychology, linguistics and semiotics.

This was an idea that dawned to me some years earlier and which I had never found the time to elaborate. It had to do with the notion of “embedding” and its possible applicability to certain higher-level questions in the historiography of mathematics (and, in as far as the history of mathematics is relevant for the philosophy of mathematics, also problems belonging to this latter domain).

The concept – that is, the specific technical meaning of the word, which itself is obviously older and more broadly used – originated in grammar (more precisely in syntax). In the sentence “the canary, which had been devoured by the cat, had been the joy of my mornings”, the relative clause “which had been devoured by the cat” is *embedded* in the main clause; as a whole it functions as an element of the main clause, and it can be replaced (with a change of word order) by a single adjective – for instance, “yellow”.

Embedding proper thus occurs when a whole subordinate clause – a sentence in itself – occupies the place of a sentence member (be it in a main clause, be it in a higher-level subordinate clause) or of some other phrase (as when a relative clause fills the place of an adjective). The most developed form of embedding is the iterative expansion of this latter type as described, for instance, in the recursive schemes of generative grammar (on recursion, cf. examples below).

Embedding-like phenomena characterize language in various shapes and at different levels, not only in sentence syntax. One, only virtual and perhaps not properly carrying the name, is the Saussurean interplay of “syntagm” – a sequence of places in a sentence – and “paradigm” – the set of possible values of the “variable” occupying a particular place.² It is related to the use of general terms (e.g., “animal”) in language that may stand for any one of a number of particulars (*in casu* “cat”, “canary”, “eel”, ...), but with the difference that the general terms are actually present in the sentence “the animal is alive”, whereas the places in a syntagm are potentialities which (in the best Aristotelian manner) are only actualized by being filled out by paradigm members – the syntagm in itself is no sentence but an abstract scheme. It is also somewhat similar to the idea of a

² That is, for example, the “syntagm” subject – verb – object, where (from the grammatical point of view) the place “subject” can be filled by any member of the “paradigm” consisting of nouns (and pronouns in the nominative), the place “verb” by the paradigm consisting of transitive verbs, and “object” by a paradigm containing nouns and accusative forms of pronouns.

Cartesian product, but in this case with the difference that here the places in the syntagm are occupied by *identical* paradigms,³ whereas the places in a sentence syntagm are to be filled out by different paradigms.

My intention in the following is to explore whether and to which extent the concept can be applied fruitfully and coherently to three areas pertaining to the structure and history of mathematics:

- The structure of numerals and the emergence of place-value and quasi-place-value notations.
- The relation of algebraic symbolism to preceding representations.
- The alleged absence of “revolutions” from the development of mathematics.

I shall tie all three discussions to the notion of “embedding”; it remains to be seen whether it is applicable *in the same or related ways* in all three cases, and thus whether applying it provides some real insight; if the three uses are unrelated, it is an empty metaphor which might as well be discarded.

Empty or not, “embedding” is a spatial metaphor. I shall close by some reflections on spatiality, language and mathematics.

Numerals

A theme which “historically interested” mathematicians are fond of treating is the emergence of place-value notations. In agreement with the “Lamarckian fallacy”⁴ so close at hand in every evolutionary thought, *our* present position is seen as the goal of preceding changes; maybe further developments shall attain even higher peaks (this was what Nietzsche supposed, adding an *Übermensch* to Lamarck’s ladder of perfection), but these will by necessity ascend *from us*. There may be blind alleys in evolution (even Lamarck supposed that animals that happened to live in the sea might develop into more perfect fishes but

³ An example would be a statistical table showing the variation of the population of a number of cities over time: for each year, the same cities are listed in the table.

⁴ The term refers to Lamarck’s original thought as set forth in his *Philosophie zoologique* from 1809 – *man* is the perfect being, and other animals strive in their development toward this perfection. The fallacy is absent from the “neo-Lamarckian” doctrines from the late nineteenth century, from which the teleological element has been eliminated, and where the originally ancillary inheritance of acquired characteristics has become the central explanatory device.

could not attain the perfection of man); but the blind alleys are always represented by “the others”.

The place-value system can be explained as an embedding in various ways.⁵ One is recursive on the level of places only,

$$\{ \text{place sequence} \} \longrightarrow \left\{ \begin{array}{l} \{ \text{place}_f \} \\ \{ \text{place sequence} \} \{ \text{place} \} \end{array} \right.$$

where $\{ \text{place}_f \}$ ($_f$ for “first”) may be filled out by any of the digits 1, 2, ..., 9 constituting the paradigm $\{ \text{digit}_f \}$, and $\{ \text{place} \}$ by any of the digits 0, 1, 2, ..., 9 (the paradigm $\{ \text{digit} \}$). In plain words, the minimal string representing a number consists of a single place that may be filled by any of the digits 1, 2, ..., 9; others number strings may be produced by appending one or more places to the right that can be filled by any of the digits 0, 1, 2, ..., 9. Because of the incomplete identity of the paradigms $\{ \text{digit} \}$ and $\{ \text{digit}_f \}$, this is no full Cartesian product, although it comes closer to this type than the linguistic syntagm-paradigm structure. For use in the following, I shall call it “type I” description.⁶

The scheme which best corresponds to current explanations of the system (“type II” description) avoids the explicit reference to places (as did studies of syntax prior to the advent of structuralism), but it still separates the writing from the arithmetical meaning,

$$\{ \text{written number} \} \longrightarrow \left\{ \begin{array}{l} \{ \text{digit}_f \} \\ \{ \text{written number} \} \{ \text{digit} \} \end{array} \right.$$

The corresponding numerical value is explained as a sum $\sum_{i=0}^n a_i \cdot 10^i$; that is, it still refers to the single places (even this is an analogue of the way a pre-structuralist syntactical analysis ascribes meaning to a sentence). A recursive definition which does not refer to the numbers of the places can be made in the shape of an algorithm

⁵ In the interest of simplicity, I shall at first restrict the discussion to the writing of positive integers. Later, the reference to historical examples will force us to introduce fractions.

⁶ If we admit an initial string of zeroes in the writing of a numeral, the distinctions between $\{ \text{digit} \}$ and $\{ \text{digit}_f \}$ and between $\{ \text{place} \}$ and $\{ \text{place}_f \}$ are evidently superfluous. The cost is that numbers no longer correspond to numerals but to equivalence classes among numerals.

with a single loop, $\langle \text{value} \rangle := \langle \text{value} \rangle \times 10 + \langle \text{next digit} \rangle$, starting with $\langle \text{value} \rangle = \langle \text{first digit} \rangle$ and ending when $\langle \text{next digit} \rangle$ is the last digit, corresponding to the formula

$$(\dots((d_n \cdot 10 + d_{n-1}) \cdot 10 + d_{n-2}) \cdot 10 \dots + d_1) \cdot 10 + d_0.$$

It is rarely pointed out that the place-value notation implies a more refined type of embedding (“type III”), which cannot be formalized as a simple recursive scheme of the type used in generative grammar, and which I shall therefore approach through examples.⁷ It has the strange property that it can be interpreted in different ways – to be “polysemic” – at the intermediate stage but to lead to the same final total meaning irrespective of the choice of intermediate interpretations.⁸ In type-I and type-II interpretation, a number of type $d|0|0|\dots|0$ (which may be understood as an additive contribution to a more complicated multi-place number) means $d \times 1|0|0|\dots|0$; apart from the recursive definition of {place sequence}, embedding is thus only present in the sense that a “1” can be

3600	600	60	10	1
	1	0	2	5
	1	0	2	5
	2	0	5	0
1	0	2	5	0
1	0	2	5	0
2	0	5	4	0

replaced by any digit. But in a multi-place number $a|b|c|d|e|f|\dots|r$, any sequence of digits may actually be taken out to represent a number counting the units at its own lowest place; thus, $a|b|c|d|e|f|\dots|r = (a|0|0|0|0|e|f|\dots|r) + (b|c|d \times 1|0|0|\dots|0)$ – less abstractly, $234875 = 234 \times 10^3 + 875 = 23 \times 10^4 + 48 \times 10^2 + 75 = 2 \times 10^5 + 3487 \times 10^1 + 5$, etc. That is, if a multi-place number is put into the place of a unity of any level, all its “overflowing” places end up where they “should” stand.

This property is essential for the simplicity of algorithms. In order to understand why this is so one may look at how addition works in the mixed decimal-seximal

⁷ Not only is this property of the place value notation rarely pointed at or explained, but mathematics teachers tend to censure students’ spontaneous taking advantage of the principle in locutions like “three point twentyfive”.

