notes

Developing CITE, a Concept Inventory Tool for Electrical Engineering

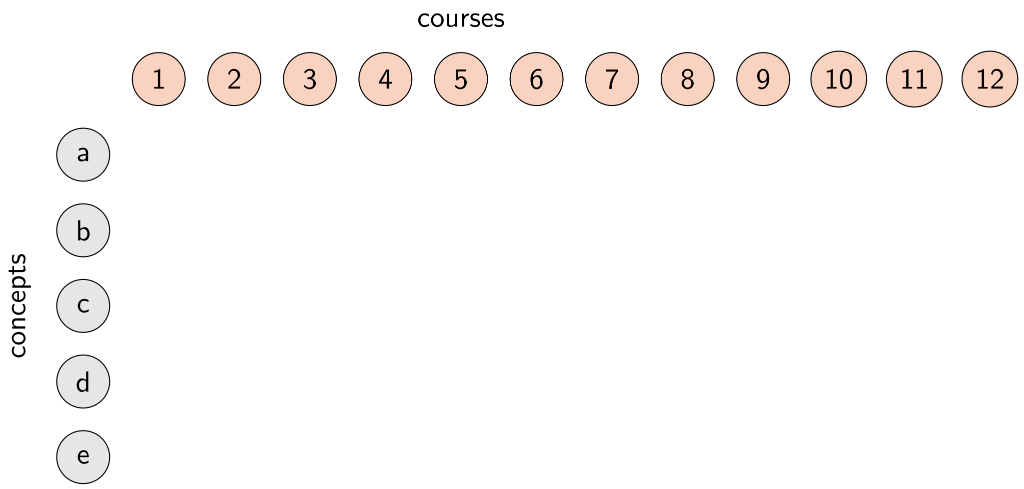
Eva Fjällström
Steffi Knorn
Damiano Varagnolo
Kjell Staffas

LULEÅ UNIVERSITY OF TECHNOLOGY
UPPSALA UNIVERSITET

notes

A whole program can be thought as a series of courses, each introducing, extending or building on top of specific concepts

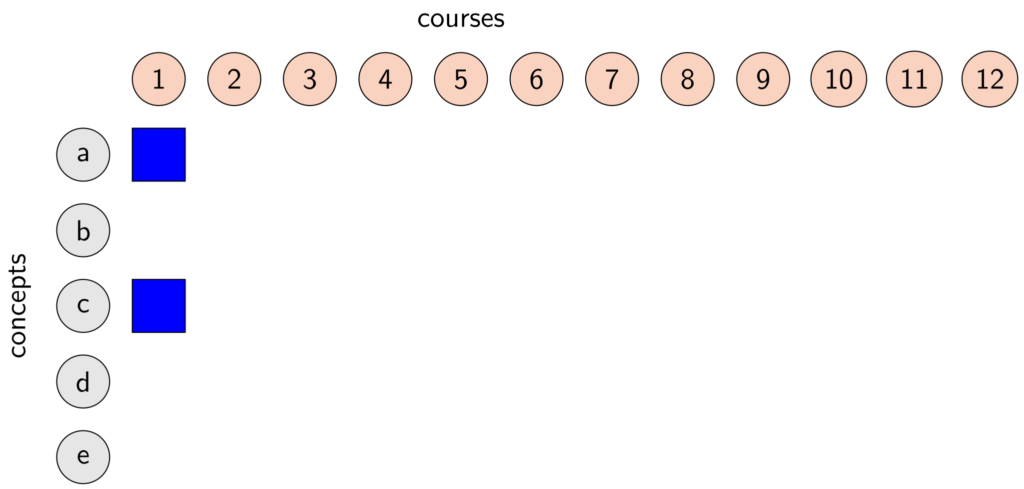
A schematic representation of a generic university program



notes

For example the first course may introduce two concepts, in this case a and c

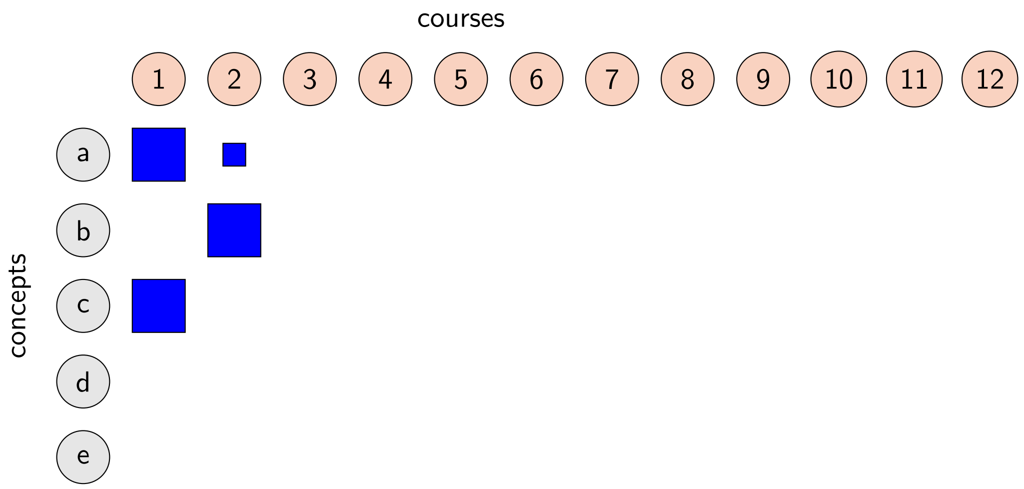
A schematic representation of a generic university program



notes

The second course may introduce concept b but ladder or extend a

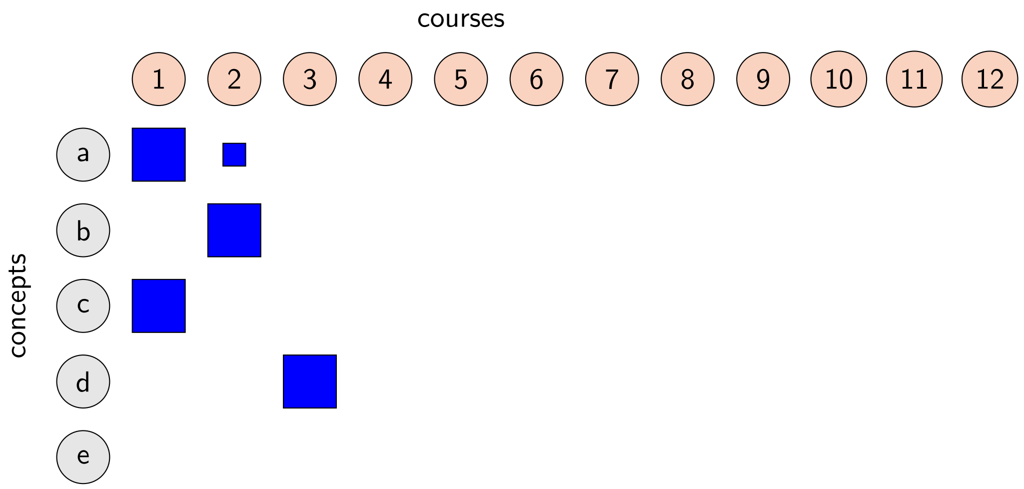
A schematic representation of a generic university program



notes

. . . and so on . . .

