Deterministic Model for objectifying the students' proving process in Discrete Mathematics

Mickaël Da Ronch[1]

[1]University of Teacher Education Valais, Saint-Maurice, Switzerland; Mickael.DaRonch@hepvs.ch

This paper presents a deterministic model for describing and objectifying the students' mathematical activity, targeting the field of discrete mathematics. This model allows to identify their proving processes in problem-solving, involving the manipulation of tangible objects, based on the audiovisual data collected. The model is described in three steps. The first one involves selecting observables relating to the actions and strategies linked to the proving process identified in the mathematical analysis that we wish to identify in the students' mathematical activity. The second one concerns the retranscription and encoding of their activity in relation to the choice of actions performed. Finally, the last one involves analysing the retranscriptions of their activity and comparing them with the strategies identified in the mathematical analysis, to find similarities in the proving processes used to solve the problem.

Keywords: Deterministic model, proving process, problem-solving, discrete mathematics education.

Background and objective of the research

The field of Discrete Mathematics is important for teaching and learning of the reasoning and the proving process. As Gravier and Ouvrier-Buffet (2022) point out, many international works on problem solving and proof use this field (e.g., Da Ronch, 2022; Grenier & Payan, 1999; Hanna & de Villiers, 2012; Hart & Sandefur, 2018; Stylianides, 2016). Discrete Mathematics encourage the students to develop an authentic mathematical activity (Lampert, 1990) similar to that mathematician, particularly in terms of the skills and know-how associated with this activity. Moreover, this field makes it easily to illustrate problems using tangible objects (see Grenier & Payan, 1999; Da Ronch, 2022). But how can we identify in students' mathematical activity and their proving processes used in problem solving when they manipulate these objects? One way of proceeding is to carry out an analysis of the traces using audiovisual data collected with the camera assistance. The audiovisual data allows to researchers to access more easily and more exhaustively to the traces of activities produced by students (Jewitt, 2012; Da Ronch, 2022). This also means that students don't have to select the traces they wish to leave behind a researcher or a teacher when it comes to submitting written work. Written traces lead to a loss of information and make the retranscription of the activity less objective. In a way, they also encourage interpretation biases in the analysis of these written traces by the researcher. The aim of this paper is therefore to propose a deterministic model for retranscribing the students' mathematical activity on discrete mathematics problems and to provide tools for analysing their proving processes. This model is a theoretical model for research in (discrete) mathematics education. It is therefore not intended to be put into practice by teachers, but rather to be used by researchers. On this occasion, we will first give some theoretical elements which will then enable us to describe the main steps in this model, particularly about the choice of observables, the retranscription method and the analysis of the latter. Several examples illustrating this model are given in Da Ronch's work (2022).

Theoretical background and research questions

Many models of problem-solving processes exist (Rott et al., 2021). The authors describe and compare existing models to characterise their potential and limitations for analysing students' problem-solving processes. Among other things, they point out that most of the models proposed are normative. They affirm that these phase models were not designed for the analysis of empirical data or to describe externally observed processes (e.g., videotaped problem-solving processes of students), which highlights the authors' need to build a descriptive model. The model developed by Rott et al. (2021) can be used to code episodes, i.e. to code "A period of time during which an individual or a problem solving group is engaged in one large task […] or a closely related body of tasks in the service of the same goal […]" (Schoenfeld, 1985, p. 292). These episodes are then coded according to a pre-defined typology (deductive approach), which is then enriched by data analysis (inductive approach). However, episodes are coded at a macroscopic level of student activity (e.g., "reading the problem; implementing a plan; verifying a solution", etc.) (Rott et al., 2021, p. 743). This model therefore does not allow us to retranscribe, at a microscopic level, the activity of the students in order to give a retranscription as close as possible to the reality observed. Moreover, this microscopic retranscription of the pupils' activity would make it possible to reduce the interpretation bias that can be induced by coding into episodes and coding the associated episode types.

