

Learning analytics in formal education: our vision, and the role of FACE-IT¹

the FACE-IT team

December 10, 2019

¹Fostering Awareness on program Contents in higher Education using IT tools
Erasmus+ Strategic Partnership project number 2019-1-NO01-KA203-060257

Different stakeholders have different needs; for example these ones may ask themselves these questions . . .

Some needs from some stakeholders

teachers: how does modifying my own course affect the overall program?

program boards: how shall we modify and assess the structure of our programs using also quantitative evidence, and not only our personal opinions?

students: how do courses connect to each other, and why is this Intended Learning Outcome (ILO) important from my learning perspective?

There are several other ones, also very important ones. Somehow, though, when our team tried to generalize them, they often boiled down to one specific thing: the various persons involved in the education process often lack of a holistic viewpoint on the program content. For examples, for teachers it is difficult to know (and keep updated) on what the other teachers do. Boards are composed by teachers, so they suffer from the same. Students may also read the contents of the courses they have to take, but it is difficult to get a grasp of what means what and how things connect. Having a good holistic viewpoint would help a lot.

Overarching need

be fast and efficient
in getting and maintaining
holistic maps of the program contents

So we want to find tools that can help the stakeholders get (and communicate) holistic viewpoints.

Overarching need

be fast and efficient
in getting and maintaining
holistic maps of the program contents

our strive: open-access tools that can serve this purpose

Preamble: we will present now a strategy that may solve the problems above, but in the development of this we will use some taxonomy levels. We may use different taxonomies (i.e., division of the knowledge into opportune levels); a very famous one is Bloom's, but actually we have to understand which one will work best. For now, though, the important is to consider that there are different knowledge levels.

Preamble: assessing knowledge levels using some taxonomy

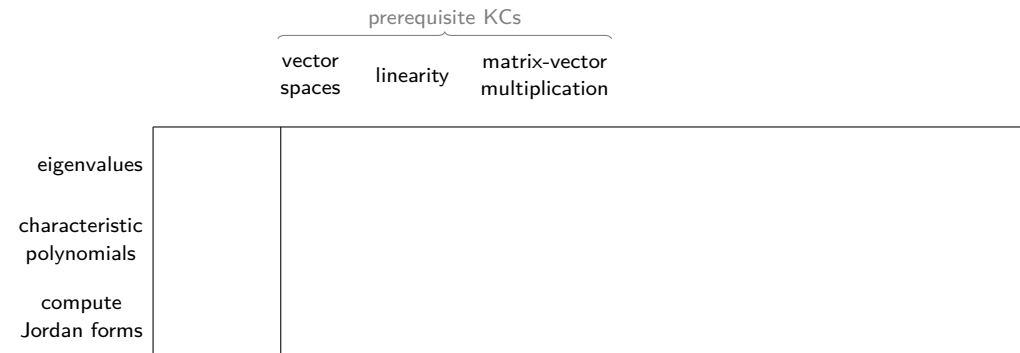
For example, Bloom's:

1	<i>Remember</i>	be able to recognize or remember facts, terms, basic concepts, or answers without necessarily understanding what they mean
2	<i>Understand</i>	be able to organize, compare, translate, interpret, give descriptions, and state the main ideas
3	<i>Apply</i>	be able to use prior knowledge to solve problems, identify connections and relationships and how they apply in new situations
4	<i>Analyze</i>	be able to break information into component parts, determine how the parts relate to one another, identify motives or causes, make inferences, and find evidence to support generalizations
5	<i>Evaluate</i>	be able to present and defend opinions by making judgments about information, the validity of ideas, or quality of work based on a set of criteria
6	<i>Create</i>	be able to build a structure or pattern from diverse elements and to put parts together to form a whole

First, some notation. The term *Knowledge Component* denotes "an acquired unit of cognitive function or structure that can be inferred from performance on a set of related tasks". In practice, facts, concepts, and procedures. On top of this term we can define the term *Domain Model*, that corresponds to the definition of some knowledge components relevant to some domain and the relations among them, plus potential maps of a set of given items to these knowledge components. *Domain modeling* means then assigning individual items to knowledge components and to model the relations among these components. Back to the slides, assume that I am the teacher of the imaginary *course X*, and I want to do *domain modelling* of my course. Having designed or taught course *X*, I know that this course is characterized by a list of prerequisites, that are defined in terms of Knowledge Components (KCs), i.e., facts, concepts, and procedures, plus a list of KCs that will be opportunely developed during the course. Let's assume thus that I write down the list of the prerequisite and developed KCs of this course in this way.