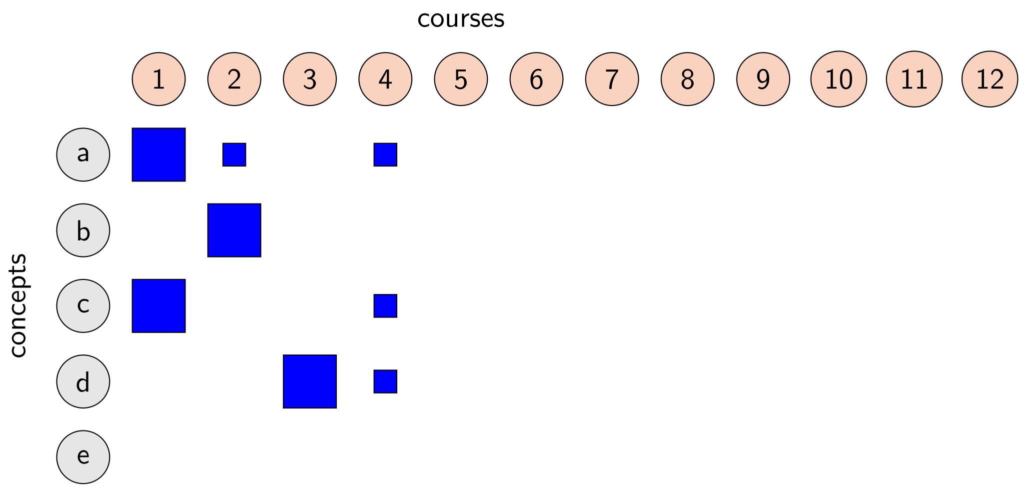
A schematic representation of a generic university program



notes

. . . and so on . . .

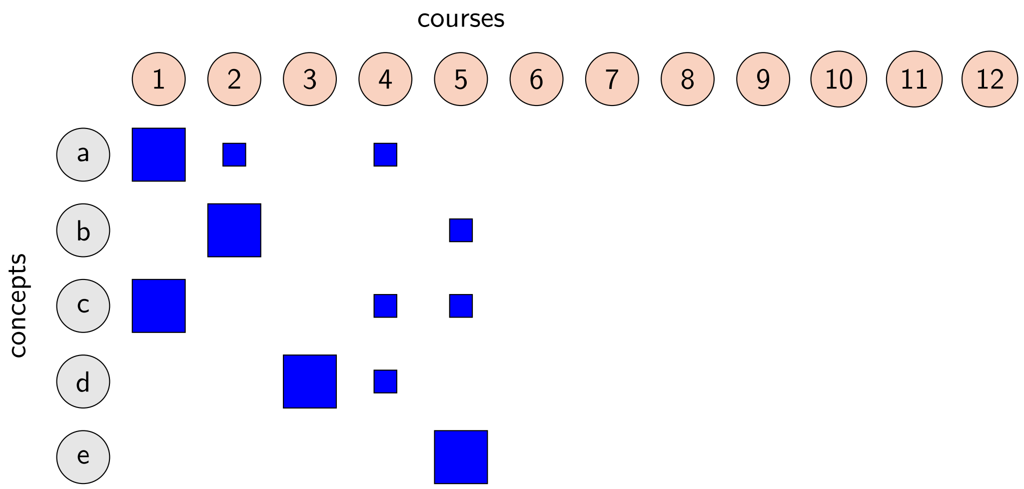
A schematic representation of a generic university program



notes

. . . and so on . . .

A schematic representation of a generic university program

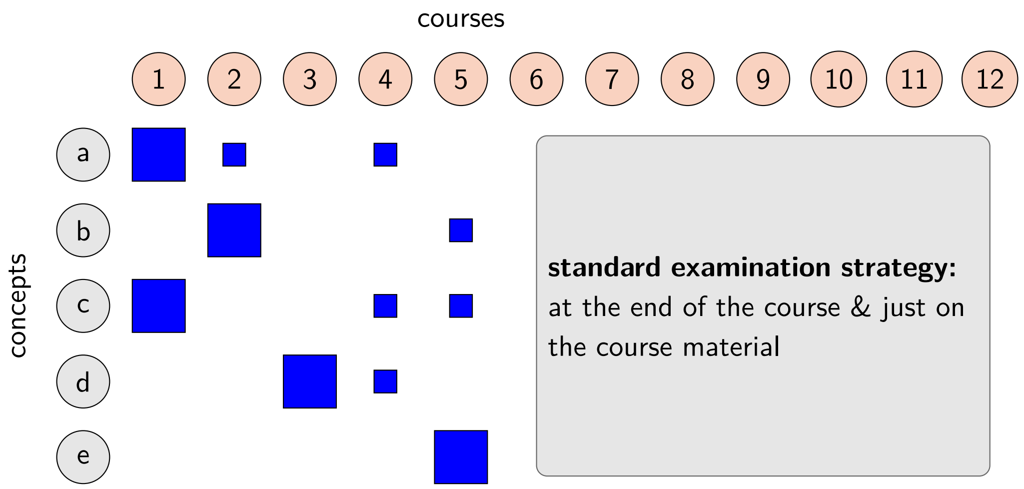


notes

. . . and on. Examinations, though, are typically:

- performed at the end of each course
- referring only to that set of concepts that have been introduced or extended in that specific course. E.g., here the examination of course 5 will be based on concepts b, c and e, and thus ignore if students have been forgetting a and d or not;
- this is obviously a natural choice.

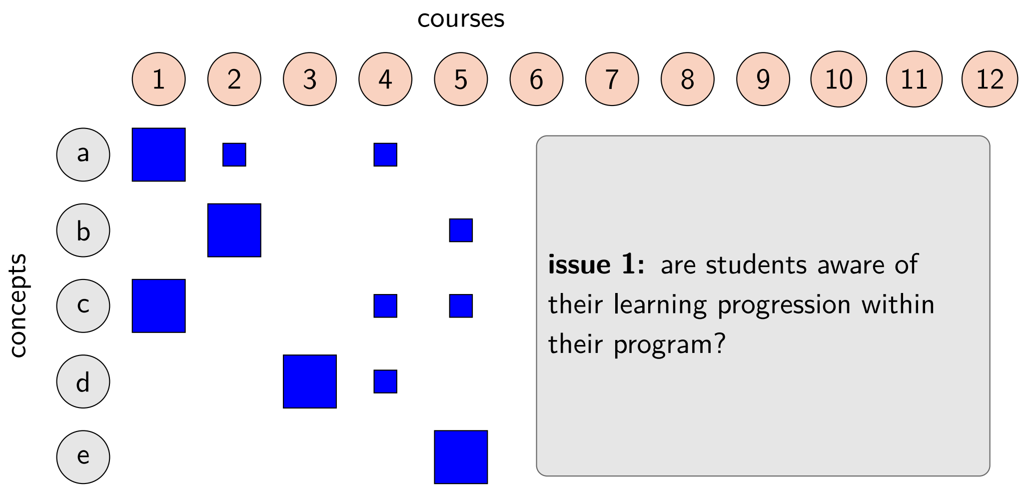
A schematic representation of a generic university program



notes

However this implies a series of issues. . .