If we want to identify students' mathematical activity in problem-solving, and their proving processes, one way of doing this is to identify a set of elementary actions considered relevant to the given problem. We are therefore at a microscopic level. Moreover, these actions are directed towards the environment of the situation, but do not all have the same function. The notion of environment named *milieu* in French didactic was developed by Brousseau (1997), Margolinas (1995) and Bloch (2006) in terms of structuring the environment by levels. Other researchers have proposed a characterisation of the environment in which students operate according to three spaces: micro-space, meso-space and macro-space (e.g., Berthelot & Salin, 1992). These three spaces simply allow to classify the environment according to the value of the variable concerning the size of the space. However, the characterisations given in the literature do not allow us to categorise the actions directed towards the environment of the situation in detail, especially as these actions do not all have the same functions (Da Ronch, 2022). It therefore seems more appropriate to categorise our environment according to its functionalities. This will enable us to propose a typology of actions according to the zone of interaction with the environment and their assigned function. For our research, the environment is therefore characterised in *functional zones*, while considering the social dimension of mathematical activity. The *Work Zone* ($\mathcal{W}_\mathcal{Z}$) is the main area of the environment where the subject act to solve the problem. Therefore, the actions in this zone are linked to actions about problem-solving. The *Tooling Zone* ($\mathcal{T}_\mathcal{Z}$) is the zone of the environment which provides to the subject(s), the objects they need to develop their actions in problem-solving. The *Information Zone* ($\mathcal{I}_\mathcal{Z}$) is the zone of the environment which communicates information about the problem: rules of the game, instructions, examples, *etc.* (Da Ronch, 2022; 2024).

In relation to the objective formulated in the first section of the paper and the characterisation given to our environment, several questions can be formulated, namely*: what choices of observables should we make? How can we retranscribe the students' mathematical activity? And how should it be* *analysed?* In the following section, we describe the steps in our model that will enable us to answer to these questions. These steps will make it possible to objectify students' mathematical activity, and particularly their proving processes based on audiovisual data collected.

Method for processing and analysing audiovisual data: a three-step model

In this section we present our model called **ORA** (**O** for observable, **R** for retranscription and **A** for analysis). The different steps of this theoretical model will be illustrated by an example based on a problem in discrete mathematics. This model has already been applied to other problems (e.g., Da Ronch, 2022) during research aimed at retranscribing and finely analysing the mathematical problem-solving activity of individuals, out of the classroom (in a context of popularisation of mathematics), with dozens of participants. In each case, coding was carried out by a single researcher based on the choice of observables made through mathematical analysis of the problem (deductive approach).

Choice of observables: typology, actions and strategies for identifying proving processes

Firstly, we have to identify elementary actions noted a_i. The mathematical analysis carried out on the problem allows us to retain only a subset $\mathbb{A}$ of this a_i-actions. Thus, these elementary actions are chosen for epistemological reasons in relation to the knowledge we have on the problem, or for didactic reasons. Furthermore, the subset $\mathbb{A}$ of these actions is finite since the corpus of audiovisual data to be processed and analysed is also finite. The characterisation of the environment by *functional zones* makes it possible to propose a typology of actions. We can therefore distinguish 4 types of specific action. The $\mathcal{W}$*-actions* which are actions that modify $\mathcal{W}_{\mathcal{Z}}$. *The* $\mathcal{T}$*-actions* that interact with $\mathcal{T}_{\mathcal{Z}}$. This may involve selecting or depositing objects in $\mathcal{T}_{\mathcal{Z}}$. *The* $\mathcal{I}$*-actions* that allow access to $\mathcal{I}_{\mathcal{Z}}$. Finally, the $\mathcal{SO}$*-actions* which are actions arising from social interactions involving exchanges between peers or to oneself. They can be actions in oral language or actions that involve gestures (Da Ronch, 2022; 2024).