Towards a quantitative tool that can give holistic viewpoints

Representing an imaginary *Course X* through a course-wide Domain Model (DM)



After having compiled these lists, as a teacher I am aware that there exist different levels of knowledge associated to the knowledge concepts above, and that these levels can be encoded through a taxonomy, e.g., Bloom's, as we said before. Now assume that my experience as a teacher, or logical considerations, make me consider that to reach the first KC students should ideally have a prerequisite learning level (again, using Bloom's taxonomy just as an illustrative example) *2 - understand* for both prerequisite concepts *vector spaces* and *linearity*.

Towards a quantitative tool that can give holistic viewpoints

Representing an imaginary *Course X* through a course-wide Domain Model (DM)

	prerequisite KCs		
	vector spaces	linearity	matrix-vector multiplication
eigenvalues	2	2	
characteristic polynomials			
compute Jordan forms			

To be able to learn about *characteristic polynomials* students should ideally have a prerequisite learning level *2 - understand* for the skill *multiply matrices and vectors*, and have reached - while studying for *Course X* - a learning level *1 - remember* about the KC *eigenvalues*. Note that since there is this potential situation for which to reach the second KC students may need to learn the first KC the table that we are constructing needs to “repeat” the KC on the side of the prerequisites.

Towards a quantitative tool that can give holistic viewpoints

Representing an imaginary *Course X* through a course-wide Domain Model (DM)

	prerequisite KCs			developed KCs		
	vector spaces	linearity	matrix-vector multiplication	eigenvalues	characteristic polynomials	compute Jordan forms
eigenvalues	2	2				
characteristic polynomials			2	1		
compute Jordan forms						

Then, to be able to learn the skill *compute Jordan forms* students should ideally have a prerequisite learning level *1 - remember* for the prerequisite *linearity* and level *2 - understand* about the knowledge concepts *eigenvalues* and *characteristic polynomials*.

Towards a quantitative tool that can give holistic viewpoints

Representing an imaginary *Course X* through a course-wide Domain Model (DM)

	prerequisite KCs			developed KCs		
	vector spaces	linearity	matrix-vector multiplication	eigenvalues	characteristic polynomials	compute Jordan forms
eigenvalues	2	2				
characteristic polynomials			2	1		
compute Jordan forms		1		2	2	

To be also complete, I also add to this table information on which knowledge level students will ideally reach when successfully passing the course. As it will be clear later on, this information will be useful to analyze the structural properties of the programs that involve course X.

Towards a quantitative tool that can give holistic viewpoints

Representing an imaginary *Course X* through a course-wide Domain Model (DM)

	prerequisite KCs			developed KCs			intended final learning level
	vector spaces	linearity	matrix-vector multiplication	eigenvalues	characteristic polynomials	compute Jordan forms	
eigenvalues	2	2					3
characteristic polynomials			2	1			2
compute Jordan forms		1		2	2		3

Finally, I may add additional information such as “amount of teaching time” dedicated to each KC – or, in general, every information that we may discover be valuable for assessing the structural properties of the course and of the related program as a whole.

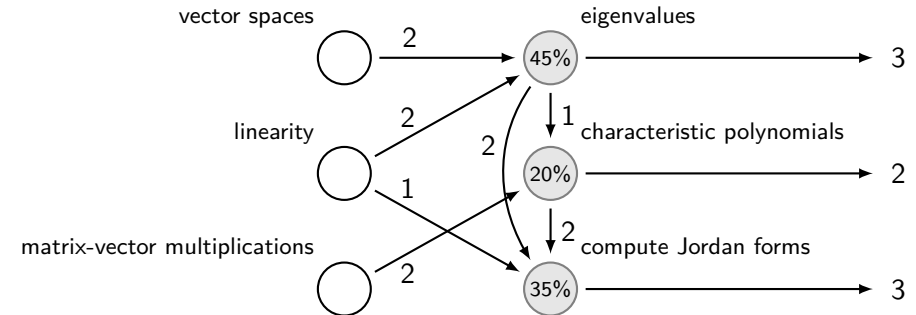
Towards a quantitative tool that can give holistic viewpoints