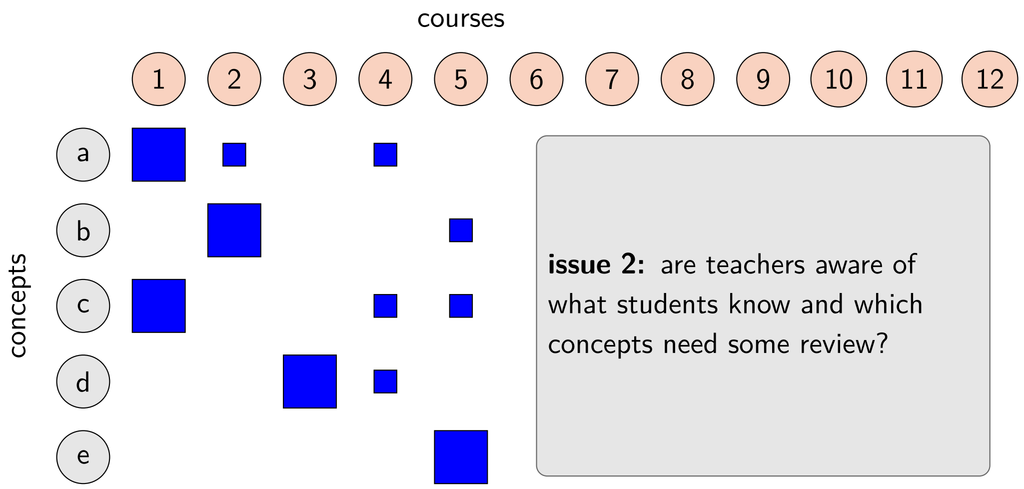
A schematic representation of a generic university program



notes

However this implies a series of issues. . .

A schematic representation of a generic university program

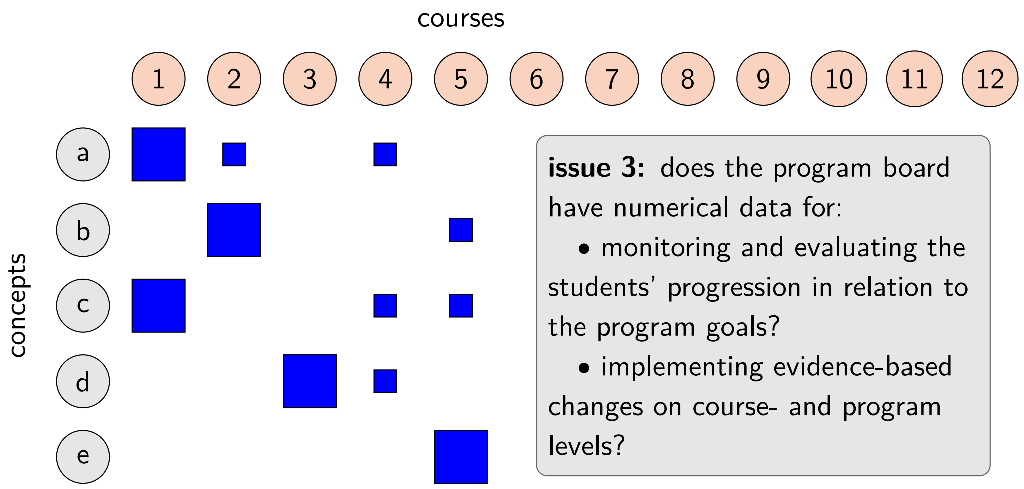


issue 2: are teachers aware of what students know and which concepts need some review?

notes

However this implies a series of issues. . .

A schematic representation of a generic university program



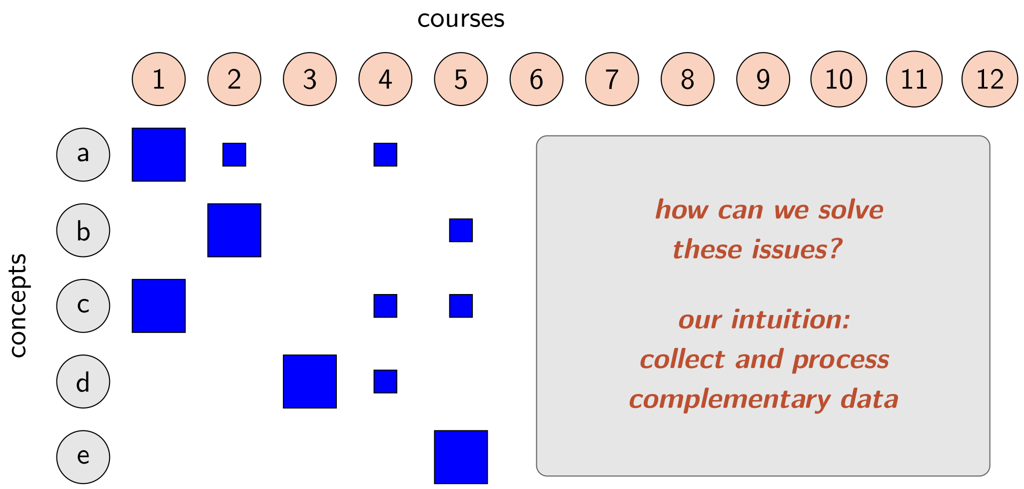
issue 3: does the program board have numerical data for:

- monitoring and evaluating the students' progression in relation to the program goals?
- implementing evidence-based changes on course- and program levels?

notes

Even more importantly, there is no systematic strategy at the program level for collecting evidence on how much and how fast students forget each individual concept. And without numerical evidence the problem-correction actions risk to be opinion-based instead of being facts-based.

A schematic representation of a generic university program

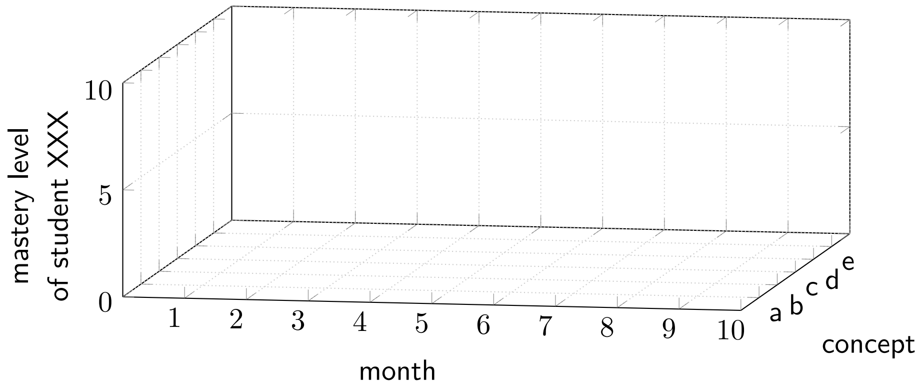


notes

From mathematical perspectives the best thing possible would be to have a situation like the one assumed here

Ideal situation (i.e., best possible one in terms of amount of information)

Unrealistic assumption: every month we can measure the knowledge level of each student on the whole program *(this assumption will be removed later on)*

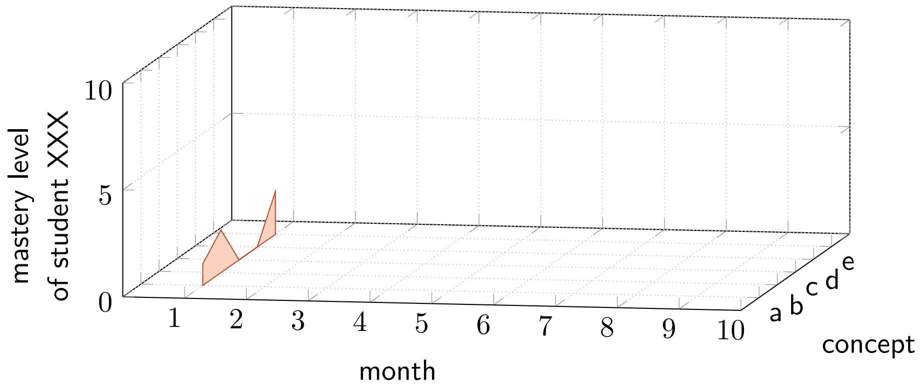


notes

Graphically speaking, we may think that at month 1 we measure this specific knowledge level for a specific student for all the various concepts of the program board

