A micoworld to implant a germ of probability

Michele Cerulli, Augusto Chioccariello, Enrica Lemut

A MICROWORLD TO IMPLANT A GERM OF PROBABILITY

Cerulli Michele, Chioccariello Augusto, Lemut Enrica

Istituto Tecnologie Didattiche - C.N.R. di Genova

This paper reports on the third part of a long term experiment concerning the introduction of 12-13 years old pupils to the concept of randomness and probability. In the first parts, the pupils analysed the concept of randomness, in daily life and with LEGO-RCX robots. The third part is based on the Random Garden Game (within the programming environment ToonTalk). The pupils are guided towards the concepts of equivalence of sample spaces and of classical probability, through temporary and personal meanings, corresponding to an evolution by means of class discussions.

INTRODUCTION

This paper accounts of some findings of a long term experiment [1] based on technological tools to introduce pupils to the concepts of randomness and probability. The experiment is composed of three phases: the first two concern pupil's introduction to the concept of randomness (Cerulli et al., 2006). The third phase focused on the concept of probability, with particular attention to: **a)** Pupils' development of the concept of sample space and of "equivalence" with respect to the events involved; **b)** Pupils' construction of theories for comparing sample spaces; **c)** Pupils' construction of the concepts of frequency, relative frequency and probability.

The focus of this paper is mainly on how we used the computer programming environment ToonTalk (Kahn, 2004) to approach issue **a)**.

THEORETICAL ASSUMPTIONS

Research literature shows difficulties related to pupils' introduction to probability (Fischbein, 1975; Pratt, 1998; Wilensky, 1993; and Truran, 2001), witnessing the failures of standard approaches. One of these difficulties concerns pupils' capability of interpreting different random phenomena according to a unifying perspective (Pratt et al., 2002; Nisbett, 1983 via Pratt). In order to avoid this difficulty, our experiment is based on the idea of presenting pupils with different random phenomena that teachers and pupils can recognize and consider as belonging to a sector, unitary and homogeneous, of the human culture, that we can identify as *experience field of aleatory phenomena* (Boero et. Al., 1995). A unifying perspective is partly achieved in the first phases of the experiment, where phenomena such as the tossing of a coin, and one-dimensional random walk, were represented by means of unique robot, the Drunk Bot (Cerulli et al., 2005). However, according to Pratt (1998) the construction of meaning within this subject lies in connecting its formal and informal views. A formal view according to Noss and Hoyles (1995) can be achieved thanks to the introduction of suitable computer microworlds in the school practice. The activity we are presenting is based on a specific microworld, the Random Garden that we built ad hoc in ToonTalk. The key idea is that the microworld can be used as a unifying model for representing and manipulating random phenomena. This model is unifying in the sense that each random phenomenon is represented as a sample space with a random extraction process, and different phenomena's representations are structurally the same. Moreover, the analysis and manipulations of the phenomena, within the microworld, are also qualitatively identical.

According to socio-constructivism, our experiment assumes that learning can be the result of active participation in both practical and social activities. However, such kinds of activities do not guarantee that the meanings constructed by the pupils are coherent with mathematics or with the teacher's educational goals. Such coherence can be achieved by means of mathematical class discussions orchestrated by the teacher (Mariotti, 2002; Bartolini Bussi, 1996).

THE RANDOM GARDEN TOOLS

The Random Garden is a microworld, for representing random extraction processes. The tool consists of a sample space (the *Garden*) a *Bird* and a *Nest*. [2] When the user gives a number to the *Bird*, a corresponding number of objects is extracted (with repetitions) from the *Garden* and deposited in the *Nest* (Figure 1).