Representing an imaginary *Course X* through a course-wide Domain Model (DM)

		prerequisite KCs			developed KCs			intended final learning level
	teaching time	vector spaces	linearity	matrix-vector multiplication	eigenvalues	characteristic polynomials	compute Jordan forms	
eigenvalues	45%	2	2					3
characteristic polynomials	20%			2	1			2
compute Jordan forms	35%		1		2	2		3

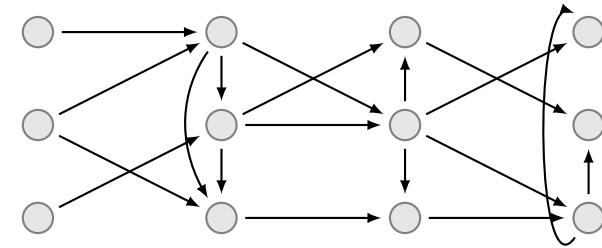
Obviously this table / domain model can be visualized as a directed weighted graph, that we call a course-wide Domain Model Graph (DMG). Already here one can see how the structure of the course can be interpreted in terms of "learning flows".

Towards a quantitative tool that can give holistic viewpoints



Assume now that all the teachers of all the courses within a program compile their own course-wide domain models (and thus graphs). Since programs are often designed in a way that the knowledge concepts developed in some courses are prerequisites for other courses, this means that we can combine the single tables / graphs of the single courses into a program-wide table / graph, as here – and notice that here we omit writing the numbers and the labels of the various knowledge concepts and skills so to keep the drawing simple. This representation is in practice a graphical description of the “*learning flows*” that students should ideally follow during their studies. But actually there is much more information in it.

Program-wide Domain Model Graphs

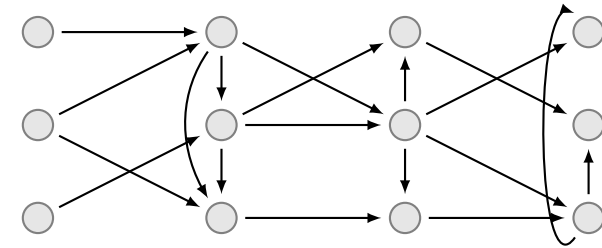


Indeed, the operation of merging the single graphs / tables into a program-wide graph / table adds information, because it enables to inspect two components that we could not inspect before:

- a logical component, i.e., the fact that courses relate logically, build up knowledge, overlap, etc.;
- a temporal component, i.e., the fact that courses tend to be taken at certain specific times and in certain specific orders.

Thus to the links of the network we implicitly associate two weights: the learning levels and how much time it passes when jumping from one knowledge concept to an other one.

Program-wide Domain Model Graphs

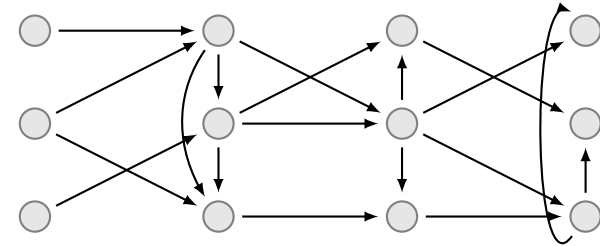


Information that is implicitly or explicitly contained in a program-wide DMGs:

- logical interconnections and overlaps among courses
- temporal progression of the learning levels

Intuitively, it is important to include the knowledge levels progressions because the programs (but also the individual courses) should be designed without having "knowledge jumps". If a course prerequisite is that students have a level knowledge 4 about a certain knowledge concept, then there should be courses that ideally lead students to that level. If the previous courses ideally bring only to level 2, then clearly there is a problem. A PW-DMG can immediately highlight this.

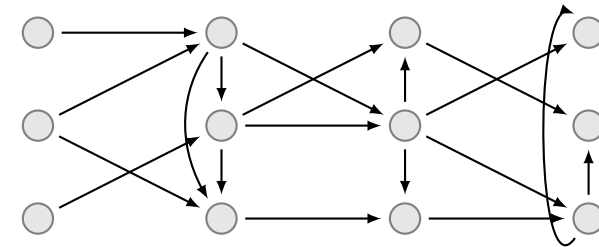
The usefulness of program-wide DMGs in detecting program's *logical* flaws



- are there “knowledge jumps”?