To identify traces of students' mathematical activity, we need to identify actions based on the knowledge or skills we have about solving of the problem. One way of doing this is to objectify their solving strategies, which require the mobilisation of knowledge, skills or know-how, by means of ordered actions carried out on objects in order to solve a given problem **P**. These strategies will be defined in *extension* or in *comprehension* (Da Ronch, 2022). When we have access to all the strategies and can describe them as a sequence of a_i-actions, on instantiated o_j-objects, we will define them in extension. A *strategy* can therefore be defined as a *word* $\boldsymbol{\omega}$ which takes the form of a sequence $a_i o_j$. On the other hand, when it is not possible to describe all the strategies because there are too many for instance, we will describe these strategies in comprehension, *i.e.,* using a set of actions on generic objects. The example below illustrates a sequence of ordered actions (additions) performed on objects (dominoes) to prove the non-existence of tiling in this configuration (Figure 1). The proof of non-existence consists of reasoning by necessary conditions until an absurdity is reached (two non-adjacent squares that cannot be tiled by a domino). The case study provides a complete proof of the non-existence of tiling.

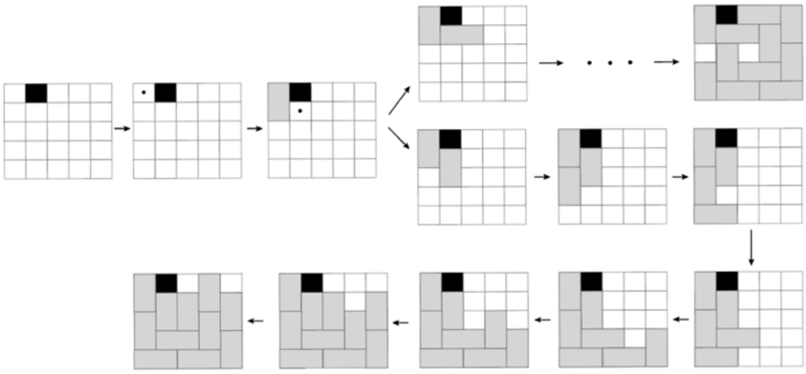

Figure 1: Strategy for proving the non-existence of this tiling

Retranscription of the experience: encoding of elementary actions and retranscription method

To retranscribe the actions that describe the strategies used by the students during the proving processes and the actions linked to the formulation of knowledge, we need to give an encoding. To make our retranscription intelligible, we bijectively encode the actions a_i of $\mathbb{A}$ in an intelligible symbolic format taken from a finite set $\mathbb{L}$ of symbols. These symbols are built as close as possible to the observed reality.

Table 1: Examples of encoding

Type of actions	a_i-action of $\mathbb{A}$	Associated symbol of $\mathbb{L}$
$\mathcal{W}$-action	To add	$+$
$\mathcal{W}$-action	To remove	$-$
$\mathcal{T}$-action	To search	$\curvearrowright$
$\mathcal{SO}$-action	To point	$\curvearrowleft$
$\mathcal{SO}$-action	To chat	

Now that we have encoded the a_i-actions, we propose a retranscription method. This method is completely deterministic, since it is not open to interpretation and could be carried out by a computer. The actions ($a_i \in \mathbb{A}$) carried out on objects ($o_i \in \mathbb{O}$ with $\mathbb{O}$ that could be empty) by the observed student ($s_k \in \mathbb{S}$) characterise the student's local activity. We define this local activity as a *micro-activity* characterised by a triplet $(a_i, o_j, s_k) \in \mathbb{A} \times \mathbb{O} \times \mathbb{S}$. Between each micro-activity, the time *t* and the discretised occurrence ($occ \in \mathbb{N}$) are specified by an occurrence and a time marker (Da Ronch, 2022).

$$\cdots \xrightarrow[occ]{t} (a_i, o_j, s_k) \xrightarrow[occ+1]{t'} (a_p, o_m, s_k) \xrightarrow[occ+2]{t''} \cdots$$

Between each marker, there may be one or several micro-activities starting at the same time or for which we are unable to give an order, since we do not know precisely when they were carried out. We therefore assume that they all start at the same time. We refer to all micro-activities between two markers as *phases*. Here a phase is not an episode in the sense of Schoenfeld (1985).

$$\cdots \xrightarrow[occ]{t} \overbrace{\{\text{List of simultaneous actions.}\}}^{\text{Phase}} \xrightarrow[occ+1]{t'} \cdots$$

Below is an example of a retranscription of the micro-activities of two students trying to solve the problem proposed in the previous section (the dialogue transcription is in Da Ronch, 2022, p 325).