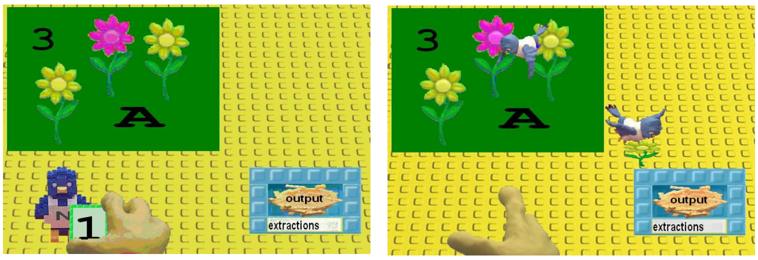

Figure 1. A number is given to the N-bird to request a random extraction of *N* objects. A new bird comes out and drops extracted objects in the output nest.

Figure 2. Eight extractions are collected in a box containing a nest (left); only the first element is clearly visible. The nest can be converted into a box with eight holes showing the ordered sequence of extracted objects (right).

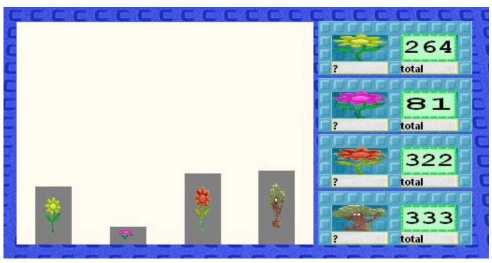

Figure 3. The *Bar Graph* (left) and the *Counters* (right), show respectively the proportions and the exact numbers of elements extracted for each kind of object.

The user can modify the garden by adding or removing objects which can be numbers, text, or images of any kind. It implies that this simple device can be used as a means for representing any kind of random phenomena. The elements extracted from the Random Garden are collected in a box containing a nest (Figures 1 and 2).

In order to visualize the whole sequence of extractions, it is possible to convert the nest into a new box with as many holes as the number of extracted objects (Figure 2). Hence, a rough qualitative view of the sequence and of its properties can be seen at a glance. If the number of extractions is large, and/or if one needs a more detailed qualitative/quantitative analysis of the data, other tools are required and provided: *Bar Graph* and *Counters* (Figure 3). These are dynamic tools, in fact numbers and bars change while the extraction is in progress. In particular, the bars oscillate at the beginning of the extraction process, and stabilize after a large number of extractions.

THE GUESS MY GARDEN GAME

If one is given a nest, or a box of extractions (see Figure 2), it is possible to address the question of what a possible composition of the Random Garden that generated the given sequence is. This key question is at the core of the Guess my Garden game [3] which is conducted as follows. One team of pupils, namely a small group, creates a *Random Garden* (using at most 12 objects), thus defining a sample space; then produces a set of boxes containing increasing numbers of extractions, for example, 2 boxes with 100 extractions and 2 boxes with 1000 extractions, and so on.

All the boxes, containing the data generated by the pupils' team, are included in a ToonTalk *notebook* named like team, and the *notebook* is published on the web, as a challenge for other players. Another team can then download the notebook and analyse the data it contains in order to try to guess the makeup of the Random Garden produced by the challenging team. The team can either simply observe the sequences of extractions, or study them using the *Bar Graph* and *Counters* tools. Once they make a conjecture concerning the garden to be guessed, they can produce a new corresponding garden and use it to produce a number of extractions that may be compared to those provided by the challenging team. Once the team is satisfied with the conjectured garden, they can publish it on the web and wait for their counterparts to validate or invalidate their answer. Finally, the challenging team checks the published answer and posts a comment to inform the other team whether they have guessed their garden correctly or not. If the garden has not been guessed, then the exchange between the pupils can continue until an agreement is reached.

THE EPISODE

The first two phases of experiment involved reflective and practical activities aiming at exploration and consolidation of some key issues related to randomness (eg. *unpredictability, fairness, indeterminism, random walks, etc.*). In the first phase pupils collected, proposed, and analysed sentences, talks, and episodes related to randomness. Pupils had to write individual reports and to discuss some of the emerging items with the rest of the class. Each considered item was discussed in terms of key questions such as "is it random or not?", or "is it predictable or not". The results of the discussions were reported in a shared class document, called *Encicloaedia of Randomness*. The second phase was based on the interaction with some LEGO robots that incorporated different aspects of the concept of randomness (see Cerulli & al. 2006, for more details). Each robot was discussed and classified as random or not random: the same questions used in phase one functioned as pivots for the teacher's orchestration of the discussions. After these phases, the pupils were very familiar with the concept of randomness and were about to be introduced to probability starting from the Guess my Garden game.