Intuitively, it is also important to inspect the temporal component because not refreshing a knowledge concept sufficiently often may lead students to forget things. Thus having a temporal component enables to analyze if it typically passes too much time for students to refresh certain knowledge concepts. A PW-DMG can immediately highlight problems in the sense of pedagogically non-ideal temporal flows.

The usefulness of program-wide DMGs in detecting program's *temporal* flaws



- are there “temporal jumps”?

Having a PW-DMG may also enable (even if not-so-immediately) to do these things. For example:

- easing the implementation of the Bologna agreements by using these representations as a common language to exchange structural information about local programs across different universities, plus infer their compatibility;
- moreover these representation may act as “communication interfaces” so that industries can tell universities what are their needs and wishes;
- then, an other important thing to notice: up to now we assumed that these tables have been compiled only by teachers. But maybe also students may compile them! Indeed if students compile their own versions, this information would provide insights on the students’ perception (and thus favor alignment with teachers and ease the spread of opinions).

A non-exhaustive list of usages of program-wide DMGs

universities: ease the implementation of the Bologna agreements

industries: ease communicating needs to the universities

students: ease detecting miscommunications

This is the scope of the Erasmus+ project that will start in September 2019. The idea is to create IT tools that can be used for all the purposes described above.

A non-exhaustive list of usages of program-wide DMGs

universities: ease the implementation of the Bologna agreements

industries: ease communicating needs to the universities

students: ease detecting miscommunications

all this is the scope of FACE-IT

Fostering Awareness on program Contents in higher Education using IT tools: participants

Financed through EU funds:

- NTNU (Norway)
- Uppsala University (Sweden)
- Otto von Guericke University Magdeburg (Germany)
- Université Libre de Bruxelles (Belgium)
- University of Padova (Italy)

Financed through own funds:

- Luleå University of Technology (Sweden)

On top of the things above, PW-DMGs may be – in our opinion – a fundamental general tool for introducing quantitative and data-driven strategies for the individualized education of any student, in the sense that we define now.

An other *futuristic* usage of program-wide DMGs

everybody: enable data-driven individual learning

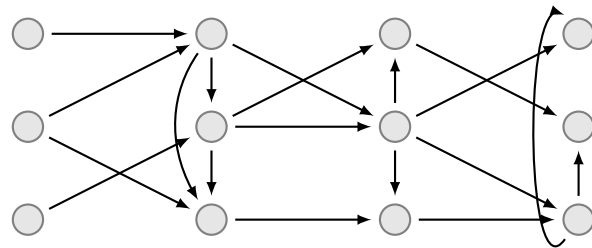
An important thing to notice is that all the information that we described up to now is "a priori", in the sense that up to now we did not involve any "collected measurement", i.e., any result from tests / exams / interrogations / exercises. Somehow, up to now the purposes that we described are independent of how the students are actually performing. We will see now how the DMGs can be ideally useful tools to provide holistic insights on the learning performance of the students.

towards a data-driven approach to education

Let's start from a numerical interpretation of the DMG. More precisely, the set of ILOs and prerequisites can be thought as the domain of a random field that has, as its codomain, the set of admissible knowledge levels present in the taxonomy that we chose to work with. Informally, the DMG can be seen as a way of summarizing how knowledge levels change in time for the various students. For example, let the node $n = 7$ (i.e., the seventh node in the graph) may be the knowledge concept "complex numbers", and the time be $t = 1$ (i.e., first week at the uni). Then if the student $s = \text{Ann}$ has a knowledge level 2 for that knowledge concept at that time, we say that $K_{s=\text{Ann}}(n = 7, t = 1) = 2$. Of course it may be that after 10 weeks of studying, Ann is such that $K_{s=\text{Ann}}(n = 7, t = 10) = 4$. But she may also forget things in time, so that $K_{s=\text{Ann}}(n = 7, t = 20) = 3$. If we generalize, we see that $K_s(n, t)$ is a description of how the knowledge for the generic student s changes in time t for the various ILOs n . Note that there exists scientific literature about this topic, that goes under the name of *domain modelling*.