⁸ This property is shared with the associative composition of group theory and thus with arithmetical addition and multiplication – $a \times (b \times c) = (a \times b) \times c$. In contrast, the algebraic expression $a + b \times (c + d) \times f$ is certainly not to be identified with $(a + b) \times c + d \times f$; nor are the logical sequences $p \Rightarrow (q \Rightarrow r)$ and $(p \Rightarrow q) \Rightarrow r$ equivalent.

system of the Babylonians.⁹ A number $1'0,2'5$ (where separation “'” stands for a factor 10 and separation “,” for a factor 6) added to itself gives $2'0,5'0$, while $1,0'2,5'0 + 1,0'2,5'0 = 2,0'5,4'0$. Multiplications of course become even more bothersome, and root extractions virtually impossible if not reduced to an implicit sexagesimal system.

Such mixed place-value systems are rare in the historical record, though their non-place-value analogues are very common in pre-metric metrologies; best known while still in use are probably the Roman numerals, with their levels 1, 5, 10, 50, The place-value example that comes to my mind beyond the Babylonian system is the Maya calendar system, which is vigesimal except for the step that ensures a unit of 360 instead of 400 days – obviously a choice dictated by actual calendarian convenience.¹⁰ It is therefore not without interest that the earliest Latin *theoretical* exposition of “Arabic” reckoning – due to Jordanus of Nemore and from the early thirteenth century – proposes an analogue of this mixed system for fractions.¹¹ Instead of describing how to compute with the sexagesimal fractions currently used by astronomers (minutes, seconds, thirds, etc.), Jordanus introduces “consimilar fractions” (in modern symbols $\sum_i^n a_i \cdot p^{-i}$), for which the factor p by which each place decreases is constant; this is an obvious generalization of the sexagesimal fractions ($p = 60$) and also encompasses decimal fractions ($p = 10$) as another special case; it might seem rather empty if it had not gone together with the introduction of another category: “dissimilar fractions”, for which the factors of decrease vary (in modern symbols $\sum_i^n [a_i \times \prod_1^i p_j^{-1}]$). The “dissimilar fractions” correspond to the “ascending continued fractions” which were commonly used in Semitic languages (Arabic, but also Babylonian – see [Høyrup 1990]) – composite fractions of the type

⁹ Mostly, historians think of the Babylonian system as a “sexagesimal” system, a place-value system with base 60. That the Babylonians themselves understood the system rather (but not exclusively) as a decimal-seximal system in the Old Babylonian period (c. 2000 BCE to c. 1600 BCE) follows from the way “intermediate zeroes” are inserted in a text from Susa and from the way roundings are made [Høyrup 2002: 15 n.19, 263]. In the Seleucid epoch (third to second century BCE), some “intermediate zeroes” normally considered erroneous suggest that the same way of thinking still prevailed.

¹⁰ See [Closs 1986: 299–307]. To be precise, the irregularity of the system occurs in what can be regarded as the fractional part, calendarian distances being counted in the vigesimal place value notation in units of 360 days; below this interval, up to 17 units of 20 days and up to 19 single days are counted.

¹¹ See [Høyrup 1988: 337].

“one half, and two thirds of one half, and four sevenths of one third of one half”. In these, the successive denominators would be chosen *ad hoc*, and therefore had to be made explicit.¹²

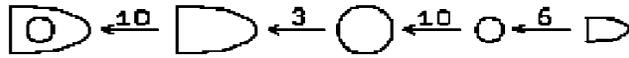
From the mathematical point of view, the dissimilar fractions constitute the general and the consimilar a “degenerate” case (sexagesimal and decimal fractions being one step further degenerate by having a predetermined and no general factor of decrease). In spite of this, the dissimilar fractions were never accepted by anybody apart from the inventor, and for good reasons: number notations are first of all tools for computation, and the criterion for acceptance is not mathematical generality or “beauty” but a compromise between computational ease and agreement with pre-existing number concepts and habits; given his non-Semitic linguistic environment, Jordanus erred on both accounts.

Place value systems constitute a special (elliptic) variant of multiplicative writings of numbers, and in this respect they correspond (ellipsis apart) to the normal way of expressing higher numerals in all languages which possess these. As spoken examples we may refer to English *sixty-four* (interpreting *-ty* as a variant of *ten*), corresponding to “types I/II”, and *two hundred sixty-four thousand three hundred and nineteen*, where the underlined part is close to the principle of “type III” (without sharing its inherent flexibility – we would never find **twenty-six myriads four thousand thirty-one-ty and nine*). In writing, multiplicative notations are known for example from younger Hieroglyphic and Middle Kingdom Hieratic Egyptian, where, respectively, 27,000,000 may be written as 270 below the sign for 100,000, and 40,000 as 4 written below the sign for 10,000 [Sethe 1916: 9]. In the Greek alphabetic notation we also find a variant related to “types I/II” – thus in Diophantos *Arithmetic* II.xxiv [ed. Tannery 1883: I, 121] Μ α.δχμ α meaning 1 (=α) myriad (Μ), 4 (=δ) thousand (,) and 641 (=χμ α). The Greek type is certainly an imitation of spoken numerals, which in Ancient Greek follow the same pattern; given the unpredictable level of the multiplicand, the Egyptian system is more likely to have been at least in part independent of spoken language.

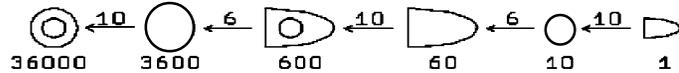
From here we may turn to the earliest beginning of writing in proto-literate Mesopotamia. Since the eighth millennium BCE, a system of characteristic tokens made of burnt clay had been used in the Near East, seemingly for accounting purposes – see, e.g., [Schmandt-Besserat 1992]. Some of these tokens (small and

¹² Either as here in words or (thus in late medieval Maghreb mathematics and in Fibonacci’s *Liber abbaci*) as $\frac{4\ 2\ 1}{7\ 3\ 2}$.

large cones and spheres) appear to have represented various standard containers (and thus measures) of grain, whereas flat circular discs probably stood for sheep and other livestock. With the advent of writing in the later fourth millennium, representations of the spheres and cones came to stand for standard units of grain. The relations were as follows [Damerow & Englund 1987: 136]:



The same signs (produced indeed by impression of the same particular stylus) were also used (presumably a *new* use) for (perhaps only “almost-abstract”) numbers, but with a different sequence of factors [Damerow & Englund 1987: 127], in which we recognize the decimal-seximal structure of the later place-value system:



We shall return anon to the reasons that these numbers should possibly be characterized as “almost-abstract” only. For the moment we observe first of all that the use of a factor sequence for the number series that differs from that of the grain measures cannot easily be explained without the assumption that at least the lower part of the number series rendered the structure of a pre-existent (and thus oral) numeral system. The writing of 600 is clearly meant multiplicatively, and corresponds to the structure of the Sumerian word for 600 (*geš.ù*, “sixty-ten”¹³); this word, however, is not attested until much later. Whether a spoken word for 3600 (with the appurtenant multiplicative word for 36000) was already in existence when writing was created is rather doubtful, as is the status of an early equivalent of *geš.ù* as a proper numeral; for certain purposes, indeed, written counting was based on a different “bisexagesimal” system with units 1, 10, 60, 120, 1200 and 7200¹⁴ [Damerow & Englund 1987: 132]. The spoken terms may therefore rather have been constructed from the written numerals; whether the decimal-seximal

¹³ I.e., ten sixties, Sumerian having postposited adjective and numeral.

