FROM LINGUISTIC REPRESENTATION TO FUZZY MATHEMATICS IN GROWN UP PEOPLE

Maria Antonietta Lepellere
Livio Clemente Piccinini
Mario Taverna
DI4A, University of Udine, Italy

Abstract. The aim of this note is to give some critical examples where even the use of the same clustering rules lead to fuzziness. It starts from poor numerical systems and compares them with the expanded Sergeyev model, where the grossone is used, as an infinite terminal element. It can be compared with terminal elements of the ancient languages, such as the Greek myriad and the Chinese wan. On them some propositions that hold in the arithmetic of the grossone are similar, while they are not meaningful for the countable system of infinity. The note shows that both the upward and downward trend are actually present in human language and in conceptual arrangements.

The note then goes on to sketch the model of evolution of Bak-Sneppen, showing two significant applications: the case of the evolution and study of foreign languages and, according to the model of Lloyd, the territorial analysis. In both cases it is highlighted how the Bak-Sneppen model becomes more stable when the universe is segmented, as already proven by the authors in previous works. The third part examines some cases of false probabilistic intuition due to incomplete perception of the phenomena, what could therefore be defined as hidden conditional probability. Interesting is the classic application of the theory of games to lotteries and ternary games, such as Chinese morra.

Keywords: Infinity, Grossone, Bak-Sneppen model, Conceptual granularity and translation, Lloyd’s problem of clustering, Probabilistic delusions.

Introduction

The authors some years ago discussed the problem of translating the mathematical text from one language to another, analyzing the perceptive implications that emerge (Piccinini et al., 2015).

Following Sergeyev (Sergeyev, 2015), it has been noted that there are problems analogous to translation when changing numerical systems from weaker instruments to more powerful one. The conceptual references can be found for a general theory of semiology in (Eco, 1975) and for the detail of translation problems in (Eco, 2002).
On the mathematical side we recall that the elementary models of Gordon (Gordon, 2004) and Pica et al. (Pica et al., 2004) are weak, while Sergeyev's expanded system is powerful (Sergeyev, 2008), so that many problems of translation may arise passing to and from ordinary arithmetic.

The note aims to capture apparently non-mathematical aspects of everyday language that are placed between these extremes and the true or presumed passage that can be accomplished passing from the real but informal world to the correct but abstract world of mathematics. Section 1 discusses the rules and the risks of this translation. Those who know the numbers associate them with the empirical quantities, to obtain a well ordered system and to compare different subjects among them. The clustering, both explicit and hidden, is a fundamental instrument for this purpose. The possibility of using different criteria is typical of classification in various sport competitions. Mathematics supplies consistent solutions to the problem, but cannot overcome its fuzziness. Some critical examples will be shown where even the use of the same clustering rules lead to fuzziness.

The second part of the note discusses the problem of clustering and its empirical consequences. The third part associates the problem of granularity with some errors of intuition that are committed in the calculus of probabilities and in its application to the elementary theory of games.

**Poor and rich systems of numbers**

The two extremes are the primitive and childish language of the three “numbers” <one, two, many> (Gordon, 2004) or of the richer system <one, two, three, four, five, many, really many> (Pica et al., 2004) and the extended system in which there is a number, the “grossone” that formalises the infinity (Margenstern, 2011). The idea is the same as the elementary system, that is, subsequent some (or many) numbers there is a last number bigger than all the other (compare Boyer, 1968). In ancient languages, such as Greek and Chinese, this number has been fixed in 10000 (myrioi = one myriad of Greeks, wan4 of Chinese), perhaps because it corresponds to the 100 x 100 square. It has always retained the double meaning of 10000 and infinity. The fundamental difference is that the “grossone” cannot be reached by a finite sum of addends, while the myriad can. As long as the calculations remain under this roof (and this often happens in practice) they coincide in the two systems.

In the original system of indoeuropean languages there were only three quantities associated with a name: singular, dual, plural. Afterwards the use of the dual has disappeared, except in Slovenisch and in some baltic languages. Let us remark that the opposition of singular and plural is not compulsory, for example, in Chinese. On the contrary the use of pseudo-numeral adjectives is widely
common in all languages. What is the meaning of the words <little>, <much>, <big>, <small>? Is there an ordinal scale of quantities? The classic case is the answer to the question “How many trees are there in this forest?”, where <many> is a correct answer, but does not allow comparison with another forest.