A class of 21 pupils from Milan starts the game by making public a set of challenges, and receiving answers from Swedish and Portuguese opponents. The first protagonists of the episode are the members of Jeka's team: Jeka (Jk), Jè (Je) and Rossana (R). They have to build a *garden* to be used to publish a difficult challenge as required by the task (Italian dialogues translated by the authors):

Jk: we could do…the same number…of flowers and trees

Jk and Rossana: three times this one, three times this one, three times this one and three times this one (*pointing to the objects in the random garden*)

The girls build their garden and ask the software to produce 100 extractions from the garden, and then comment on the results:

Jk: yellow flower 25 extracted times…

R: …but they are all the same?! (*looks surprised*)…more or less…25, 24, 25 and 26 …ah, yes, of course, we put (*in the garden*) all the same numbers (*of flowers and tree*)…(*she looks around to stress that she is stating something obvious and her pals nod*).

The girls take note of the obtained result and a researcher (M) intervenes to investigate what strategy the pupils are using to build a difficult challenge:

R: we multiplied each object of the garden by 3 (*pointing to the monitor*)…we tripled

M: Why do you think this is difficult to be guessed? (*reads one of the written questions pupils are supposed to answer in order to accomplish their task*)

R: no, it is not difficult, we just tried…

M: but it is not easy to guess this …*(he is promptly interrupted by Jeka)*

Jk: exactly! Because…one may think of two (*objects*) maybe…

R: …yes…. (*thoughtful*)

Jk: I would think of two (objects of each kind in the garden)

The original idea (to triplicate the numbers of flowers) of Jeka begins to be clearer and becomes more explicit when the teacher (A) asks them for an explanation:

A: why do you think this garden is difficult…?

Jk: maybe because with the resulting numbers *(after the extraction)* one may…one may get confused

A: why? What answer could you get?

Jk: maybe two

M: you mean two…

Jk: I mean, if one sees 20

A: 2, 2, 2, 2? *(meaning a garden with 2 objects of each kind; Jk, Jè and R nod)*

After some reflections and discussions, the girls decide to publish the garden they had produced, made of 3 objects of each kind (Figure 4). They believe that their opponents may think that the garden is made of 1 object (or 2 objects) of each kind.

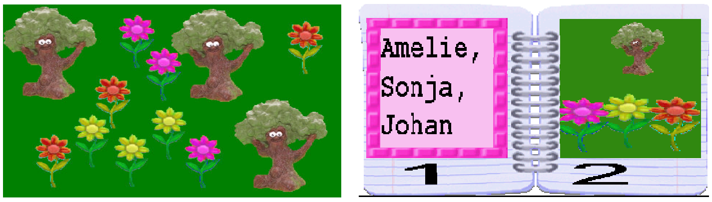

Figure 4. On the left it is possible to see the garden defined by Jeka's team, while on the right it is possible to see the attempt of guess sent by a Swedish team.

The following week the class goes back in the computer laboratory and each team finds an answer from a Swedish team. Jeka's team finds the answer given by Amelie's team (Figure 4), who conjectured that the *garden* contained 1 object for each kind, instead of 3, as expected by Jeka's team. The answer is considered wrong by Jeka's team, as they write in the message they send to the Swedish opponents.

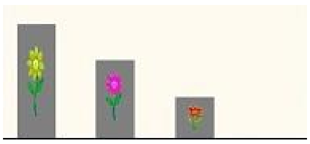

Figure 5. The results of 1200 extractions from M's challenge.