¹⁴ The signs were written thus (increasing values toward the right):

structure of spoken numerals extended to three levels (from which a systematic unlimited expansion is easily derived) before the intervention of writing is thus quite dubious. It is a fair guess, in any case, that the utmost-left grain unit is a fresh emulation of the multiplicative structure of “normal” numerals, as are the writings of 1200 and 7200 in the “bisexagesimal” system.¹⁵

Before the invention of writing, not only grain accounting but also the counting of livestock was made “concretely”, 2 sheep (e.g.) being indicated by two sheep-discs. The introduction of numerals had as its purpose to change this, and in written accounts the same meaning was indicated by juxtaposition of a drawing of the sheep-disc and the numeral 2. In this sense, the numbers can be regarded as abstract. Two reasons suggest that we should perhaps add an “almost”. One is the existence of the bisexagesimal system. Since we do not understand the exact bureaucratic procedures within which it was used, we cannot say whether its existence has any implications for the number concept that make it less abstract than ours; it may imply nothing more than our habitual counting of wine bottles in dozens and, more recently, of bytes in units of 1024 and 1,048,576 spoken of for convenience as 1000 (k) and 1,000,000 (M). The other is the use of the numbers without reference to a unit when the dimensions of rectangles are indicated. This habit of leaving implicit a “basic” unit (in length measures the *nindan* or “rod” of c. 6 m) stayed alive for millennia in Mesopotamian mathematics, and can hardly be taken as evidence for failing understanding (we also tell that something happened on the third [day] of October [in year number] 1989 at 2 [hours] of [the] clock); but it does demonstrate that the users of the system did not feel it was compulsory to separate quantity systematically from quality. From a contemporary mathematical perspective it is tempting to see this as a symptom of “primitivity”, that is, of a not fully unfolded number concept – forgetting that even we omit the quality in certain cases where it is implied unambiguously by the context.

An interpretation of the place-value system in this light may result in unexpected insights. We may compare the system we actually use (now for

¹⁵ In the case of the grain units, however, this *is* nothing beyond a fair guess. Sometimes tokens are provided with a circular punching, which almost certainly gave them a specific meaning, possibly a larger value. The impression of a small circle in writing could be an emulation of this punching. Even if this should be the case, however, the precise meaning “ $\times 10$ ” may well have gone together with the creation of writing.

fractional numbers) in “type-II”-interpretation with what Stevin proposed in *La Disme* in 1585 [ed. Sarton 1935]: to write “our” 375.72 as 375 $\textcircled{0}$ 7 $\textcircled{1}$ 2 $\textcircled{2}$. As in the Maghreb notation for the “dissimilar” or “ascending continued” fractions or in the usual notation for angular minutes and seconds, the value of each place – its “quality” – is made explicit separately from its “quantity”, the digit. This makes explanation of the meaning more obvious but prevents easy recursiveness.

In “type-I” or “type-II” interpretation, our present notation already “recedes” into “primitivity” when compared with Stevin’s original proposal; in “interpretation III”, where unambiguous recursiveness can no longer be formulated, and where the same number may be interpreted as an embedding in several different equally valid ways (which, as pointed out, is the very reason that convenient algorithms can be formulated), we are even farther removed from any clear distinction between qualitative and quantitative dimensions (not to speak of making the structure explicit). The best linguistic analogue is the kind of contact language which speakers familiar with an ergative deep structure may conceptualize in *their* way, and which speakers whose mother tongue has an accusative deep structure may without difficulty understand as *they* are accustomed to.¹⁶

Historically, all place-value systems probably arose through transformation of preceding systems where the multiplicative structure was clear, that is, where digits multiplied values of identified places or their analogues.¹⁷ As in spoken language (where, to mention simple English examples, *thirteen* is found instead of **three-ten* and *twenty* instead of **twain ten* or **two tens*), such mathematical rigour is worn off in use. In both cases, “embedding” is a reconstructed deep structure, no longer (and perhaps never historically in complete form) a clean surface structure. Even in the case of spoken numerals, the deep structure is likely to be only a possible or at best a highly plausible reconstruction, not the only logical possibility.¹⁸

¹⁶ See [Silverstein 1971] on Chinook Jargon.

¹⁷ Such analogues may be columns in an abacus – or they may be the values of specific signs like the early Mesopotamian signs for 1, 10, 60 etc., in which case the “digits” are the fixed patterns in which specific numbers of such signs are organized.

¹⁸ Similar ambiguities can be found in other linguistic domains. I think in particular of the doubts whether it is meaningful to refer to a “verb phrase” (and thus to split the sentence into subject and predicate, the latter containing verb+object) in creole and certain other languages. For creoles, see [Bickerton 1981: 53 and *passim*]; for Dyirbal in its relation to related languages, [Dixon 1977: 382]; for Sumerian, [Gragg 1973: 91].

For historians of mathematics, these observations imply a moral: there is no reason to see the introduction of a place value system as an indubitable intellectual progress. For purposes of practical computation, the progress is not to be doubted; nor is it, indeed, in the Babylonian “primitive” deletion of standard units. But conceptual ambiguity – be it pragmatically adequate ambiguity – is not what mathematicians normally see as the aim of *their* specific enterprise.

Symbols and other symbols

In [1842: 302], Nesselmann proposed in his *Algebra der Griechen* a three-stage scheme for the history of algebra. His “first and lowest” stage is that of “rhetorical algebra”, in which everything in the calculation is explained in full words.¹⁹ The second, “syncopated algebra”, makes use of standard abbreviations for certain recurrent concepts and operations, even though “its exposition remains essentially rhetorical”. The third is “symbolic algebra”; here, “all forms and operations that appear are represented in a fully developed language of signs that is completely independent of the oral exposition”.

Al-Khwārizmī’s *Algebra* (from the early ninth century CE) is pointed out to represent (together with other Arabic works²⁰) the most consistent version of the rhetorical principle, as even numbers are written in full words. Iamblichos and “the oldest Italians and their disciples, for instance Regiomontanus” are counted in the same category in spite of their use of non-verbal numerals. Diophantos and later European algebra until the mid-seventeenth century is syncopated,²¹ “although already Viète has sown the seeds of modern algebra in his writings, which however only sprouted some time after him” (the following pages mention Oughtred, Descartes, Harriot and Wallis).

Nesselmann’s stages (or types) are regularly cited, even though many histories of mathematics interpret any use of abbreviations as “symbolization”. Worse, even those who cite him rarely notice Nesselmann’s main point: that symbolization allows

¹⁹ Here and in the followings, all translations into English are mine when nothing else is indicated.

²⁰ The introduction of syncopation in late Medieval Maghreb algebra was indeed only discovered by Franz Woepcke in [1854].

²¹ P. 304, n. 15 points out that some parts of Diophantos’s *Arithmetic* are written without any use of abbreviations, and are thus purely rhetorical.

operations directly on the level of the symbols, without any recourse to thought through spoken or internalized language.²² It may hence be of some value to rethink the scheme; as we shall see, our present framework is useful for that purpose.

In Diophantos's *Arithmetic*, we find symbols for the unknown number (the *arithmós*) and its powers, spoken of as "signs" (σημειου). The unknown itself is written with a simple sign, close to ς ; for all other powers (*dynamis* = ς^2 , *kybos* = ς^3 , *dynamodynamis* = ς^4 , etc.), phonetic complements are added to the symbol (Δ^r , K^r , etc.);²³ complements are also added to the sign for the monad ("power zero"), and for numbers that stand as denominators in fractions,²⁴ except when fractions are written in compact form ($\frac{5}{16}$ meaning $\frac{16}{5}$). Addition is indicated by juxtaposition, subtraction and subtractivity by \wedge (λειψις, "missing" etc.). One sign only is used for direct operation: the designation of the "part denominated by" n (better, the reciprocal of n , since non-integer n occur), explained in the introduction to be indicated by a sign \times for powers of the unknown. In III.xi, a number is posited to be ς^\times . When then ς turns out to be $\frac{77}{41}$ (that is, $\frac{41}{77}$), the number itself is stated immediately to be $\frac{41}{77}$ ($\frac{77}{41}$). Diophantos thus knows *at the level of symbols* (and

²² In Nesselmann's own words

We may execute an algebraic calculation from the beginning to the end in fully intelligible way without using one written word, and at least in simpler calculations we only now and then insert a conjunction between the formulae so as to spare the reader the labour of searching and reading back by indicating the connection between the formula and what precedes and what follows.

Since then, (mis)use of logical arrows in sequence has sometimes eliminated the conjunctions – but already because the arrows are misused, they are mere abbreviations and cannot serve as symbols in Nesselmann's sense (cf. note 8).

²³ One should remember that Diophantos wrote without distinguishing between capital and small letters. Apart from reminding of the phonetic reading, the complements thus also indicated that the symbols for the *dynamis* and the *kybos* were not to be read as 4 and 20, respectively.

It has been suggested (thus [Heath 1921: II, 457]) on the basis of the way the sign is written in Medieval manuscripts that the sign for the $\alpha\rho\iota\theta\mu\omicron\varsigma$ comes from a contracted $\alpha\rho$. However, the form in the papyrus P. Mich. 620 (probably from the early second century CE, and not known to Heath), *viz* ζ [Vogel 1930: 373], does not support this reconstruction.