Towards a data-driven approach to education

Program-wide DMGs as random fields

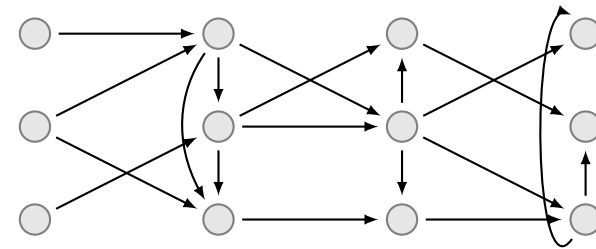


Mathematical interpretation: the DMG is a random field $\mathcal{K} : (s, n, t) \mapsto \ell$ describing the knowledge level ℓ for the generic student s at time t and for the ILO n

Now when a student studies / participates to a lesson / discusses with somebody, she/he is actively affecting the field, in the sense that (hopefully) she/he increments the field somewhere. However a student may forget something, i.e., her/his field may decrement somewhere if she/he is not refreshing the knowledge somehow. Note that there exists some scientific literature also about this topic, that goes under the name of *learner* or *student modelling*.

Towards a data-driven approach to education

The effect of studying

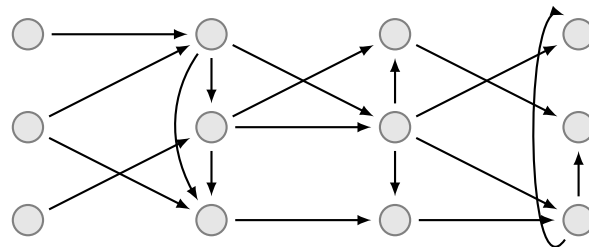


Intuitions:

- student s studies the KC n at time $\bar{t} \implies \mathcal{K}_s(n, t)$ increases for $t \geq \bar{t}$
- student s doesn't study anymore the KC n from time \bar{t} on $\implies \mathcal{K}_s(n, t)$ decreases for $t \geq \bar{t}$

In other words, if I study Fourier transforms, that are defined using complex numbers, then I will likely refresh a bit also complex numbers. Formally, these effects can be cast as a series of probability transitions characterizing the field dynamics, transitions that may be estimated from data – if we have data. Of course this is a complex problem, since for sure these transitions are affected by a lot of individual socio-psycho-physiological factors. However, the intuition is that the DMG formalism can help getting quantitative models of how learning works in an “individualized program / individualized person” perspective. This is the realm of the scientific branches of *adaptive learning* and *intelligent tutoring systems*.

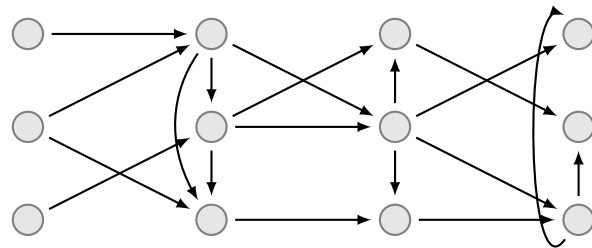
First interesting question: how do KCs correlate from a learning perspective?



Intuition: increasing the knowledge about n should improve / refresh the knowledge about what leads to n

Under the previous light the program flow graph is an educated guess on which causality relations hold among KCs for the average person. In a sense, it is a summary of personal interpretations from the various teachers of which knowledge concepts correlate from a learning perspective, and of how the knowledge levels are supposed to increase in time. This means that a DMG is a prior on the structure of the statistical dependencies among the various scalar variables composing the random field \mathcal{K} .

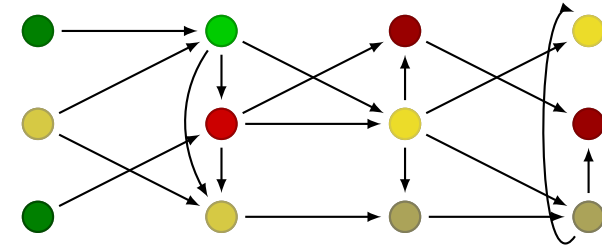
What is actually a DMG?



First intuition: if seen as a graph, then it is an educated guess on which causality relations hold among the KCs

Different persons have not only different learning dynamics: they have also different current “knowledge state”. Here every node n in the graph has a different color, to indicate a different knowledge level ℓ at a specific time, say $t = 10$. This set of levels is a function, in brief $\mathcal{K}_{s=\text{Ann}}(n, t = 10)$, and it is a summary of what Ann knows 10 weeks after having started her university program about the whole program.