Ideal situation (i.e., best possible one in terms of amount of information)

Unrealistic assumption: every month we can measure the knowledge level of each student on the whole program *(this assumption will be removed later on)*

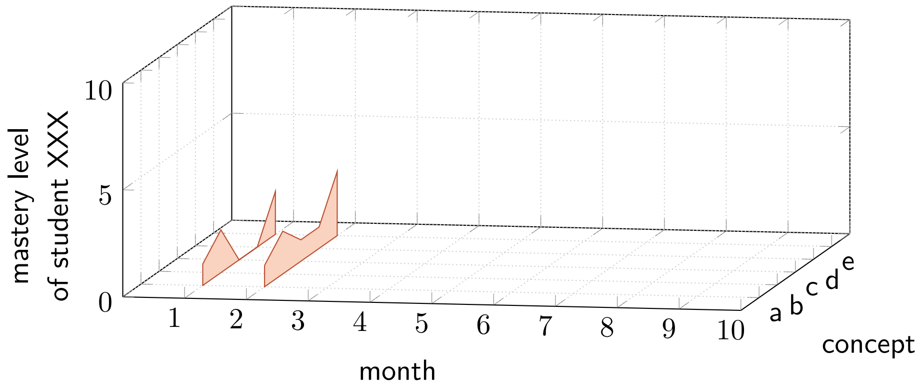


notes

Then after one month we measure this...

Ideal situation (i.e., best possible one in terms of amount of information)

Unrealistic assumption: every month we can measure the knowledge level of each student on the whole program *(this assumption will be removed later on)*

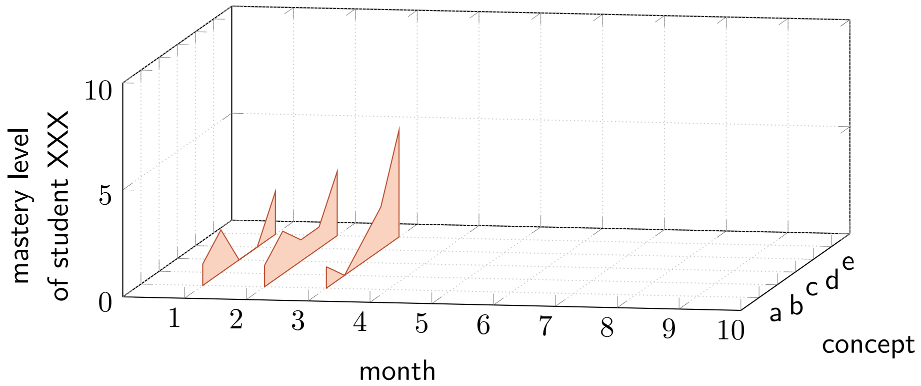


notes

Then after an other month we measure this. . .

Ideal situation (i.e., best possible one in terms of amount of information)

Unrealistic assumption: every month we can measure the knowledge level of each student on the whole program *(this assumption will be removed later on)*

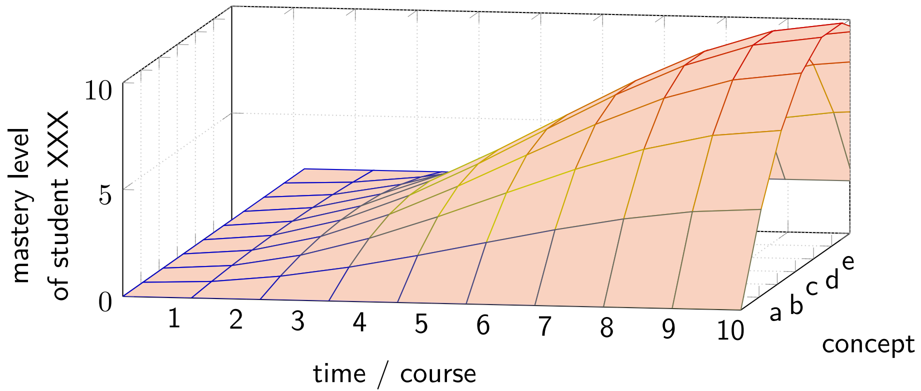


notes

Continuing like this we would get the whole "surface" of how the knowledge is evolving in time for that specific student

Ideal situation (i.e., best possible one in terms of amount of information)

Unrealistic assumption: every month we can measure the knowledge level of each student on the whole program *(this assumption will be removed later on)*

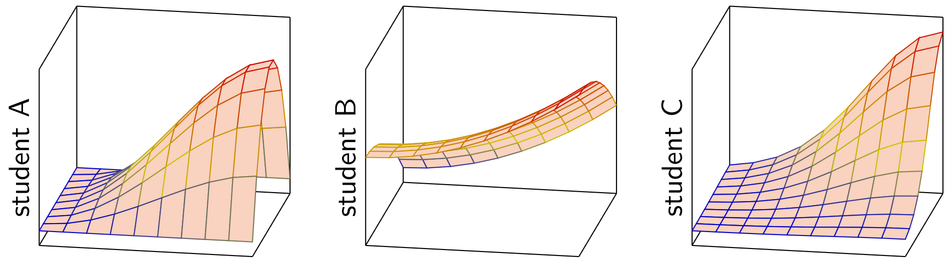


notes

But having this information for all the students, one would also have the average behavior, the standard deviations, etc., and have a lot of information that could be used in a lot of different ways.

Ideal situation (i.e., best possible one in terms of amount of information)

Unrealistic assumption: every month we can measure the knowledge level of each student on the whole program *(this assumption will be removed later on)*

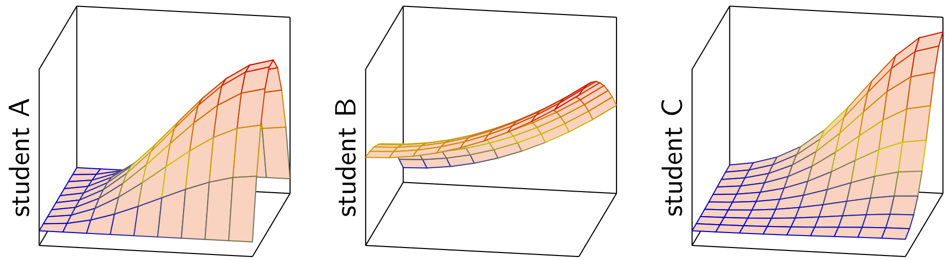


notes

But having this information for all the students, one would also have the average behavior, the standard deviations, etc., and have a lot of information that could be used in a lot of different ways.