²⁴ In I.23 [ed. Tannery 1893: I, 92], $50/23$ appears as $\bar{\nu}$ $\kappa\nu^{\text{ου}}$, "50 of 23rds", and $150/_{23}$ slightly later as "150 of the said part"

supposes his reader to understand) that $(\zeta^\times)^\times = \zeta$, and that $\left(\frac{q}{p}\right)^\times = \frac{p}{q}$.

From here we may jump to late medieval Italy. In Dardi of Pisa’s mid-fourteenth-century *Aliabrea argibra*²⁵ we find on fol. 3^r this explanation of how to multiply a square root (\mathbb{R}) by a number:

If you wish to multiply \mathbb{R} of number by number, as 6 times \mathbb{R} of 3, you should reduce the number to \mathbb{R} , that is 6, which makes 36, which you should multiply by 3, it amounts to \mathbb{R} of 108, and so much makes \mathbb{R} of 3 times 6 or 6 times \mathbb{R} of 3, that is \mathbb{R} of 108, which \mathbb{R} is surd [indiscreto], and so we prefer it as a surd number, that is, \mathbb{R} of 108.

\mathbb{R} of 3 — times 6 — that is \mathbb{R} of 36 times \mathbb{R} of 3 — makes \mathbb{R} of 108.

We notice that \mathbb{R} is used not only where we would use a mathematical symbol $\sqrt{\quad}$ but also in the discursive text, and that it is followed in all functions by the preposition “of” exactly as the fully written word “root”/ *radice* would be.

Somewhat closer to symbolization is the summary of the explanation of the multiplication of binomials *in croze* (that is, cross-multiplication; this example fol. 18^v):



Here, *c* stands for *cosa* (“thing”), that is, the first power of the unknown, and ζ for *censo*,²⁶ its second power; *e* means “and”. Drachmas are used for “power zero”. The scheme imitates one which is used to explain the multiplication $(10 - 2) \times (10 - 2)$ on fol. 4^v, itself modelled after the explanation of the cross-multiplication of two-digit numbers; it may thus be regarded as an extension of “type-I” embedding in which digits are replaced by algebraic monomials. Similar but more fully developed schemes are still found in Stifel’s *Arithmetica integra* from 1544 and other sixteenth-century works.

²⁵ On Dardi and his algebra, see [Van Egmond 1983]. I use the earliest extant manuscript Vatican, Chigi M.VIII.170, written in Venetian in c. 1395 (referring to the recent stamped foliation).

²⁶ In the fourteenth century Venetian dialect of the manuscript – probably that of the original – the writing of this word would probably have been *çenso* (but it never appears in full). In the fifteenth century, north eastern Italian dialects would mostly write *zenzo*.

Other treatises from Dardi's century go somewhat further, and write divisions by polynomials as fractions. Thus we find in *Trattato dell'alcibra amuchabile* [ed. Simi 1994: 42]:

$$\frac{100}{\text{per una cosa}} \quad \frac{100}{\text{per una cosa e più 5}}$$

corresponding to our $\frac{100}{x} + \frac{100}{x+5}$. The solution of the problem $\frac{100}{x} + \frac{100}{x+5} = 20$ is then explained verbally with reference to the operations performed on the symbolic expression in parallel to the addition $\frac{24}{4} + \frac{24}{6}$ (the aim being of course that the trained reckoner be able to operate directly on the formal fractions). Here, as we see, the places of numbers in a more intricate arithmetical expression may be taken over by algebraic polynomials.

The limits of this one-level embedding are illustrated by the way complex entities are expressed in Cardano's *Ars magna* from 1545 (this example from [Cardano 1545: 34^r]). What *we* would express as $\sqrt[3]{42 + \sqrt{1700}} + \sqrt[3]{42 - \sqrt{1700}}$ appears here as "R V: cubica 42 p: R1700 p. RV: cubica 42 m: R1700" – "p." representing *più*/"plus", "m." *meno*/"less", "R" "radice", and "V" (for *unita*/"united" or *universale*) indicating that the root is taken of two members. "RV" thus expresses that a two-member expression is embedded at the place of the radicand. However, the whole notation is so cumbersome that mental operation at the level of symbols is impossible; it facilitates writing but not understanding – as most algebraic syncopation it calls for a translation into the corresponding full verbal expression if one is to penetrate its mysteries. Only Bombelli's *L'algebra* from [1572] (claimed very adequately by the author to be primarily a rewriting in understandable form of what Cardano and other precursors had already done but set forth opaquely) introduces algebraic parentheses for composite radicands (written L. ... J, and used for multiple nesting) and an arithmetical notation for powers in which *J* represents our x^n . The latter notation is obviously akin in spirit to Stevin's almost contemporary notation for decimal fractions. The absence of a "placeholder" or manifest representative of the unknown makes it unfit both for operation with several unknowns and for embedding of a whole algebraic parenthesis at the place of an unknown; it can be seen to share the strength as

well as the weaknesses of a Stevinian notation for consimilar as opposed to a Fibonaccian notation for dissimilar fractions (see note 6). Bombelli thus provides one of the essential building stones for the creation of a fully symbolic algebra, but he uses it only for composite radicands and stops short of producing this algebra himself.

Strictly speaking, algebraic embedding does not begin with the incipient use of syncopation for symbolization purposes but rhetorically. In one problem of the algebra-chapter of Leonardo Fibonacci's *Liber abbaci* from 1228 [ed. Boncompagni 1857: 422], a *census* (the Italian *censo*) is re-baptized *res* ("thing", becoming *cosa* in Italian), which allows Leonardo to speak of its square as *census*. The same trick is inherent in a number of fifteenth- and sixteenth-century Latin and European vernacular terms for the higher powers; as an example chosen at random we may quote Pedro Nuñez's *Libro de algebra* [1567: 24]:

The first of those quantities which we call *dignidades*, which are ordered thus in proportion, is the *cosa*, which for this reason was given unity as denomination. The second is the *censo*, to which fell 2 as denomination. The third is the *cubo*, which has 3 as denomination. The fourth is the *censo de censo*, which has 4 as denomination. The fifth is called *relato primo*, whose denomination is 5. The sixth is *censo de cubo*, or *cubo de censo*, and its denomination is 6.

We find the same system in Luca Pacioli's *Summa* and Stifel's *Arithmetica integra*. Diophantos's system is different, however; here [ed. Tannery 1893: I, 6] the sequence is *arithmós*, *dýnamis*, *kýbos*, *dynamodýnamis*, *dynamokýbos*, *kybokýbos*. This is also found in al-Karajī's *Fakhrī* (Arabic list quoted in [Woepcke 1853: 48]) and in ibn Badr's *Recapitulation of Algebra* [ed., trans. Sánchez Pérez 1916: 18] – and still in Viète's *In artem Analiticen Isagoge* [ed. Hofmann 1970: 3].

The former system, as obvious also from the grammatical form "census of cube", etc., is built on embedding, the latter not. In the former system, the treatment of – say – second-degree problems where the unknown is itself a cube are therefore immediately seen to be reducible, and it is indeed equivalent to Leonardo's positing of a census as a thing; in the latter, reducibility has to be understood, it is not exhibited directly by the terminology.

But this embedding is of course rudimentary; it allows the easy treatment of biquadratics and similar problems, but does not permit that a power of the unknown

be replaced by a polynomial or other composite expression. Diophantos and al-Karajī treat the various powers as *entities*; in modern terms, Nuñez, Pacioli and Stifel have come to consider the power as a *function* or an operation, but only when the argument is another power – in other cases it remains an entity. The complete change of powers to being operations could only be effected when a generalization of Bombelli’s parenthesis function was combined with a convenient notation for powers. Only Descartes has both in his *Geometrie*.²⁷

The development of algebraic embedding thus turns out to go together with the full shift to symbolization in Nesselmann’s sense. This is not strange; only the development of certain symbolic notations allowed the unambiguous expression of embedding, and the avoidance of monsters corresponding to

$$(a+b^{k-[m]+n}\times c)-d$$

(in principle, punctuation in a written rhetorical exposition could serve, but a sufficiently consistent punctuation also did not exist in the sixteenth century – perhaps not even today). On the other hand, only the use of embedding made it possible to handle complex expressions so conveniently that computations could be made without recourse to extensive verbal explanations.