After responding to their Swedish pals, the Italian pupils are required to discuss how to respond to M's challenge. Such challenge is analysed by two teams who obtain the graph of Figure 5 using the *Bar Graph* tool.

Figure 6. On the right: Lollo's team's garden; on the left Jeka's team's garden.

The two teams that answer to M's challenge give different guesses: a) 2 red flowers, 4 pink flowers and 6 yellow flowers (Jeka's team); b) 1 red flower, 2 pink flowers and 3 yellow flowers (Lollo's team).

Which of these two answers is correct? The teacher poses this question a couple of weeks later within a class discussion where the pupils couldn't access the computers and the Random Garden tools. In order to support the class discussion, the teacher provides a set of cards representing the gardens proposed by the two teams, as shown in Figure 6.

The teacher recalls the two teams' different answers and asks the class which of them is right or wrong, or if they are both wrong or right. A brief discussion follows where pupils agree that the chances to get a red flower from garden A or from garden B are the same, as expressed by C1:

C1:	they are the same; in fact they are all doubled (*meaning that the number of flowers in garden B is twice the number of flowers in garden A*)

The gardens are now regarded as objects and their constitutive properties are compared, and garden B is now regarded as a sort of "double" of garden A. This idea reminds Jeka of the challenge they proposed to their Swedish pals.

Now the terrain seems to be ready for planting the seed of equivalence and the teacher takes this occasion to introduce explicitly the word "equivalent":

A:	[…] thus these two gardens, in theory, are equivalent?
C1:	yes, they are equiv…
A:	[or] Are they equal?
C2:	they are equal
A:	[or] Are they identical?
C1:	they are equal
Jk:	they are equivalent
C1:	they are equiv…

A noise of chat among pupils follows which ends with C1 stating:

C1:	equivalent!
Jk:	they are equivalent! (*Someone in the background says "equal"*)
C1:	equiva…equivalent…*(seems to be doubting)*
A:	will you explain me what you mean by equivalent?
Jè:	not equal because there are not the same elements in the two gardens
A:	Thus M surely had one or the other *(meaning that if M had one of the proposed gardens, he couldn't have both of them but only one)*
Jè:	equivalent because they have the same values….in practice…we can say so!

After a while, the position expressed by Jè (when she states that the two gardens are not equal because they have different elements) opens the space for discussing new criteria (different from pure "equality") for comparing gardens. Jè introduces her criterion of equivalence, based on the "values" [4] of the gardens. At this point it is not clear what is meant by the word "values", but from the context we deduce that Jè refers to some kind of result produced by the gardens that it could be the graphs, the numbers in the *Counters*, or the sets of extractions. In the meantime, other pupils, like Bo, propose their interpretation or alternative "definition" of equivalence:

Bo:	the percentages are the same…I believe it is because the percentage is more or less the same.
A:	the same of what?
Bo:	of…of twelve…of all the flowers…for instance
A:	give me an example
Bo:	in the first garden, the garden A, the percentage is 6…in the first garden the percentage is 6, six is …it is …
A:	the percentage? The total?
Bo:	yes, the total, the 6 is like the 100, and the red is 1, thus it is ….the red is 1 over 6…oh god!
A:	1 over 6
Bo:	1 over 6 and the second (garden) is 2 over 12 which is like 1 over 6
A:	uhm, thus you say "they are equivalent as the two fractions" that you said?
Bo:	yes
A:	right?
A:	but they are not the same, so we can say that these two answers are equivalent […] but without knowing Michele's garden we cannot give a definite verdict, ok?
Bo:	because even if we could make the extraction….surely the numbers would not be the same, they would be almost the same!
A:	almost the same, and the colums [of the graphs]?
Bo:	equ….with the same heights I think
A:	in the two gardens?
Bo:	in the two gardens the heights of the columns would be the same but the values different

Bo's idea of equivalence, different from Jè's, is based on an analysis of the constitutive elements of the garden, and he tries to formalize it by associating fractions to the garden. The fractions are only by chance coherent with the classical definition of probability which at this stage is not known to these pupils [5]. However, Bo also agrees that the two gardens would produce the same graph.