Postulate 2 of Sergeyev (Sergeyev, 2015) states that:

*Following the naturalistic approach of physicists, we will not venture into saying what mathematical objects are but we will build tools (such as natural numbers or other more or less powerful systems) that will allow us to improve our ability to observe and describe mathematical objects.*

Therefore the answer “many” of the previous example was not wrong, only the system used was poorly selective.

Combinations of generic terms quickly lead to fuzzy situations due to different clustering and different granularities. Furthermore, when arithmetic operations are attempted, the most frequent perceptive error is to believe that <little> + <little> = <little>. The absurd is therefore that <many> * <little> = <little>. The contrary absurd is that <many> - <little> - … - <little> = <many>, regardless the number of times <little> is subtracted. The problem of course does not exist for the “grossone”, because it is greater than any number, hence for example it makes sense to speak of 1 “grossone” minus 6 elements, that still is greater than any number (compare also Sergeyev, 2008). On the contrary in the children's speech we find the tendency to binary opposition between the two terms <all> and <none>, as exemplified in Piccinini-Indelli 1980-81 at middle school level; however this orientation is also found among adults in emphatic discourses.

It is usual to translate from the qualitative scale to some form of numerical scale. For example, he scholastic system in many countries translates the judgments linguistically expressed into a sequence of numbers, but it often happens instead that it is precisely the number that is taken as the label of a judgment. This does not alter the fact that there are strong oscillations between a linguistic-cultural area and another, both in space and in time.

Although originally the scale is qualitative, in all the rankings both socio-economic and sporting there is the immediate numerical transposition, through systems of indicators on which it is possible to perform operations of comparison, sum, and mean. Clearly, it is arbitrary how to eliminate variables considered not meaningful and, on the other hand, how to assign weights to significant ones, as well as the choice of aggregation and granularity modalities.

The extreme (lowest) case of granularity is the binary one of direct elimination, where the only certainty (at least in theory) is the one that wins the strongest, even if the ranking for subsequent positions is fuzzy.

For this reason there is the technique of creating the *seeded players*, that is to say to select the best eight (according to some previous category) and to avoid
their meeting in the first rounds. The eight players can be assigned by lot, or, if there is a ranking among them, the correct technique is to follow this order:
- Quarterfinals 1-8, 5-4; 3-6, 7-2
- Expected semi-finals 1-4, 3-2
- Expected final 1-2.
In direct elimination the information that is derived is the following:
- The losers of the semifinals are at the same level third/fourth;
- The losers of the quarterfinals are at the same level from fifth to eighth;
- No further information can be derived.
In the case of a random match, however, there is no certainty that the finalist who loses is the second in the ranking. In fact, in the case of the output 1-2, 3-4; 8-7, 6-5, the second would already be eliminated in the quarterfinals.
The problem of creating a ranking has no obvious solution, and it is a NP-hard problem, and must be done on the basis of the results, even quantitative, of direct comparison. The mechanism is as follows: a square matrix $M$ is created, where all the results translated into a number are recorded. After that, a new order is assigned, equal for the rows and columns, constructed so that the superdiagonal of the matrix assumes the maximum possible value and the subdiagonal the minimum one. In general it will not happen that the subdiagonal is null, as this only occurs when there is at least one complete ordering (and in this case the problem is reduced to a polynomial one). A powerful solution method is described in the work of Piccinini & Chang 2007, while its meaning in the macroeconomic models is analyzed by Chang, Piccinini, Iseppi in (Chang et al., 2013). Usually, for practical reasons, simpler proxies are used, as it happens in sport classification.