What is actually a DMG?

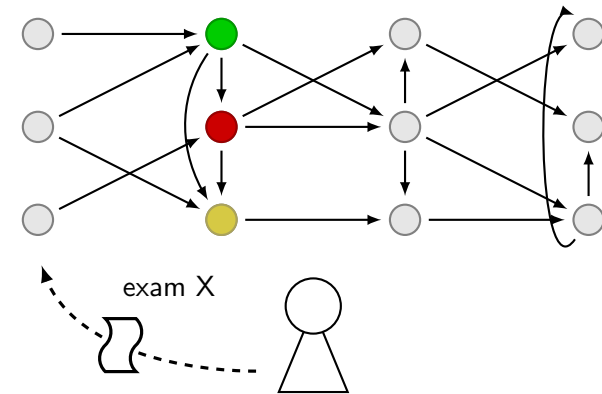


Second intuition: if seen as a function, then it is *the status of the knowledge of a student about the whole program*

The diagram illustrates a neural network with four layers of nodes. The first layer (input) has 3 nodes, the second layer has 3 nodes, the third layer has 3 nodes, and the fourth layer (output) has 3 nodes. Arrows indicate the flow of information from left to right. A stick figure is positioned below the second layer, representing a user or agent interacting with the network.

For example, the exam of course X may reveal something about the ILOs of course X.

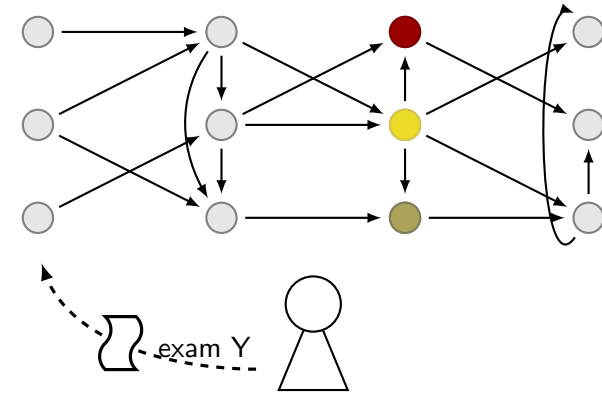
What is actually an exam?



Third intuition: any assessable test is a *noisy and aggregated measurement of the knowledge of a student on part of the program*

And the same for the exam of course Y. In a sense, every test works in a sense as a "sensor" transforming the state of the system into an output. And even if the current strategies of examining students (and processing of the relative results) are structured in a way that it is unclear how to map an exam / test / exercise into a "measurement" of the various individual variables of the field, however the underlying process can be conceptualized in this way.

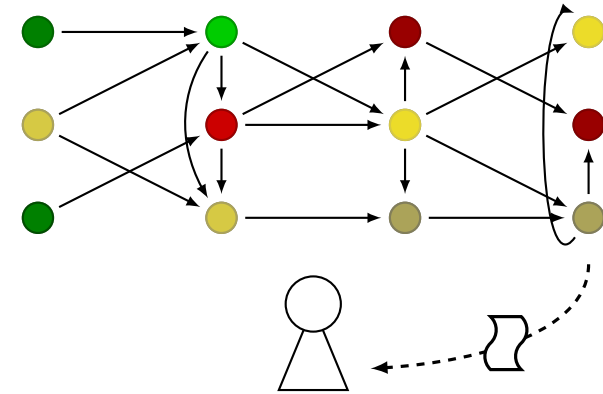
What is actually an exam?



Third intuition: any assessable test is a *noisy and aggregated measurement of the knowledge of a student on part of the program*

Even more importantly, in this context every learning activity that is suggested to the student can be seen as an operation of improving some part of this field.

What is actually an exercise?



Fourth intuition: any exercise is a *indirect push to improve some part of the DMG*

The fact is that with the framework suggested here is the one of the learning analytics / educational data mining communities, that eventually may be interpreted as going towards a control-oriented approach to education: if we know what the student knows and what not, and if we have a model (that shall be intended as a quantitative idea of what will be the effect of doing some specific exercise or generic learning activity), then we can compute what are the “best” learning activities (i.e., what to study) that that student should follow. This is what is meant with “intelligent tutoring systems”, that provide individualized and automatic suggestions to the learners. This is not a new concept, since it has been studied by several authors from the aforementioned *adaptive learning* and *intelligent tutoring systems* communities. However, to the best of our knowledge it has not been studied by people that come from the automatic control community. And we believe that our tools and methods can be very useful to bring improvements in these communities. More specifically, we as automatic control engineers may help with ...