The class needs to establish criteria to validate responses to challenges. This leads them to introducing some idea of equivalence. Such idea is still fuzzy, but nevertheless it turns out to be useful in the following excerpt where the class discusses the case of Jeka's team and their Swedish opponents. According to Jeka, the strategy used by M is the same as the one used by her team. They both produced a garden whose results were "the same" as the results of other gardens, so that their opponents' chances to guess are low. This issue relates to the idea of equivalence of gardens, so the teacher approaches it just after the discussion of M's challenge, and puts it in terms of validation of the Swedish answer:

A: they [Jeka's team] answered to the Swedish pupils "you did not guess", are we leaving this answer or are we going to write to them again?
C: I think it is better to re-write it and state that they did well but the values…
Jk: they did wrong
C: they did wrong but with respect to the values…they did wrong but they were right because the original value was 3, 3, 3 and 3, the one they had to guess, and they put 1, 1, 1, 1 but it is not their fault because…
Jk: they did wrong
C: they did right, it is only that the bars were the same, so they could put any number
Jk: I think we should re-write it stating that they did wrong, but that the two gardens are equivalent like the other one (referring to M's challenge)

According to C, the Swedish team's answer should be considered correct because their garden produces the same graph as Jeka's team garden. This idea is re-formulated by Jeka in terms of equivalence between the two gardens, thus the word "equivalent" (ita. "equivalente") begins to be used by the pupils as a tool for validating responses to the challenges. However, the meaning associated to the word "equivalent" needs some more clarification:

A: what is the phrase you would write to explain them…?
C: I would write, you did wrong…should I say also the solution?
A: suppose that what you say is going to be sent to the Swedish pupils
C: you did…you did wrong…you didn't guess our garden but you found another one that has the same value of the one we did, apart from…

The words "values" (ita.: "valori") and "equivalent" appears as strictly tied, and their relationship is definitely cleared by C in the following excerpt:

A: what do you mean by "value"?
C: I mean that the graphs of the extractions are equal to the graphs of our garden
A: perfect
C: only, our garden had different quantities of objects
A: so, if they give the same extractions they are equivalent
C: there are 3 equivalent gardens…there are 2 equivalent gardens…there is our garden, 3 red flowers, 3 yellows, 3 pinks and 3 trees, your garden, one for each object, and a third one containing to elements for each object .

Finally, the meanings of the words "values" and "equivalent" are cleared and the class agrees on the following criterion for validating answers to responses: the answer is correct if the garden proposed by the responder is equal to the original garden; the answer is almost correct if the garden of the responder is equivalent to the original garden, in the sense that they produce the same graph.

What we find interesting in this story is that the "germ" of the idea of equivalence appeared in the form of the strategy employed by Jeka's team for winning the game. It then reappears in the discussion of M's challenge which was designed on purpose to exploit the ambiguities related to the equivalence of gardens; in this case the idea of equivalence appears in the form of a criterion for deciding which of the two proposed answers is correct. Each of these steps corresponded to an evolution of pupils' idea of equivalence of gardens by means of reflections and class discussions that are clearly motivated and driven by the needs of the game. The two needs that drive such evolution are basically: the need for finding a principle to validate answers; the need to produce "difficult" challenges.

CONCLUSIONS

In this paper we show how the third phase of the long experiment contributed to pupils' development of an idea of equivalence of sample spaces, which we assume to be basic for developing a definition of probability. We observe also that in this kind of activity pupils implicitly cope with matters related to the law of large numbers.

The experiment continued by involving the pupils in tasks of comparisons of sample spaces, questioning on how easy (probable) it was to pick a given flower from a given Random Garden. Starting from these tasks, new class discussions were set up, which led pupils to define their own operative strategies for comparing sample spaces, in terms of choosing the sample space which would more likely generate a specific event. All of these strategies consisted in some kind of computation associating a "result" to each sample space: the comparison of such results would help choosing the "best" sample space. One of the strategies proposed by pupils was consistent with the classic definition of probability. This made room for introducing the classic definition as a means for studying and comparing random phenomena.