### Clustering and Bak-Sneppen models

Piccinini, Lepellere, Chang and Iseppi in (Piccinini, 2015) discussed some applications of Bak-Sneppen model, which describes the essential concepts of the progress of biological components in an interactive framework (Bak & Sneppen, 1993). The basic idea recalls clustering: every element of the system has a score (called fitness, to remind biological original aim of the model). At each period the worst element is forced to change, obtaining a new score randomly. Proceeding in this way, in the long run a suboptimal distribution would be achieved but not very realistic. Another possibility is to consider a natural random decay of all the elements, but the model provides a more interesting choice: when the worst element changes, some of its neighbours are induced or forced to change at random, even if their score was already high. The overall system gets an improvement but complete optimality can never be achieved. The authors in (Piccinini et al., 2013) and in (Piccinini et al., 2014) have shown that the construction of cluster boundaries, though perhaps non democratic, can improve
the average levels and allows from time to time sudden phenomena of overtaking. The basic Bak-Sneppen model poses a dramatic question: why who are involved in the process of improving weaker species can get worse? A first answer is obvious, because someone has to bear the costs, but also the change of an element may induce damage on the neighbours or stimulate them to change.

Good examples are found in linguistics and in learning foreign languages. Words can be grouped into clusters according to areas of meaning. This takes place in the construction of the linguistic thesaurus, often used in the systematic teaching of a language or in the dialogues of specific situation (for example “Greetings”, “Journey by plane”, “Customs”, etc.) The thesaurus system and the pertinences it creates are widely used in machine translation programs to settle cases of polysemy and homophony. The Chinese uses this system even in the writing of many ideograms, where often the semantic area is described by an ideogram called radical, which is glued with another ideogram that instead has a similar pronunciation and therefore their pictorial structure does not allow to derive pronounce. When the semantic cluster is progressively learned each new acquisition perturbs the meaning to be attributed to the next terms. This sometimes leads to neglecting its neighbors and supporting the new word in place of those already known. In the creation of languages, the passage from Latin to Neo-Latin languages is full of examples. We will remember pullus (young animal, bud, scion) that passing (already in Plautus) through pullus gallinaceus (son of hen) becomes chicken (Pollo, Italian and Spanish; Poulet, French) with this exclusive final meaning.

Another interesting problem is Lloyd’s problem. Its simplest version is this: in a territory a certain number of equivalent service centers must be installed. Each user will use the nearest service center; the centers are not in competition with each other, unlike the Hotelling model of the two ice-cream makers (Hotelling, 1929). The goal of the project is to minimize the burden of the route for all users. A good measure of this burden is the square of the distance, even if other functions can be used to evaluate it. This problem is called Lloyd's problem.

To solve it (in an iterative way) the Voronoi diagrams are used (Voronoi, 1908): given N nodes, we must divide the domain into N parts, that represent the attraction basin of each node. The division must be such that for any point of the attraction basin its generating node is the next of all. If we use the normal Euclidean distance, the result is a subdivision of the domain into N convex polygons, each containing a node. If the domain is not limited, one or more of the polygons will be unrestricted. The methods to perform the analysis are numerous; refer, for example, to (Aurenhammer, 1991).

The problem is linked to the Bak-Sneppen model when a new node is inserted in a cell. In fact, it is involved not only the polygon in which the new node falls, but also some of the adjacent ones. Only if the boundary consolidation
is carried out, the innovative action concerns only the polygon in which insertion takes place.

Lloyd’s problem requires that the centroids of the Voronoi diagram be chosen so as to minimize the sum of the overall distances. In case of a quadratic Euclidean distance and a uniform territory, the optimal point is the center of gravity of the polygon. A solution can then be found by the following iteration:

a) Given arbitrarily N nodes;
b) Evaluate the N Voronoi polygons;
c) Calculate the center of gravity of these polygons;
d) Use them as a new node assignment;
e) If they coincide (almost) with the previous set of nodes, stop. Otherwise restart from b).

In the paper by Du et al. (Du et al., 2006) it can be found a proof of the convergence of the above described iteration.

But does this solve the problem? The algorithm provides a local minimum, but not all the possible minima. This can be checked dividing a square into two domains. The minima may be achieved dividing horizontally or vertically according to the starting position of the first nodes, but the minimum is the same. Suppose now that you construct a rectangle, perturbing the original square by a small amount. There are still horizontal and vertical solutions, but only one provides the actual minimum! This shows that unfortunately Lloyd’s problem is not convex and may have a plurality of local minima. Usually they are not very different, so they are reasonable suboptimal solutions, but the problem remains fuzzy.