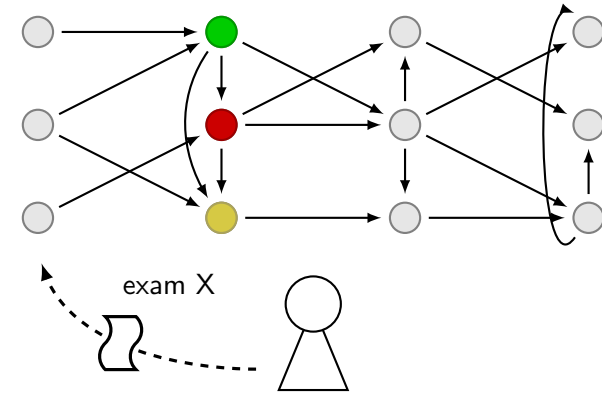
Towards learning analytics and intelligent tutoring systems: towards a control-theoretic approach to education

LA / ITSs is, in essence, feedback:

- ① estimate the current status of the individual student
- ② (if possible) model how that individual student learns
- ③ based on the estimated status and learning model, suggest tailored learning activities

To implement the approach above there is the need to develop methodologies and mathematical tools that solve the needs above. The first need is “estimation”. This means that we need tools to transform the noisy measurements coming from the results of exams and tests into estimates of the current status of the whole field for a specific student.

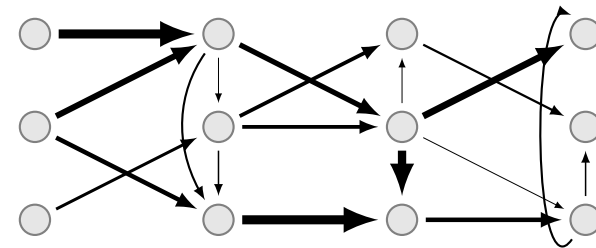
Research topic 1: estimation



How can we reconstruct the whole $\mathcal{K}_s(n, t)$ from the results of past and current exams and tests?

The second need is “modelling”. This means that we need tools to transform the noisy measurements coming from the results of exams and tests into estimates of how learning / forgetting a specific topic correlates with learning / forgetting an other topic. In other words, get an individual quantitative (and obviously approximated) description of the learning process. Note that it may be that we obtain only “average” models, i.e., not individualized ones, and that they may also be not very accurate descriptions of how these process work.

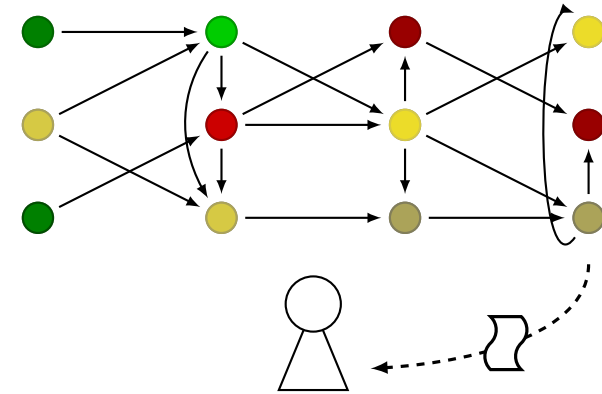
Research topic 2: modelling



How do we compute the learning model (i.e., the interactions among the variables in $\mathcal{K}_s(n, t)$) of each individual student?

The third need is “control”. This means that we need tools to transform the uncertain estimates of the current knowledge and uncertain models of how learning activities will affect the current knowledge into learning activities plans. This is a particular instance of the overarching research problem within the realm of automatic control.

Research topic 3: control



Which current learning activity shall we suggest to each individual student?

Summarizing, we foresee that DMGs can lead to these benefits, and that they may help implementing *adaptive learning* and *intelligent tutoring systems* at the university level.

Conclusions

- program-wide Domain Model Graphs can lead to holistic viewpoints and constructive alignment
- Domain Model Graphs can help implementing program-wide learning analytics (*specially if we add the “evaluations from exams” ingredient to the picture*)