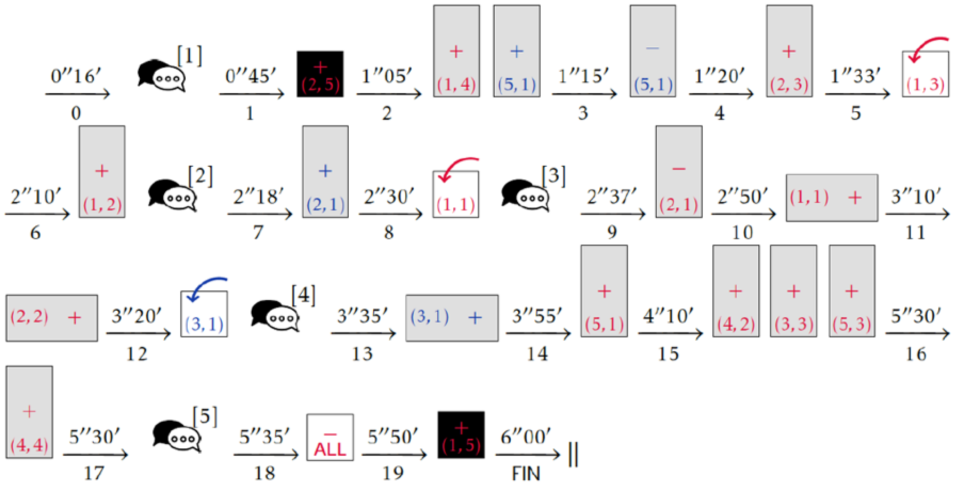

Figure 2: Example of a retranscription of the activity of two students s_1 and s_2

In this extract, two students (s_1 and s_2) are trying to tile a 5×5 grid with dominoes. This grid has a hole (black unit square) and this hole is positioned by the students. In the retranscription, the students are identified by the red and blue colours corresponding to s_1 and s_2 respectively. In addition, when a student puts (+) or removes (-) a domino on the grid, we represent this domino in the colour corresponding to the student, with the plus (+) or the minus (-) symbol and coordinates. To quickly identify the position of the dominoes on the 5×5 grid, we take the bottom left square of the grid as the origin (1,1). For example, on the first occurrence, s_1 puts the unomino (hole) in position (2,5). Then, both students simultaneously put two dominoes vertically in positions (1,4) and (5,1). This means, for example, that the south square of the domino placed by student s_1 on the grid is in position (1,4).

We then propose to enhance our transcription with a *summary*, *i.e.,* an "enhanced" photograph of the Work Zone ($\mathcal{W}_Z$) at a given moment, providing a static view of the situation at a given time. This summary allows us to retrace the students' reality at a *t* time. It also allows us to know exactly the number of $\mathcal{W}$-actions performed by the student knowing the number of $\mathcal{W}$-actions performed by the other students. This gives the number of $\mathcal{W}$-actions carried out between two effectives $\mathcal{W}$-actions and also the distribution of these actions: did s_1 perform more, fewer or as many $\mathcal{W}$-actions as s_2, for example? In the summary below, taken from the retranscription, we can see the reality of the students at a given moment in the retranscription (Figure 3). The symbols of type $\boldsymbol{a}|_{\boldsymbol{b}}$, where $\boldsymbol{a}$ and $\boldsymbol{b}$ are integers, represent the number of $\mathcal{W}$-actions of s_1 (respectively s_2) knowing the number of $\mathcal{W}$-actions of s_2 (respectively s_1). For example, during its first $\mathcal{W}$-action, s_1 puts the unit square in position (2,5), knowing that s_2 has not yet performed a $\mathcal{W}$-action (Figure 2). We thus have coded $\mathbf{1}|_{\mathbf{0}}$.

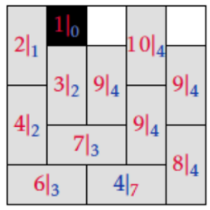

Figure 3: Example of a summary linked to the retranscription (see Figure 2)

The summary reproduces the Work Zone in which the students find themselves at a given moment, which is a perfectly complementary to the retranscription.