Let consider now a square to be divided into four Lloyd’s regions. A solution is dividing it along horizontal strips, another along vertical strips, another into four isosceles rectangular triangles. These are unstable solutions. A good stable solution is to divide it in four coordinate squares of side one half the original. But the best solution is achieved taking a triangle in the interior, and three confining particular irregular figures along the boundary. This plenty of solutions can be compared with economic expansion of activities as shown in (Chang & Iseppi, 2011; Chang et al., 2014) for the optimal central clustering, while in (Chang & Iseppi, 2012) and (Chang et al., 2015) the linear solution is chosen, even if it seems to be more unstable due to the limited resilience.

**Fine tuning in Probability**

When using probability, intuition is very dangerous. In particular results are very sensitive to the granularity of the framework. Experience could lead to lifelong memory of correct solutions as Freudenthal reminds in (Freudenthal, 1991). This part of the note associates the problem of granularity with the intuitive errors
that are made in the calculation of probabilities. We do not speak here of the critical problems of conditional probabilities and independent events, which require a more refined logical reflection, but we limit ourselves to some classic examples taken from game theory and to some evolution problems. A trap lies in hidden conditional probabilities, and it is shown in the last example.

A first case is to ask to write a random 10-digit sequence. The two samples were constituted in one case by adults of a good professional level and in the other case by engineering students. It was found that the percentage of choice of equal consecutive digits was zero for adults and significantly less than 1/10 for students. Obviously the probability that the next digit is equal to the previous one in an unbiased process is equal to the reciprocal of the number of digits. So, if we use decimal notation, the probability is 1/10 for each pair of consecutive digits.

A plausible interpretation is that the probability of equality is perceived as too low to be significant, which is true if the subject uses a scale of the type <impossible> <improbable> <possible> <probable>. Anyhow the most likely interpretation is that the random sequence is “felt” as a sequence of draws without replacement, as happens in the popular game of bingo. In the case of only ten digits and a long sequence, the possibility of avoiding any repetition is very low. As a result, usually after a few different items an old one is drawn again, but not near another equal item.

The intuitive scale becomes sufficient in the binary case, since the alternate sequence 101010…clearly cannot be considered random, so it is necessary to have some repetitions. How long should they be, in the average? The problem is solved by a Porfirio tree which each branch there is the possibility of change the sequence and start a new one. So the probability of changing after just one shot is ½; this means that the probability of moving forward is ½. At this point there is the probability ½ of ½ of changing, leaving only a possibility of ¼ to move on. So after 2 draws the probability of stopping is ¼, after 3 draws 1/8, in general after n draws is 1/2^n. The average length is therefore given by the sum of the sequence n/2^n, that is 2, as calculus teaches.

Already in the case of three digits most people seem to like the construction of preferential rather than equiprobable sequences. This is proven experimentally by the impossibility of winning in the game of Chinese morra with an expert system that analyzes the behavior of the human adversary.

Let us recall that in this game each player has three moves, which are also physically represented: the card = extended hand, the stone = closed fist, the scissors = two extended fingers and the other three closed. Players show their move at the same time. The card wins on the stone, the stone wins on the scissors, the scissors win on the card. In the case of an equal move, neither wins.

The winning strategy among skillful players is to make every move at random with the same probability. However, if a player changes the next move
after each move, as soon as the opponent realizes this mistake, he has a winning strategy. In fact, after a stone the first player will choose between paper and scissors, so the opponent who plays scissors is sure not to lose and has a ½ chance to win.

The answer rule is therefore:

<table>
<thead>
<tr>
<th>Stone</th>
<th>Scissors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper</td>
<td>Stone</td>
</tr>
<tr>
<td>Scissors</td>
<td>Paper</td>
</tr>
</tbody>
</table>

The expert system is needed to analyze the behavior of the first player and understand what are his privileged sequences. In correspondence with these he builds the winning strategy. However, it may be noted that the gesture of the scissors requires a more elaborate movement and therefore in a quick game it may appear less often, so playing paper with a probability greater than 1/3 can be a good strategy.

The game between two computers with the same pseudo-random number construction algorithm, on the other hand, ends in parity.