Ideal situation (i.e., best possible one in terms of amount of information)

Unrealistic assumption: every month we can measure the knowledge level of each student on the whole program *(this assumption will be removed later on)*

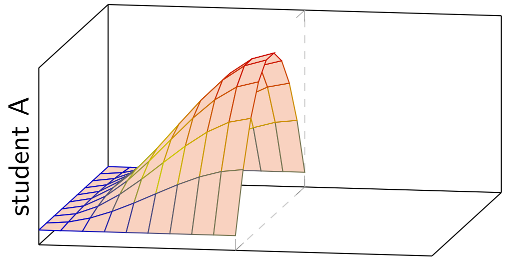


how could we use this data?

notes

The test that we want to develop should help students answering these questions . . .

Information that we could extract - students side



- do I know enough to take the next courses?
- how am I currently doing with respect to the program goals?
- what do I tend to forget, and how fast?
- where should I focus?
- what should I eventually know at the end of the program?

notes

. . . and help teachers answering these questions . . .

Information that we could extract - teachers side

- what can I expect my students to know?
- which variety exists in the prior knowledge among students and across the various years?
- where shall I start teaching?
- how shall I adapt to this specific class?
- what, how much & how fast do students forget what I teach?
- *(if having comparative data)* how do changes in how I teach affect long term performances?
- *(if having comparative data)* what are the best changes to do?

notes

. . . and program boards answering these questions

Information that we could extract - program boards side

notes

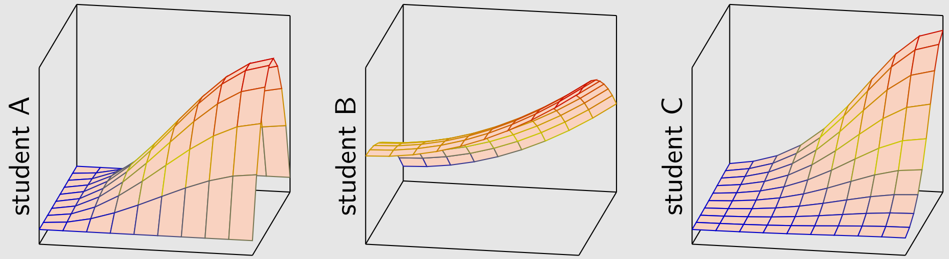
. . . and program boards answering these questions

Information that we could extract - program boards side

- how does the performance vary among students, also across the various years?
- is there correlation / causation among performances of different courses?
- is it better to teach a course before or after an other one?
- where and when students drop?
- *(if having historical / comparative data)* how do changes in the program / teaching strategies transform into long term effects in the performances?
- *(if having comparative data)* what are the best changes to do?

notes

the main message is this one



first message: having this data would be **very** useful

notes

Of course measuring the knowledge level of every student every month seems infeasible. Nonetheless testing strategies that are not as detailed as the one above may still provide information useful for everybody

How do we collect this data?

notes

Thus we want to develop a test that can be implemented, and that –even if it is an approximation of the ideal test– gives information useful to everybody. And which kind of purpose would we like to serve?

How do we collect this data?

CITE, a Concept Inventory Tool
for Electrical Engineering

notes

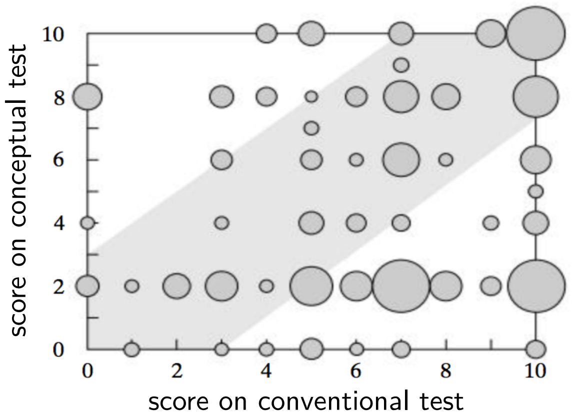
The first point is understanding how these tests should look like

First point: which knowledge shall be assesssed?

notes

Scholars have been making students of math-related classes take two types of tests: classic ones, and “conceptual” ones, where the “conceptual” ones are as we will show later on. Evidence shows that statistically if a student gets a good grade in the conceptual tests then she/he is very likely to get a good grade also on the classic ones, but not the viceversa. This suggests the potential (and intuitive) interpretation that if one has understood a concept then that person is more likely to solve classical procedural exercises. At the same time if one has understood how to solve a procedural exercise, that person may not have understood the concept. This suggests us to make tests that assess conceptual knowledge, and not procedural one.

First point: which knowledge shall be assesssed?



[From Mazur, E. (1997). Peer Instruction: A User's Manual. Prentice Hall, Upper Saddle River]

(radius of the circles = number of students represented by that circle)

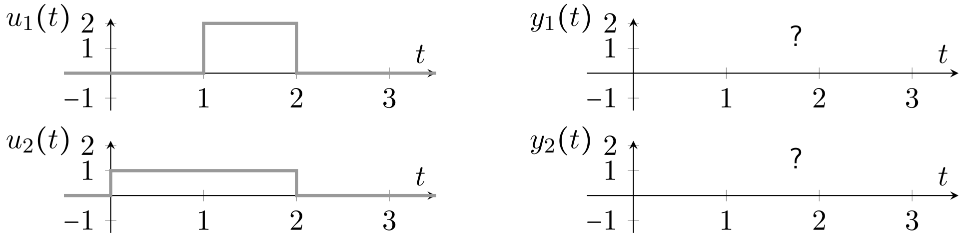
notes

The currently foreseen approach is to use *concept inventory tests*, i.e., multiple-choice tests that should specifically assess if a person has understood a specific concept. If the person has misunderstood something then that person is likely to choose a wrong answer, since the alternative answers are designed to trick the person

Example of conceptual question (1)

Consider an LTI system for which to this input signal $u(t)$ corresponds the following output signal $y(t)$:

Sketch the output signals $y_1(t)$ and $y_2(t)$ corresponding respectively to the following input signals $u_1(t)$ and $u_2(t)$, which would be fed into the same system as above:



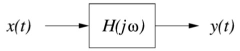
notes

In this second concept inventory test example if one has understood what a Bode plot is then that person will answer correctly immediately.

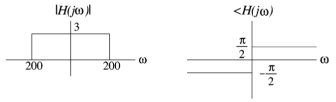
Example of conceptual question (2)

Question 6

Consider the LTI system with input $x(t)$ and output $y(t)$ shown in Fig. 3(a). The magnitude response $|H(j\omega)|$ and phase response $\angle H(j\omega)$ (in radians) of the system are shown in Fig. 3(b).



(a) System for Question 6.



(b) Magnitude and phase response of the system in Question 6.

Suppose that the input $x(t) = \cos(50t)$ for all time. What is the output $y(t)$?

(a) $3\cos(50t + \frac{\pi}{2})$
(b) $\cos(50t + \frac{\pi}{2})$
(c) $3\cos(50t)$
(d) $3\cos(200t)$

notes

Specifically, our project focuses on doing the following actions

Objectives of CITE

1. develop the test items
2. develop plans for the format and implementation of the tests
3. develop tools to provide relevant results to all stakeholders
4. do full scale test pilots

notes

And up to now we did this . . .