With this in mind, we may take a look at two notations that are neither rhetorical nor participants in the development toward Modern European symbolic algebra – first the Indian notation, which Nesselmann refers to as an earlier case of symbolic algebra.

As examples we may consider two equations from Bhāskara II (b. 1115) – see [Datta & Singh 1962: II, 31f]. An equation which in our terminology translates

$$5x+8y+7z+90 = 7x+9y+6z+62$$

is actually expressed]:

$$\begin{array}{cccc} y\hat{a} & 5 & k\hat{a} & 8 & n\hat{i} & 7 & r\hat{u} & 90 \\ y\hat{a} & 7 & k\hat{a} & 9 & n\hat{i} & 6 & r\hat{u} & 62 \end{array}$$

²⁷ Descartes did not need to combine the two, and therefore does not do so; what he needs and uses repeatedly are polynomial parentheses multiplied by powers of a variable. But Descartes made the tools available for those who were to need them for wider purposes in the following generations.

whereas our

$$8x^3+4x^2+10y^2x = 4x^3+0x^2+12y^2x$$

corresponds to]:

$$\begin{array}{r} y\hat{a} gha 8 \quad y\hat{a} va 4 \quad k\hat{a} va y\hat{a}.bh\hat{a} 10 \\ y\hat{a} gha 4 \quad y\hat{a} va 0 \quad k\hat{a} va y\hat{a}.bh\hat{a} 12 \end{array}$$

As Datta and Singh quote David Eugene Smith [1923: II, 425f], this notation is “in one respect [...] the best that has ever been suggested” because it “shows at a glance the similar terms one above the other, and permits of easy transposition”. It corresponds well to the single-level embedding of place-value numbers in Stevin’s notation.²⁸

What it does not permit is direct multiple embedding, for instance replacement of $y\bar{a}$ by a polynomial. Nor is it intended for that, it is only used for linear reductions of equations, as one will discover looking at the context of Bhāskara’s formula in the *Vijā-ganita* [trans. Colebrooke 1817: 248]; the rest of Bhāskara’s argument is syncopated, though more systematic in its use of abbreviations than Diophantos.

Indian schemes (if not Bhāskara’s whole text, but the same could be said about Viète) are thus justly seen as a symbolic notation by Nesselmann; but Smith is right that it is the best “in one respect” only – namely within the restricted framework of problems actually dealt with by Bhāskara and his fellows, and for the even more restricted use made of it within this framework; it was not open-ended. In this way, it presents us with a parallel to the “short-coming” of the place-value system for fractions as compared with Fibonacci’s more flexible and more explicit but less handy notation for dissimilar fractions.

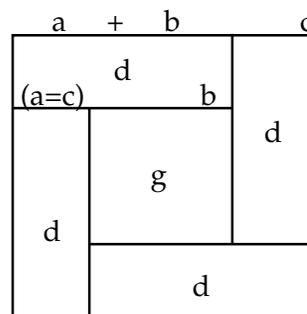
The other example is European, and borrowed from Jordanus of Nemore’s *De numeris datis*, written somewhere around 1220–30. The work is a quasi-Kantian critique of the procedures of algebra, modelled after Euclid’s *Data*; it tries so to speak to demonstrate that what is currently done “empirically” in Arabic and post-Arabic *al-jabr* can be made by theoretically legitimate methods based on arithmetical theory – see [Høystrup 1988] and [Puig 1994]. I translate one of the propositions from the Latin text in [Hughes 1981: 58] (the diagram is added in the interest of intelligibility, in agreement with the exposition in [Puig 1994] – nothing similar is

²⁸ One may add that the use of abbreviations for unknowns, powers and operations prevents that arithmetization of the designation of powers which reduces the multiplication and division of powers of the unknown to a purely formal process.

found in the original²⁹):

If a given number is divided into two and if the product of one with the other is given, each of them will also be given by necessity.

Let the given number abc be divided into ab and c , and let the product of ab with c be given as d , and let similarly the product of abc with itself be e . Then the quadruple of d is taken, which is f . When this is withdrawn from e , g remains, and this will be the square on the difference between ab and c . Therefore the root of g is extracted, and it will be b , the difference between ab and c . And since b will be given, c and ab will also be given.



The working of this is easily verified in the following way. For instance: Let 10 be divided into two numbers, and let the product of one with the other be 21, whose quadruple is the same as 84, it is taken away from the square of 10, that is, from 100, and 16 remains whose root is extracted, which will be 4, and that is the difference. It is taken away from 10 and the remainder, that is, 6, is halved; The half will be 3, and this is the minor part, and the major is 7.

Not uncommonly, the use of letters have made interpreters see this as an early instance of symbolic algebra. Nothing could be more mistaken – in terms of the caption of the present section, these letters are indeed “other symbols”. The letters serve to make the argument general, and are thus a parallel to the line segments of geometrical demonstrations. But the argument cannot be made by manipulations of the symbols, in particular because every new step is expressed in new symbols that have to be identified verbally; even that rudimentary embedding is avoided which consists in conserving the name $4d$ for the outcome of a multiplication of d by 4.

It is of some interest that Jordanus’ letter notation in these proofs may have been inspired by the algorithms for computation with place-value numbers in type-I interpretation – see [Høyrup 1988: 337]. When presenting demonstrations for these algorithms, Jordanus uses letters for digits, not for numbers; they thus represent the place in which the digit has to be inserted. In this sense embedding is of course also a feature of the proofs of *De numeris datis* – the letters represent places where

²⁹ Whether Jordanus thought of something similar is uncertain but possible (the failure to point out at first that a is meant to equal c might suggest that this was evident from a diagram); in any case the diagram may be helpful for a modern reader. The proof when read in the context of the treatise as a whole does not need it: even though there are no explicit references, the unexplained jumps build on propositions that are proved earlier on in Jordanus’s *Elements of Arithmetic*.

any number can be inserted instead of the letter; but embedding in this sense is inherent in any attempt to formulate an argument or statement in general terms, be it arithmetical, geometrical, or ethical – cf. above, after note 2.

Embedded theoretical domains?

In 1968, Raymond L. Wilder formulated as the last of 10 “laws’ governing the evolution of mathematical activity” that

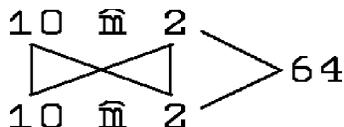
Mathematical evolution remains forever a continuously progressing affair limited only by the contingencies [of] the opportunities for diffusion, such as may be provided by a universally accepted symbolism, increased outlet for publications, and other means of communication[, the] Needs of the host culture [and the long-term stifling effects of a] static cultural environment [or] an adverse political or general anti-scientific atmosphere

(quoted from the reprint [Wilder 1978: 200f]). In a similar list of ten “laws”, Michael J. Crowe [1975: 165f] also proposed as the tenth that “revolutions never occur in mathematics”, in the sense that no previously accepted entity is ever “overthrown [or] irrevocably discarded”. He gave as an example that “Euclid was not deposed by, but reigns along with, the various non-Euclidean geometries”, and added that his law does not preclude the existence of revolutions in “mathematical nomenclature, symbolism, metamathematics (e.g. the metaphysics of mathematics), methodology (e.g. standards of rigor), and perhaps even in the historiography of mathematics”. In contrast, the shift from Ptolemaic to Copernican astronomy is seen as a genuine revolution.

Comparable formulations abound, not least among mathematicians reflecting on the history of their field. They do not deny that new things happened when (e.g.) the range of numbers was enlarged so as to encompass complex numbers or quaternions or when non-Euclidean geometries were accepted in the nineteenth century; but they go on to tell that quaternions contain the complex numbers as a subset, for which the usual arithmetic of complex numbers holds good; that complex numbers encompass real numbers as a subset, for which ...; ... and that the integers contain the positive integers as a subset for which the rules of *Elements* VII–IX remain valid; similarly, it is normally held, not that Euclidean geometry “reigns along with the various non-Euclidean geometries” but that all of these turn out to be special cases of an all-encompassing geometry.

The alleged absence of revolutions in mathematics is thus explained as an embedding of old theories within more general frameworks. It could be added that

the possibility of embedding older domains has often been used as a defining condition when the scope of mathematical theories was widened. One case was already referred to above – a very simple case, which is not suspect of being motivated by explicit concerns for embedding. When Dardi of Pisa wants to show (fol. 4^v) that $(-2) \times (-2)$ *must* be +4, he uses this scheme:



The idea is that cross-multiplication should still be valid, just as if we multiply “10 plus 2” by itself. Obviously, 10×10 is still 100, and $10 \times (\text{less } 2)$ as well as $(\text{less } 2) \times 10$ must reasonably be “less 20”. Since the whole product (8×8) has to be 64, $(\text{less } 2) \times (\text{less } 2)$ must hence be +4 – it can be neither -4 nor 0, since then the whole product would be 56 or 60.