The opposite phenomenon occurs in the falsification of accounting books, in which the last significant amount is not randomly distributed, but are dominant 0, 5 and 9, the latter for commercial and psychological reasons.

The case in which the random number is required is that of certain games where one has to guess the sequence of 5 or 6 distinct numbers chosen among the possibilities of the bingo game. In theory, the events are equally probable, so the obvious sequence of type 1, 2, 3, 4, 5 has the same probability of 3, 89, 32, 17, 2. However, if the game requires that the bet is divided between all those who guess the winning sequence extracted on the bench, it is in the interest of the player to bet on a sequence that is probably not chosen by any other player, a bit like it happens with the numeric keys of the safes.

Let us now come to a very subtle example taken from the book of Métivier (1968) in which the case of the two brothers is discussed. It is given that the probability of a male or a female among the juvenile population is approximately equal. Then we consider the males present in the universe and identify those that belong to pairs of two brothers. At the end it is asked if the other member of the couple is male or female. Most observers expect MM responses and MF responses to be about the same. Instead it is not so because the entire universe, writing first the older brother, consists of pairs MM, MF, FM, FF, averagely equiprobable. In our sample the FF pairs have not been consulted, so the MF answer, which has been considered equiprobable to the FM, appears with double frequency of the MM. The experimental verification made among the students of electronic engineering nevertheless gave the surprising result that MF and MM had approximately the same frequency. This is probably due to a considerable distortion of the sample, since evidently the presence of a sister diversified the
cultural landscape of the male, reducing its propensity to electronic engineering, while the presence of another male increased it.

Conclusions

The advice that should be given to those who read this note is to keep in mind the saying “The best is the enemy of good”. The Bak-Sneppen model presented with some of its applications in the second section shows which price can be paid for an improvement even by those who already adequately comply with current regulations.

Operational research and business logistics offer numerous cases in which there is a global improvement, but the price to pay can be high, both in terms of training time and in terms of obsolescence of existing equipment and skills acquired. On the other hand, even young people who enter directly into the new system can not easily use the skills of the olderly and will be forced to reflect on the knowledge that is imparted by the new system.

The first section, on the other hand, shows how the mathematical models can be altered upwards or downwards, and how these variations conform to the way of thinking and the language generally used. The affirmation of science that the more the tools are perfected the more knowledge is possible, it stops at the practical level with the finite of our ability to work mathematical and conceptual (first axiom of Sergeyev (2015). Information science teaches us that it is essential to adapt the tools to the proposed objectives (and also to the times when problems must find an appropriate solution).

The third section, instead, opposes Sergeyev's third axiom to the first, which re-establishes that even in a complex system “the totality is greater than every single part”, which should be read in the other verse “every part is smaller than the total”. The last section recalls the errors made in the use of probability calculus and in game theory when parts of the universe in which we move are omitted.

Summary

In everyday life we use a small amount of numbers. After all, even if we used 100 numbers a day in our entire life we would remain at a few millions, but most likely the numbers actually used will be a few hundred, some of which are repeated very often. There are more or less precise quasi-numerical adjectives, such as <little>, <very>, <still> ... and in general a “final number” that is often called in the modern languages Billion, but in Greek there is the Myriad (10000). The final number is “fuzzy” because it is often replaced by the expression <full>. It is interesting, because if it is diminished by one unit it remains unchanged: if we say that the theater was full of people, after our neighbor came out, we continue to say that it is full. This from the mathematical point of view leads to contradictions, and they are overcome in the Sergeyev system, object of the first paragraph. In this system it is allowed to specify that the first moment it was actually <full> while in the second moment it was <full> minus 1.
In the second paragraph the Bak-Sneppen evolution model is recalled and applied to the clustering problem, which in turn is closely related to the creation of reduced quasi-numerical systems. A common application is the teaching of foreign languages, but more striking effects can be seen in models of urban planning such as the case of the Lloyd process. It shows in particular the non-uniqueness of clustering solutions and their substantial arbitrariness.

The third paragraph concludes with examples of intuitive, but distorted, use of probability. They are explained by the insufficient precision of the description system that is used. It may become ambiguous when it is translated into a mathematical model that happens to be too poor or too fuzzy for giving a correct representation.

References