Status of the project @ September 2018

- developed a methodology (see the next slide)
- developed a database of questions for electrical engineering students
- developed a `CITEsuite` software for the generation of the tests
- performed several initial tests in Uppsala and Luleå on volunteering students
- interviewed students about their preferences on the implementation strategies
- . . . going to have the first full-scale test in 1 week

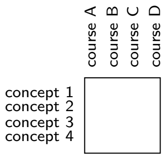
notes

We have been proposing and following this methodology. First, create the CCM (i.e., the courses-concepts matrix), that can be in practice an excel sheet where the rows are labeled with the names of the concepts, the columns labeled with the various courses of the program, and every element of this matrix being either:

- 0, to indicate that that specific course does neither make use nor introduce that specific concept;
- 1, to indicate that that specific course makes explicit use of that specific concept;
- 2, to indicate that that specific course teaches or heavily rely on that specific concept.

Actually, the very same process of compiling the CCM generates a lot of potentially useful information – think, for example, at the fact that if two teachers have different opinions on how the CCM should be compiled this means that they have some misaligned interpretation of the program. However in CITE we focus more on the test, and thus for now neglect potential spin-off research issues related to the CCM.

Our methodology

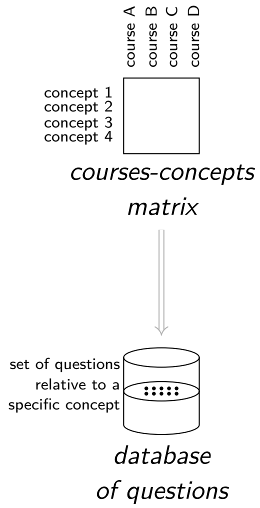


courses-concepts matrix

notes

Once the CCM is compiled, one can understand the relative importance of the various concepts within the program by integrating the CCM values by column, and taking the resulting number as an indication of how much that concept "weights" in the program. This means that one can then focus on creating a database of conceptual questions where for each concep there exist a number of questions that is proportional to the importance of that concept (since more important concepts will be assessed more often and more extensively)

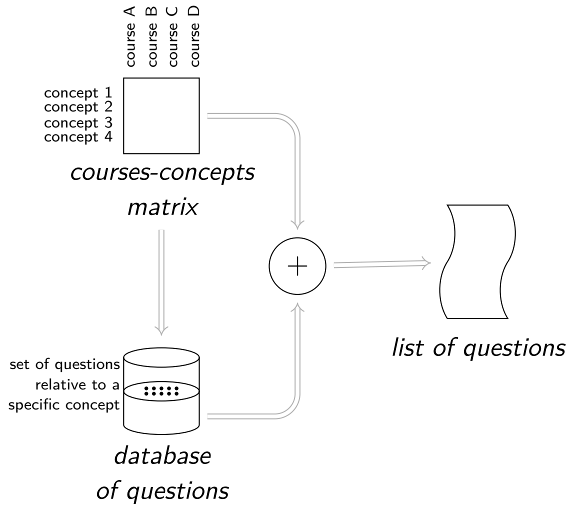
Our methodology



notes

From the CCM and the database one can then extract an actual test. As hinted before, the test should comprise questions extracted through a probability density that accounts for how much important a concept is, how often that concept has been assessed, and how well the students in average score on that concept. Note that we still have no clear picture of how the mathematics should look like, but we currently feel that there should be some "adaptive questions extraction" mechanism

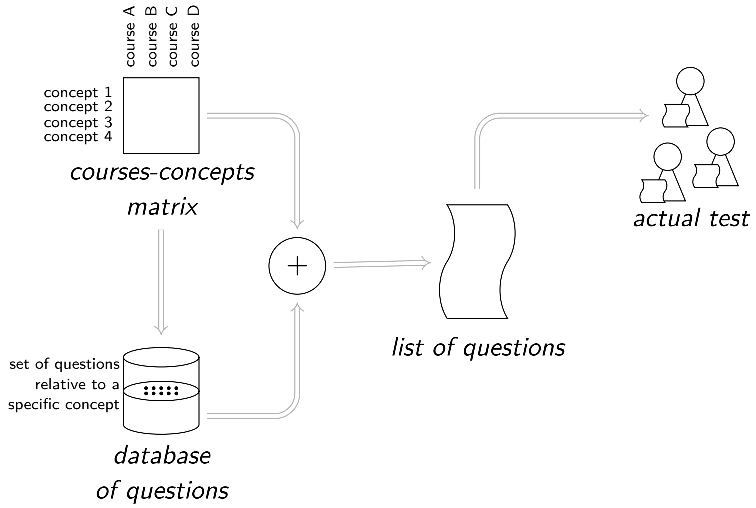
Our methodology



notes

Students can then take the test – how, we still have to understand (for now we are using web-based tests)

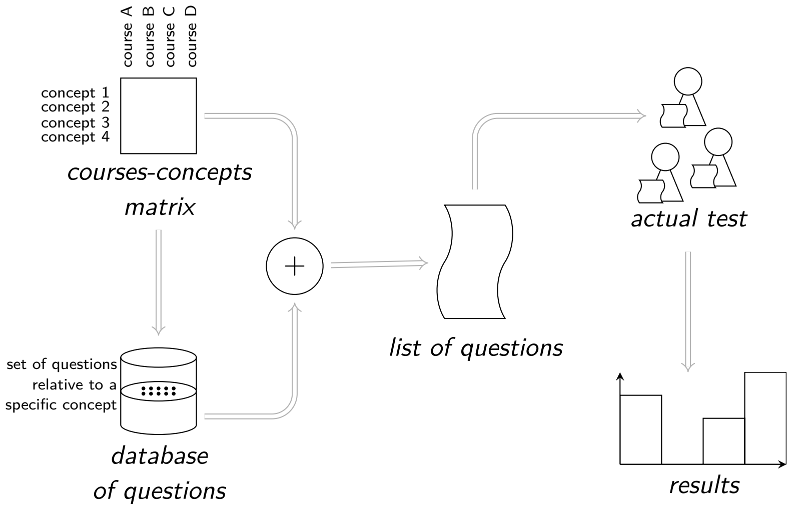
Our methodology



notes

We can then compute the results, ...

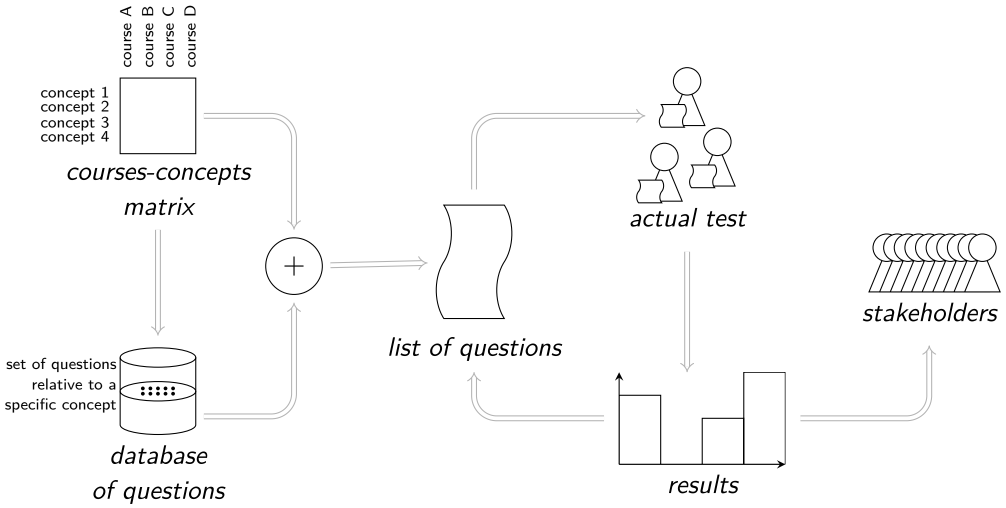
Our methodology



notes

. . . and use them to both inform the stakeholders and influence how the next test should look like

Our methodology



notes

The CCM, though, isn't useful *only* to fill up the database. The matrix here represented is a piece of the CCM corresponding to the Engineering Physics program at Uppsala University, and filled by Uppsala's teacher in early 2018. A CCM can be considered a collection of personal opinions of these teachers about the program where they teach – in other words, every column answers the personal question "what do I think, as the teacher of course XXX, my course is about?".