To claim that no revolutions occur in mathematics amounts to asserting that *all* theoretical shifts in mathematics either consist in such embeddings of the old within something larger or in the addition of new fields of interest, corresponding to the addition of spectroscopy to existing physics (this latter example is given in [Crowe 1975: 165]).

The current way to prove this claim is a *petitio principii*. If one wants, for instance, to prove that the geometry of *Elements* II is “covert algebra”³⁰ and thus isomorphic with a substructure of modern (or Cartesian) algebraic theory, then he has to strip the text of all those features that do not fit the claim, by declaring them to be non-essential and mere results of the unfortunate limitations of the framework within which the ideas had to be expressed. In the actual case this not only implies that we take it to be non-essential that Euclid’s theorems deal with equalities of areas and lengths and not with numbers (which could still be defended by the observation that areas and lengths can be mapped isomorphically onto the set of positive real numbers) but also that we neglect the fact that propositions 5 and 6 are algebraically though not geometrically identical (we just have to switch some

³⁰ Thus Hans Freudenthal’s characterization of *Elements* II.5 [1976: 189].

names), as are propositions 9 and 10 (4 and 7 are so if we use proposition 1).³¹

Similar arguments could be used in many other cases. On the level of whole theoretical domains, embedding thus does not describe the actual historical process, since what is embedded is *not the conceptual network of the old theory but a substructure of the new theory itself* which has some superficial similarity with *certain features* of the old theory; in the best cases it is homomorphic with *those features of the old theory which the new theory wants to conserve*. This is no different from the conservation of epicycles in Copernicus's theory, the conservation of Copernicus's heliocentricity in Kepler's, or Newton's conservation of Kepler's idea that the same physics should hold true below and above the moon. From this perspective there was thus no revolution in early Modern astronomy.

All in all, the purported protective embedding of everything once made by earlier mathematicians by their successors is rather an expression of the *prevalent ideology* of mathematicians – and probably not of an intra-scientific ideology alone. In a survey of the political opinions of US university faculty, Everett Carll Ladd and Seymour Lipset [1972: 1092] found mathematicians to be somewhat more conservative than physicians, considerably more than physicists, and far more than teachers of the social sciences, the humanities, and even law. Probably, intra-scientific and extra-scientific ideologies reinforce each other. Thomas Kuhn once

³¹ In symbolic translation, *Elements* II.1–10 can be expressed as follows ($\square(a)$ stands for the square on the segment a , and $\square\square(p,q)$ for the rectangle contained by p and q):

1. $\square\square(a, p+q+\dots+t) = \square\square(a, p) + \square\square(a, q) + \dots + \square\square(a, t)$.
2. $\square(a) = \square\square(a, p) + \square\square(a, a-p)$.
3. $\square\square(a, a+p) = \square(a) + \square\square(a, p)$.
4. $\square(a+b) = \square(a) + \square(b) + 2\square\square(a, b)$.
5. $\square\square(a, b) + \square\left(\frac{a-b}{2}\right) = \square\left(\frac{a+b}{2}\right)$
6. $\square\square(a, a+p) + \square\left(\frac{p}{2}\right) = \square\left(a + \frac{p}{2}\right)$.
7. $\square(a+p) + \square(a) = 2\square\square(a+p, a) + \square(p)$.
8. $4\square\square(a, p) + \square(a-p) = \square(a+p)$.
9. $\square(a) + \square(b) = 2[\square\left(\frac{a+b}{2}\right) + \square\left(\frac{b-a}{2}\right)]$.
10. $\square(a) + \square(a+p) = 2[\square\left(\frac{p}{2}\right) + \square\left(a + \frac{p}{2}\right)]$.

If b is replaced by $a+p$ in propositions 5 and 9, propositions 6 and 10 result; if b is replaced by p in proposition 4, and if we use proposition 1 to show that $\square(a) + \square\square(a, p) = \square\square(a, a+p)$, proposition 7 results when $\square(a)$ is added to both sides. Application of similar small modifications will show that all propositions 4–10 if seen as algebraic identities are trivially equivalent.

stated [1963: 368] that scientists, though “*trained* to operate as puzzle-solvers from established rules, [...] are also *taught* to regard themselves as explorers and inventors who know no rules except those dictated by nature itself”. But mathematicians, as we see, are often not taught so; instead, they learn that progress in their field has always consisted in “changing in order to conserve”, in agreement with a famous slogan of political conservatism.

Embedding and spatiality

Charles Darwin emphasized that evolution often makes use of existing organs which are put to new use. One example is the swim bladder, which in certain fish was so well furnished with veins that it could serve for supplementary breathing; when adequate circumstances occurred, selection pressure gave rise to the development of genuine lungs.

Obviously, the human language faculty has made use of a pre-existing organ – namely the brain. We may ask which specific “organ” within the brain was made use of, but since many brain centres are involved in the use of language, no exhaustive answer is likely to emerge.

If we ask for syntax alone, however, at least a partial answer exists. According to Ron Wallace [1989: 519f], “In all mammals except humans, both sides of the hippocampus are cognitive-mapping structures³² [...]. In humans, the right hippocampus is specialized for mapping, the left for the production of verbal material”. The evidence that the “cognitive-mapping system could function as a deep structure for language” is analyzed in [O’Keefe & Nadel 1978: 381–410].³³ Moreover, in almost all cases where the origin of grammatical case systems can be traced, they derive from frozen spatial metaphors – see [Anderson 1971]: something was done *by* X, that is, when X was close by and therefore probably responsible; and it was done *for* Y, that is, in front of Y, and hence probably for Y’s sake. Other languages, not least those where case is grammaticalized as inflection, underscore the point – see also [Anderson 2006: 115–148].

³² [As explained in the abstract of the article (p. 518), “cognitive maps” are neurological models of space, which “probably characterize all mammal species. The human cognitive map appears to be unique, however, in being closely related to communication”. / JH]

³³ Cf. also a commentary by Ron Wallace in [Burling 1993: 43f], and William Calvin’s arguments [1983: 121] that “enhanced throwing skill could have produced a strong selection pressure for any evolutionary trends that provided additional timing neurons. This enhanced timing circuitry may have developed secondary uses for language reception and production”.

It thus seems fully justified to speak of the “embedding” for instance of relative clauses, using metaphorically a term whose genuine meaning is spatial (namely to place something within another thing or material): the syntactical operations in language really seem to imitate spatial activity.³⁴

In [1980: 248f] Noam Chomsky suggested in passing that “certain forms of mathematical understanding – specifically, concerning the number system, abstract geometrical space, continuity, and related notions” belonged, along with the language faculty, to a set of domains in which “humans seem to develop intellectual structures in a more or less uniform way on the basis of restricted data”. Already in [1975], James Hurford analyzed numerals from the point of view of generative grammar, in what he later characterized as “a paradigm example of Kuhnian normal science” [Hurford 1987: 43].

In this last-mentioned work, Hurford was led (p. 305f) by analysis of universals and universal irregularities in the formation of number systems and comparison with other features of language to the conclusion that only one of a list of five innate contributions to the number faculty – namely the “Cardinality Principle”, the “disposition to make the sizeable leap from a memorized sequence of words to the use of these words expressing the cardinality of collections”³⁵ – “is special to numeral systems; the rest are very familiar in human language more generally”; further, that these five

innate capacities are sufficient to determine the number faculty in Man, but insufficient to determine the universal morphosyntactic peculiarities found in the human linguistic systems that express number. Man has the capacity for language and for number, capacities which his ancestors at some stage lacked. Children, born with the capacity to acquire language and number, acquire them simultaneously, and this simultaneity is significant.³⁶ Language is the mental tool by which we exercise control over numbers. Without language, no numeracy. [...]. The capacity to reason about particular numbers, above about 3, comes to humans only with language.

³⁴ According to scattered observations by Piaget, it is also “a characteristic of operatory thought that it achieves at the level of thought the same decentration, reversibility, and composability which was achieved at the sensori motor level during the second year of life”, as summed up in [Høyrup 2000: 249].