Analysis method of retranscription for objectifying students' proving processes

To analyse the retranscription (Figure 2) to objectify the students' proving processes, we need to remove some unnecessary information. This could be, for example, some dominoes which have been added and then removed, or some showing actions which are not significant in identifying a strategy linked to the proving process. We are therefore looking for a subword in this retranscription to highlight a strategy linked to a proving process identified in the mathematical analysis (Figure 1). This subword is shown in Figure 4.

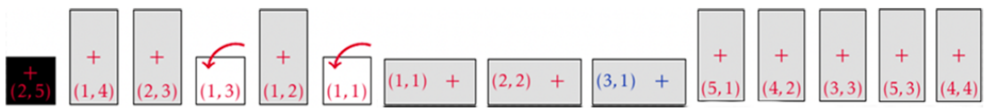

Figure 4: Subword from the retranscription sequence (see Figure 2)

This subword shows a similarity with a part of the strategy for proving the non-existence of tiling in this configuration (figure 1). However, it does not describe the entire strategy given by the mathematical analysis. In fact, the students dealt with only one of the two cases to prove the non-existence of tiling. Moreover, it is reasonable to say that it is only s_1 who has developed a solving strategy leading to a partial proving process of non-existence. Indeed, the subword described shows that s_1 has carried out almost all the actions needed to develop this partial strategy. In addition, the transcription of the dialogue shows that s_1 reasons well, using reasoning based on necessary conditions (“we have to”, see Da Ronch, 2022, p. 325) until it is no longer possible to add dominoes. Furthermore, the end of the retranscription shows that s_1 removes all the dominoes and changes the position of the hole (black unit square). This suggests that he has not realised that his strategy for proving the non-existence of tiling was incomplete. Finally, this subword shows a correct proving process but incomplete that implicitly uses a proof by non-contradiction and a reasoning by necessary conditions without an exhaustive study of the cases. Even if this proof is partial, s_1 has nevertheless developed a mathematical activity. Of course, this example of transcription is given for a particular instance; we would need to have the full transcription to be able to identify other strategies that would also be indicative of a mathematical activity.

Results and discussion

We have presented a deterministic model described in three steps: observable, retranscription and analysis. This model makes it possible to describe and objectify students' strategies in detail, and particularly their proving processes in problem-solving. We illustrated this on a tiling problem in

Discrete Mathematics. This enabled us to show, on an example, that this theoretical model was operational for describing and recognising significant elements of the proving process in problem solving. (for other examples, see Da Ronch, 2022). Moreover, in this model, we can clearly locate the interpretation part. It is precisely at the level of the choice of observables (choice of actions and strategies that seem relevant to what we want to observe). The remainder is completely deterministic, excepted sometimes at the level of retranscription analysis, since the strategies identified may be more or less close to those identified in the mathematical analysis. With this model, a single coder is sufficient, as there can be no interpretation in the coding of an elementary action, unlike episode coding (Schoenfeld, 1985), which requires at least double coding to measure reliability. This model also shows its relevance and effectiveness when it comes to solving problems using tangible objects, whose actions are similar to those carried out in "paper and pencil". Hence, Discrete Mathematics seems to be a suitable field, since a lot of mathematical objects (graphs, tiles, polyominos, *etc.*) can easily be represented in the form of tangible objects (Da Ronch, 2022; Gravier & Ouvrier-Buffet, 2022). In addition, it is probable that such problems, represented in a numerical format, facilitate this retranscription but open up the question of computer transposition. It would be appropriate to consider this type of processing and analysis in an automated way using digital tools, Artificial Intelligence and image recognition, for instance. This would make it possible to search large volumes of data, for a wide variety of users and over a very long period. By the way, recent researches is being carried out by the TWEAK team (LIRIS in Lyon, France), with the kTBS4LA (kernel for Trace Based Systems for Learning Analytics) platform. Based on *Learning Analytics*, kTBS4LA enables observable traces to be collected, stored, manipulated and analysed in order to calculate indicators of individual activity (Casado et al., 2017; Lefevre, 2018). Our model can also be used to perform quantitative analyses by mining data collected on a specific action for instance. It also has its limitations, particularly when the execution time of the actions is different, making it possible the reading and the writing of the retranscription complex. This model also raises questions about its operationality towards complex abstract objects that cannot take on a tangible format, giving an encoding more complex. Finally, it is important to specify that this theoretical model is not intended to be used by teachers, but only to provide researchers with the necessary tools to identify the traces of mathematical activity in problem solving.