Since a CCM is a matrix of numbers, it is naturally prone to quantitative analyses. The problem is then: what can we say about a university program through analyzing its CCM?

Important spin-off: which information can we extract from a CCM?

	1TE705 Introduction to Elect. Eng.	1TE704 Components & Circuits	1MA008 Algebra & Vector Geom.	1TE667 Elect. Circ. Theory
complex numbers			2	2
vectors			2	1
systems of lin. eq.			2	2
Ohm's law		2		2
Kirchoff's laws		2		2
potential voltage		2		2
linearity				
matrices	1	1	2	2
work, energy		2		2
integral calculus	1			

notes

To understand what we can we say about a university program through analyzing its CCM we exploit the natural intuition of considering a university program as a flow of information. With this intuition the dichotomy courses/concepts can then be used to alternatively (but equivalently) represent a CCM as a bipartite graph, where one set of nodes are the concepts, and the other set of nodes are the courses

From Courses-Concepts-Matrices to Courses-Concepts-Graphs (CCGs)

complex num.

vectors

sys. of lin. eq.

Ohm's law

Kirchoff's laws

pot. voltage

linearity

matrices

work, energy

int. calculus

1TE705 *Intro to El. Eng.*

1TE704 *Components & Circuits*

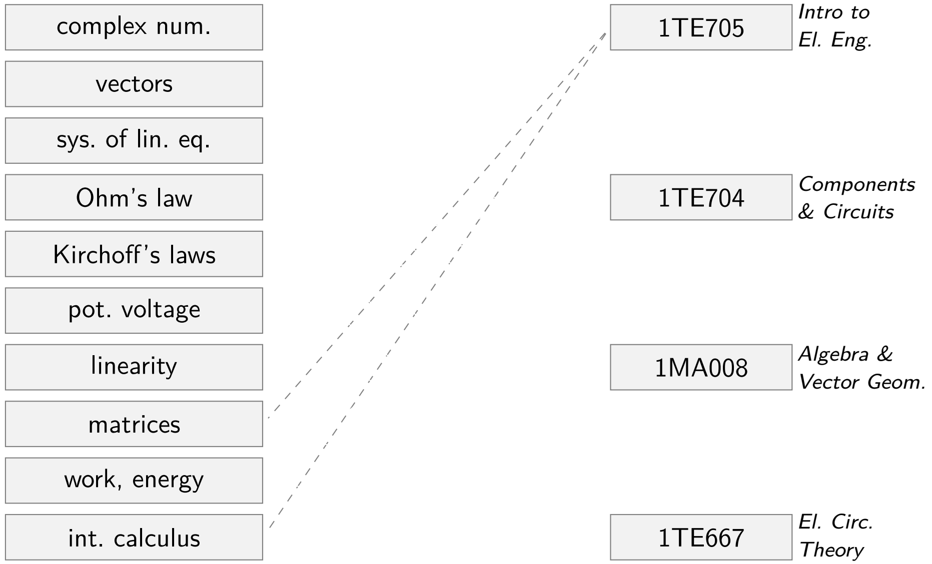
1MA008 *Algebra & Vector Geom.*

1TE667 *El. Circ. Theory*

notes

Then each column in the CCM is a set of weighted links from that specific course (i.e., column) to the various concepts

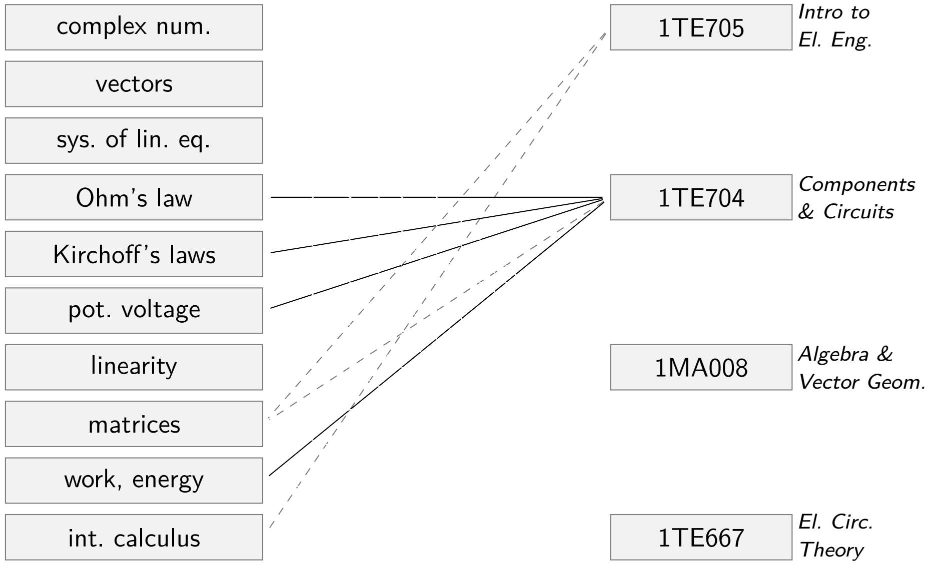
From Courses-Concepts-Matrices to Courses-Concepts-Graphs (CCGs)



notes

Then each column in the CCM is a set of weighted links from that specific course (i.e., column) to the various concepts

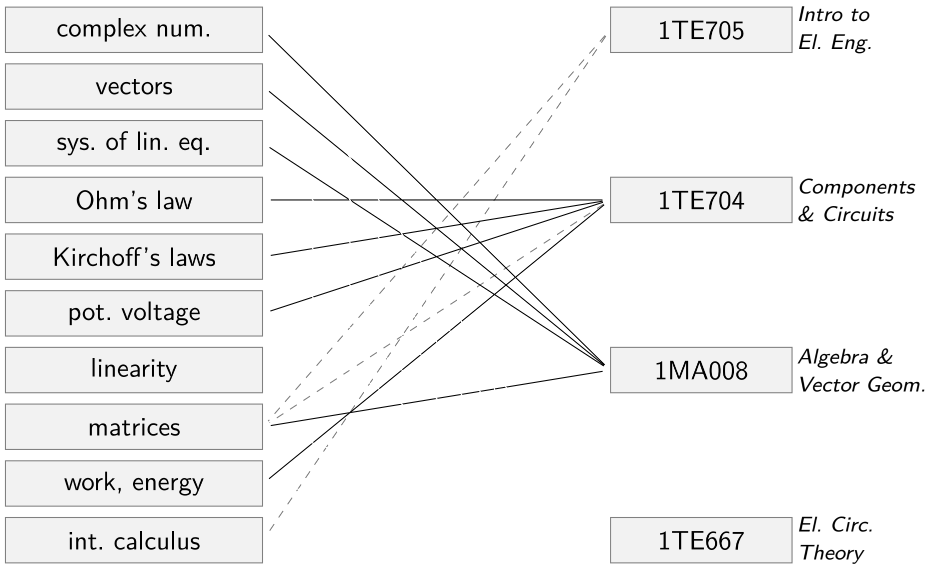
From Courses-Concepts-Matrices to Courses-Concepts-Graphs (CCGs)



notes

Then each column in the CCM is a set of weighted links from that specific course (i.e., column) to the various concepts