The strategies proposed by pupils are, as a matter of fact, not dependent on the context, in the sense that the flowers of the random garden can be substituted with other objects/symbols. Thus the Random Garden becomes a unifying model to represent random phenomena, as suggested by the final activities of the project, where pupils used the Random Garden tools to reproduce on the screen the behaviours of the LEGO robots they experienced in the first phases of the long experiment.

NOTES

1. We acknowledge the support European Union. Grant IST-2001-32200, for the project "WebLabs: new representational infrastructures for e-learning" (see http://www.weblabs.eu.com/).

2. See a tutorial: http://www.weblabs.org.uk/wlplone/Members/augusto/my_reports/Report.2004-06-21.4151

3. Rules and data: http://www.weblabs.org.uk/wlplone/Members/augusto/my_reports/Report.2005-01-04.5407

4. Consider that the Italian "valori" (used by Jeka and here translated with "values") can have several meanings, among which it can represent either the values of the parameters of the input of a process or the results of the same process.

5. We should mention that the class had previously discussed on what strategies can be employed to decide, given two gardens, which of them is more likely to extract a certain kind of flower. The strategies proposed by pupils include "measure" of "chances to extract" of various kinds among which we find the fractions employed in this case by Bo.

REFERENCES

Bartolini Bussi M. G. (1996). Mathematical Discussion and Perspective Drawing in Primary School. *Educational Studies in Mathematics*, 31 (1-2), 11-41.

Boero, P., Dapueto, C., Ferrari, P., Ferrero, E., Garuti, R., Lemut, E., Parenti, L. & Scali, E. (1995). Aspects of the Mathematics-Culture relationship in mathematics Teaching-Learning in compulsory school. *Proceedings of PME XIX*, Recife, Vol. 1, pp. 151-166.

Cerulli, M., Chioccariello, E. & Lemut, E. (2007) Randomness and LEGO robots. (In Bosh, M. (Eds), *Proceedings of CERME 4*. IQS Fundemi Business Institute, Sant Feliu de Guixols, Spain, ISBN: 84-611-3282-3)

Fischbein, E. (1975). *The Intuitive Sources of Probabilistic Thinking in Children*. (Reidel, Dordrecht)

Kahn, K. (2004). ToonTalk-Steps Towards Ideal Computer-Based Learning Environments. (In Tokoro, M. & Steels, L. (Eds.), *A Learning Zone of One's Own: Sharing Representations and Flow in Collaborative Learning Environments*. Ios Pr Inc.)

Mariotti, M. A. (2002). Influences of technologies advances in students' math learning. (In English, L. D. (Eds.), *Handbook of International Research in Mathematics Education* (chapter 29, pp. 757-786). Lawrence Erlbaum Associates publishers, Mahwah, New Jersey.)

Nisbett, R., Krantz, D., Jepson, C. & Kunda, Z. (1983). The Use of Statistical Heuristics in Everyday Inductive Reasoning. *Psychological Review*, 90(4), 339-363.

Noss, R. & Hoyles, C. (1996) *Windows on mathematical meanings learning cultures and computers*. (Kluwer Academic Publishers, Dordrecht/Boston/London.)

Pratt, D. (1998). The Construction of Meanings IN and FOR a Stochastic Domain of Abstraction. Dissertation University of London, Institute of Education.

Pratt, D. & Noss, R. (2002). The Microevolution of Mathematical Knowledge: The Case of Randomness. *The Journal of The Learning Sciences*, 11(4), 453-488.

Truran, J. M. (2001). The teaching and Learning of Probability, with Special Reference to South Australian Schools from 1959-1994. Dissertation, University of Adelaide.

Wilensky, U. (1993). Connected Mathematics-Building Concrete Relationships with Mathematical Knowledge. Dissertation, MIT.