³⁵ To the last point one should perhaps add the qualification that “Every language has a numeral system of finite scope” [Greenberg 1978: 253]. In contrast to what occurs in certain written number systems (not least in place value systems), the recursiveness in any actual spoken language, though possibly high, is never unlimited.

If the human number faculty itself is largely a by-product of innate linguistic capacities, the linguistic subsystems dealing with number are shaped by further principles, which are not innate in individuals. The two main such principles are:

Languages and their subsystems grow gradually over time. Their structures exhibit traces of this growth in the form of discontinuities and irregularities.

Pragmatic factors make certain forms favoured for communication and such pragmatic preferences become grammaticalized, that is regarded by new acquirers as having the status of grammatical rules.

So, according to Hurford, the understanding of the possibility to count a collection of items (the “cardinality principle”) may be a language independent universal (where he overlooks the integrations with ordinality, cf. note 36); but in his view the embedding involved in the construction of higher numerals is transferred from the corresponding structure in general language, and has no independent status.

If we regard number systems alone, it is indeed close at hand to regard their pragmatic characteristics (*thirteen* instead of **three-ten*, Italian *sedici* for 16 but *diciasette* for 17, etc.) as manifestations of features that characterize language in general. Possibly, this analysis might even be projected upon written number systems and abacus-type representations and their kin – for instance, the cancellation of Stevin’s place identifications might be seen as an analogue of the English deletion of the relative pronoun when it occurs in object position.

However, the inclusion of symbolic algebra in the panorama suggests a somewhat different interpretation. As we have seen, it is exactly when symbolism leaves language efficiently behind that it develops the capability of multiple embedding. Moreover, this embedding refers *directly* to spatiality. This is true of the new root sign that replaced \mathbb{R} , and which allows that we write

$$\sqrt[3]{42 + \sqrt{1700}} + \sqrt[3]{42 - \sqrt{1700}}$$

instead of Cardano’s “ \mathbb{R} . V: cubica 42 p: \mathbb{R} 1700 p. \mathbb{R} V: cubica 42 m: \mathbb{R} 1700”; but it is also true of the various types of parentheses, which all suggest an actual enclosing –

³⁶ [Actually this simultaneity should probably be formulated differently. Both according to Piaget’s results and my own observations, the integration of cardinality and ordinality, the certainty that loops are not permitted in the number jingle, and the immediate rejection of repetition of the same item twice in counting, only turn up around the age of five to six. Language, of course, is acquired before; but the acquisition of recursive syntax, not least the use of relative clauses, occurs around the same time [Romaine 1988: 232ff]. / JH]

not only our modern (), [], { } and < > but also Bombelli's L. ... J³⁷ and Descartes' $\left. \begin{matrix} +p \\ -q \\ +r \end{matrix} \right\}$. No mathematician ever had the idea to enclose something in)...(or)...{ or anything similar. The use of < and > for "smaller than" and "larger than" are also directly linked to that possibility of repeated embedding which corresponds to the transitivity (and actual spatial meaning) of the relations. Other symbols derive from abbreviations (e.g. Σ and \int for "sum", δ and Δ for "difference"), but it appears that symbols that have a spatial interpretation are directly iconic, and that their character is in disagreement with Saussure's principle that the relation between the linguistic signifier – the actual shape of a word – and its signification is generally arbitrary.³⁸ Mathematical symbolism seems to be tied *directly* to our capacity for processing spatial information, and not only indirectly through our syntactic capacity.

Seen in this light, even the number faculty may be less subordinated to the language faculty than concluded by Hurford for anything going beyond the cardinality principle, and connected to spatial connection directly and not only through the mediation of general language. It might perhaps be involved with "abstract geometrical space, continuity, and related notions", as suggested by Chomsky.³⁹

These (at least partially distinct) couplings of language, number and algebraic symbolization to our faculty for processing spatial information suggests that the

³⁷ Actually, Bombelli's manuscript shows that he intended the even more explicit $\lfloor \quad \rfloor$ with embedding $\lfloor \lfloor \quad \rfloor \rfloor$, but this was asking for more than the typesetter would accept. See the reproduction of a manuscript page in [Bombelli 1966: xxxiii, fig. 2].

³⁸ See the various contributions to [Haiman 1985] for examples of similar iconic exceptions to the general rule in the domain of syntax proper.

³⁹ In an article published after the original version of the present paper was prepared [Hauser, Chomsky & Fitch 2002], Chomsky and two collaborators go even further, reaching a position close to what is suggested here. As summarized in the abstract, they hypothesize that the FLN, the "language faculty in the narrow sense" (which excludes such things as the general sensori-motor and the conceptual-intentional systems) "only includes recursion [which is furthermore] the only uniquely human component of the faculty of language. We further argue that FLN may have evolved for reasons other than language, hence comparative studies might look for evidence of such computations outside of the domain of communication (for example, number, navigation, and social Relations)".

Recursion, as also evident in the above formalizations of the placevalue system, is closely related to embedding, in particular to multiple embedding (as also made clear in [Hauser, Chomsky & Fitch 2002: 1577]).

shared notion of “embedding” is *more* than a gratuitous metaphor in as far as these three domains are concerned. The “embedding” of theories, however, even in cases where it describes real generalization and is no mere expression of conservative ideology, is not easily linked to spatiality proper, and should probably be understood as a different phenomenon. It should instead be linked to Eugene Wigner’s famous “Unreasonable Effectiveness of Mathematics in the Natural Sciences” [1960] – both show that there is some kind of objective truth in mathematics, which does not coincide with the single theory but conditions it.

Bibliography

- [1] Anderson, John M. 1971. *The Grammar of Case. Towards a Localistic Theory*. Cambridge: Cambridge University Press.
- [2] Anderson, John M., 2006. *Modern Grammars of Case*. Cambridge: Cambridge University Press.
- [3] Bickerton, Derek, 1981. *The Roots of Language*. Ann Arbor: Karoma Publishers.
- [4] Bombelli, Rafael, 1572. *L’Algebra*. Bologna: Giovanni Rossi.
- [5] Bombelli, Rafael, 1966. *L’algebra*. 1. ed. integrale. Intro. di U. Forti, pref. di E. Bortolotti. Milano: Feltrinelli.
- [6] Boncompagni, Baldassare (ed.), 1857. *Scritti di Leonardo Pisano matematico del secolo decimoterzo*. I. Il *Liber abbaci* di Leonardo Pisano. Roma: Tipografia delle Scienze Matematiche e Fisiche.
- [7] Burling, Robbins, 1993. “Primate Calls, Human Language, and Nonverbal Communication”. *Current Anthropology* 34, 25-37, discussion 37-53, 165-170.
- [8] Calvin, William H., 1983. “A Stone’s Throw and its Launch Window: Timing Precision and its Implications for Language and Hominid Brains”. *Journal of Theoretical Biology* 104, 121–135.
- [9] Cardano, Girolamo, 1545. *Artis magna sive de regulis algebraicis, liber unus*. Nürnberg: Petreius.
- [10] Chomsky, Noam, 1980. *Rules and Representations*. Oxford: Blackwell.
- [11] Closs, Michael P., 1986. “The Mathematical Notation of the Ancient Maya”, pp. 291–369 in Michael P. Closs (ed.), *Native American Mathematics*. Austin: University of Texas Press.
- [12] Colebrooke, H. T. (ed., trans.), 1817. *Algebra, with Arithmetic and Mensuration from the Sanscrit of Brahmagupta and Bhascara*. London: John Murray.