References

Berthelot, R., & Salin, M. H. (1992). *L'enseignement de l'espace et de la géométrie dans la scolarité obligatoire* [Teaching space and geometry at school] (Doctoral dissertation, Université Sciences et Technologies-Bordeaux I). https://theses.hal.science/tel-00414065

Bloch, I. (2006). *Quelques apports de la théorie des situations à la didactique des mathématiques dans l'enseignement secondaire et supérieur* [Contribution of the TDS to the study of the teaching of mathematics at the transition between secondary school and University] (HDR, Université Paris-Diderot-Paris VII). https://theses.hal.science/tel-00012153/

Brousseau, G. (1997). *Theory of the Didactical Situations in Mathematics*. Kluwer Academic Publishers. https://doi.org/10.1007/0-306-47211-2

Casado, R., Guin, N., Champin, P. A., & Lefevre, M. (2017). KTBS4LA: une plateforme d'analyse de traces fondée sur une modélisation sémantique des traces [KTBS4LA: a trace analysis platform based on semantic trace modeling]. In *Méthodologies et outils pour le recueil, l'analyse et la visualisation des traces d'interaction*. https://hal.science/hal-01637548/

Da Ronch, M. (2022). *Pratique de l'activité mathématique en médiation: modèles didactiques et conception d'ingénieries* [Practice of mathematical activity in mediation: didactic models and engineering design] (Doctoral dissertation, Université Grenoble Alpes [2020-....]). https://cnrs.hal.science/tel-04089443/.

Da Ronch, M. (2024, July). *Model for retranscribing and analysing the mathematical activity of subjects in problem solving.* 15th International Congress on Mathematics Education (ICME-15), Sydney, Australia. https://hal.science/hal-04642665

Gravier, S., & Ouvrier-Buffet, C. (2022). The mathematical background of proving processes in discrete optimization—Exemplification with Research Situations for the Classroom. *ZDM–Mathematics Education*, *54*(4), 925–940. https://doi.org/10.1007/s11858-022-01400-3

Grenier, D. & Payan, C. (1999). Discrete mathematics in relation to learning and teaching proof and modelling. In I. Schwank (Ed.), *Proceedings of CERME 1*, Vol. 1 (pp. 143–155). Osnabrück.

Hanna, G. & De Villiers, M. (Eds.) (2012). *Proof and proving in mathematics education*. Springer. https://doi.org/10.1007/978-94-007-2129-6

Hart, E. & Sandefur, J. (Eds) (2018). *Teaching and learning discrete mathematics worldwide: Curriculum and research*. Cham: Springer. https://doi.org/10.1007/978-3-319-70308-4

Jewitt, C. (2012). *An introduction to using video for research*. NCRM Working Paper. Link

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal, 27*(1), 29–63. https://doi.org/10.2307/1163068

Lefevre, M. (2018, April 3). *KTBS4LA a kernel for Trace-Based Systems for Learning analytics*. Rencontre des jeunes chercheurs en EIAH [Conférence]. Besançon. https://rjceiahatelierdne.files.wordpress.com/2018/04/demo-ktbs4la.pdf

Margolinas, C. (1995). La structuration du milieu et ses apports dans l'analyse a posteriori des situations [Environmental structuring and its contribution to a posteriori situation analysis]. *Les débats de didactique des mathématiques*, *La Pensée sauvage Grenoble* (pp.89–102). https://shs.hal.science/halshs-00418815/

Rott, B., Specht, B., & Knipping, C. (2021). A descriptive phase model of problem-solving processes. *ZDM–Mathematics Education*, *53*, 737–752. https://doi.org/10.1007/s11858-021-01244-3

Schoenfeld, A. H. (1985). *Mathematical problem solving*. New York Academic Press. https://doi.org/10.1016/C2013-0-05012-8

Stylianides, A.J. (2016). Proving in the elementary mathematics classroom. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198723066.001.0001