- [13] Crowe, Michael J., 1975. "Ten 'Laws' Concerning Patterns of Change in the History of Mathematics". *Historia Mathematica* 2, 161–166.
- [14] Damerow, Peter, & Robert K. Englund, 1987. "Die Zahlzeichensysteme der Archaischen Texte aus Uruk", Kapitel 3 (pp. 117–166) in M. W. Green & Hans J. Nissen, *Zeichenliste der Archaischen Texte aus Uruk*, Band II (ATU 2). Berlin: Gebr. Mann.
- [15] Datta, Bibhutibhusan, & Avadhesh Narayan Singh, 1962. *History of Hindu Mathematics. A Source Book*. Parts I and II. Bombay: Asia Publishing House. ¹Lahore: Motilal Banarsidass, 1935–38.
- [16] Dixon, R. M. W., 1977. "The Syntactic Development of Australian Languages", pp. 365–415 in Charles N. Li (ed.), *Mechanisms of Syntactic Change*. Austin: University of Texas Press.
- [17] Freudenthal, Hans, 1977. "What is Algebra and What Has It Been in History?" *Archive for History of Exact Sciences* 16, 189–200.
- [18] Gragg, G. 1973. "Linguistics, Method, and Extinct Languages: The Case of Sumerian". *Orientalia*, NS 42, 78–96.
- [19] Greenberg, Joseph H. 1978. "Generalizations about Number Systems", pp. 249–295 in Joseph H. Greenberg (ed.), *Universals of Human Language*. Vol. 3. *Word Structure*. Stanford: Stanford University Press.
- [20] Haiman, J. (ed.), 1985. *Iconicity in Syntax*. Amsterdam: John Benjamins.
- [21] Hauser, Marc D., Noam Chomsky & W. Tecumseh Fitch, 2002. "The Faculty of Language: What is It, Who Has It, and How Did It Evolve". *Science* 298 (2002), 1569–1579.
- [22] Heath, Thomas L., 1921. *A History of Greek Mathematics*. 2 vols. Oxford: The Clarendon Press.
- [23] Hofmann, Joseph E. (ed.), 1970. François Viète, *Opera mathematica recognita à Francisci à Schooten*. Hildesheim & New York: Georg Olms Verlag.
- [24] Høyrup, Jens, 1988. "Jordanus de Nemore, 13th Century Mathematical Innovator: an Essay on Intellectual Context, Achievement, and Failure". *Archive for History of Exact Sciences* 38, 307–363.
- [25] Høyrup, Jens, 1990. "On Parts of Parts and Ascending Continued Fractions". *Centaurus* 33, 293–324.
- [26] Høyrup, Jens, 2000. *Human Sciences: Reappraising the Humanities through History and Philosophy*. Albany, New York: State University of New York Press.
- [27] Høyrup, Jens, 2002. *Lengths, Widths, Surfaces: A Portrait of Old Babylonian Algebra and Its Kin*. New York: Springer.

- [28] Hughes, Barnabas B., O.F.M. (ed. trans.), 1981. Jordanus de Nemore, *De numeris datis*. A Critical Edition and Translation. Berkeley etc. : University of California Press.
- [29] Hurford, James R. 1975. *The Linguistic Theory of Numerals*. Cambridge: Cambridge University Press.
- [30] Hurford, James R. 1987. *Language and Number*. Oxford: Blackwell.
- [31] Kuhn, Thomas S. 1963. "The Function of Dogma in Scientific Research", pp. 347–369 in A. C. Crombie (ed.), *Scientific Change*. London: Heinemann.
- [32] Ladd, Everett Caril, Jr. & Seymour Martin Lipset, 1972. "Politics of Academic Natural Scientists and Engineers". *Science* **176**, 1091–1100.
- [33] Nesselmann, G. H. F. 1842. *Versuch einer kritischen Geschichte der Algebra*. Nach den Quellen bearbeitet. Erster Theil, *Die Algebra der Griechen*. Berlin: G. Reimer.
- [34] Nuñez, Pedro, 1567. *Libro de Algebra en Arithmetica y Geometria*. Anvers: En casa de los herederos d'Arnaldo Birckman.
- [35] O'Keefe, John, & Lynn Nadel, 1978. *The Hippocampus as a Cognitive Map*. Oxford: Clarendon Press.
- [36] Puig, Luis, 1994. "El *De Numeris Datis* de Jordanus Nemorarius como sistema matemático de signos". *Mathesis* **10**, 47–94.
- [37] Romaine, Suzanne, 1988. *Pidgin and Creole Languages*. London & New York: Longman.
- [38] Sánchez Pérez, José A. (ed. trans.), 1916. *Compendio de algebra de Abenbéder*. Texto arabo, traduccion y estudio. Madrid: Centro de Estudios Históricos.
- [39] Sarton, George, 1935. "The First Explanation of Decimal Fractions and Measures (1585). Together with a History of the Decimal Idea and a Facsimile (no. XVII) of Stevin's Disme". *Isis* **23**, 153–244.
- [40] Schmandt-Besserat, Denise, 1992. *Before Writing*. I. *From Counting to Cuneiform*. Austin: University of Texas Press.
- [41] Sethe, Kurt, 1916. *Von Zahlen und Zahlworten bei den Alten Ägyptern, und was für andere Völker und Sprachen daraus zu lernen ist*. Straßburg: Karl J. Trübner.
- [42] Silverstein, Michael, 1971. "Language Contact and the Problem of Convergent Generative Systems: Chinook Jargon", pp. 191–192 in D. Hymes (ed.), *Pidginization and Creolization of Languages*. Cambridge: Cambridge University Press.
- [43] Simi, Annalisa (ed.), 1994. Anonimo (sec. XIV), *Trattato dell'alcibra amuchabile* dal Codice Ricc. 2263 della Biblioteca Riccardiana di Firenze. Siena: Servizio Editoriale dell' Università di Siena.

- [44] Smith, David Eugene, 1923. *History of Mathematics*. I. *General Survey of the History of Elementary Mathematics*. II. *Special Topics of Elementary Mathematics*. Boston: Ginn, 1923–1925. Reprint New York: Dover.
- [45] Tannery, Paul (ed. trans.), 1893. *Diophanti Alexandrini Opera omnia cum graecis commentariis*. 2 vols. Leipzig: Teubner, 1893–1895.
- [46] Van Egmond, Warren, 1983. “The Algebra of Master Dardi of Pisa”. *Historia Mathematica* **10**, 399–421.
- [47] Vogel, Kurt, 1930. “Die algebraischen Probleme des P. Mich. 620”. *Classical Philology* **25**, 373–375.
- [48] Wallace, Ron, 1989. “Cognitive Maps and the Origin of Language and Mind”. *Current Anthropology* **30**, 518–526.
- [49] Wigner, Eugene, 1960. “The Unreasonable Effectiveness of Mathematics in the Natural Sciences”. *Communications in Pure and Applied Mathematics* **13**:1, 1–14.
- [50] Wilder, Raymond L., 1978. *Evolution of Mathematical Concepts*. Milton Keynes: Open University Press. 1968.
- [51] Woepcke, Franz, 1853. *Extrait du Fakhri, traité d’algèbre par Aboû Bekr Mohammed ben Alhaçan Alkarkhî; précédé d’un mémoire sur l’algèbre indéterminé chez les Arabes*. Paris: L’Imprimerie Impériale.
- [52] Woepcke, Franz, 1854. “Recherches sur l’histoire des sciences mathématiques chez les Orientaux, d’après des traités inédits arabes et persans. Premier article. Notice sur des notations algébriques employées par les Arabes”. *Journal Asiatique*, 5^e série **4**, 348–384.

Contact Details:

Prof. Jens Høyrup

Section for Philosophy and Science Studies

Roskilde University 03.1.3

Box-260

DK-4000, Roskilde

Denmark

E-mail: jensh@ruc.dk

Generalization or widening or extension of meaning is a process contrary to specialization. In such cases the meaning of a word becomes more general in the course of time. It means that initially the word had only one lexical meaning and was used in a narrow sense but then got wider sense and more than one lexical meaning. Due to it such words usually become more frequently used. This research has helped me to understand the semantic structure of English words at the present stage of their development. This development is always a source of quantitative and qualitative development of the vocabulary. The constant development of trade, intercultural communication, technology, medicine, etc. usually brings new notions into our life. But languages do not have words to name them. If their students are to develop mathematical proficiency, teachers must have a clear vision of the goals of instruction and what proficiency means for the specific mathematical content they are teaching. They need to know the mathematics they teach as well as the horizons of that mathematics—where it can lead and where their students are headed with it. appropriate ways to respond. A teacher must interpret students' written work, analyze their reasoning, and respond to the different methods they might use in solving a problem. Teaching requires the ability to see the mathematical possibilities in a task, sizing it up and adapting it for a specific group of students. Ameliorative development of meaning or amelioration, or elevation is an improvement of the connotational component of meaning. It is a semantic shift undergone by words due to their referents coming up the social scale. For example, OE *cwān* 'a woman' > MnE queen; OE *cniht* 'a young servant' > MnE knight. The meanings of these words have been elevated through associations with aristocratic life. Polysemy is inherent in the very nature of words and notions, as they always contain a generalization of several traits of the object. Some of these traits are common with other objects, hence the possibility of identical names for objects possessing common features. Thus polysemy is characteristic of most words in many languages.