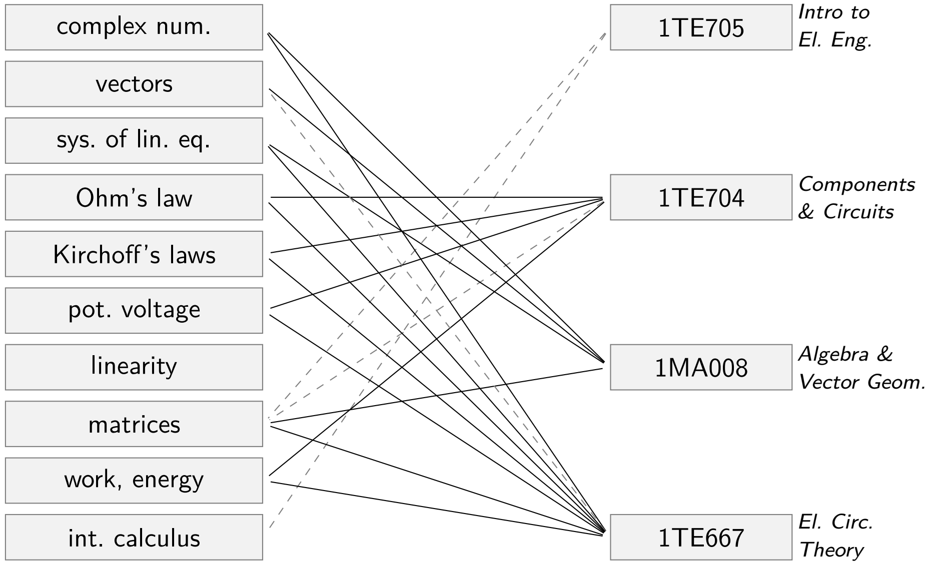
From Courses-Concepts-Matrices to Courses-Concepts-Graphs (CCGs)



notes

Then each column in the CCM is a set of weighted links from that specific course (i.e., column) to the various concepts

From Courses-Concepts-Matrices to Courses-Concepts-Graphs (CCGs)



notes

Something we have recently started analyzing is how to use classical graph connectivity and nodes centrality indexes as proxies for program connectivity and courses/concepts centrality indications.

Work-in-progress: program analysis via graph connectivity & nodes centrality indexes investigations

notes

For example, in literature one may find a lot of different indexes.

Work-in-progress: program analysis via graph connectivity & nodes centrality indexes investigations

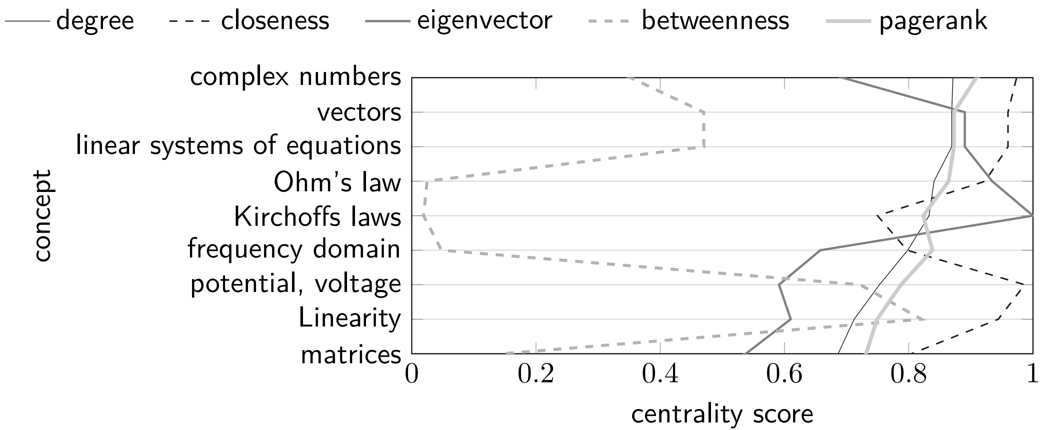
Example: *closeness centrality* := average length of the shortest path between a specific node and all other nodes in the graph

notes

Given that we have been collecting information on the CCMs / CCGs for our programs, we have been able to calculate these classical indexes.

Work-in-progress: program analysis via graph connectivity & nodes centrality indexes investigations

Example: *closeness centrality* := average length of the shortest path between a specific node and all other nodes in the graph

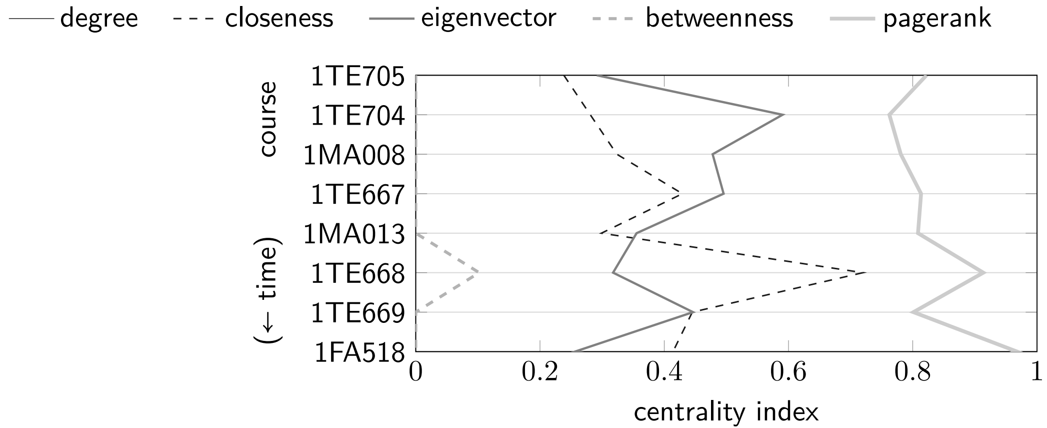


notes

Given that we have been collecting information on the CCMs / CCGs for our programs, we have been able to calculate these classical indexes.

Work-in-progress: program analysis via graph connectivity & nodes centrality indexes investigations

Example: *closeness centrality* := average length of the shortest path between a specific node and all other nodes in the graph



notes

The (research) issue is though: what do these numbers indicate? And how can we use them?

Work-in-progress: program analysis via graph connectivity & nodes centrality indexes investigations

Example: *closeness centrality* := average length of the shortest path between a specific node and all other nodes in the graph

What do these numbers indicate?
How can they be used to improve the programs?

notes

We have some preliminary results that have been just submitted to the European Control Conference 2019. We send the interested reader back to the corresponding pre-print in the authors' webpages.

Work-in-progress: program analysis via graph connectivity & nodes centrality indexes investigations

Example: *closeness centrality* := average length of the shortest path between a specific node and all other nodes in the graph

What do these numbers indicate?
How can they be used to improve the programs?

Preliminary results: *Courses-Concepts-Graphs as a Tool to Measure the Importance of Concepts in University Programmes*, submitted to *ECC 2019*

notes

It is important to realize that the information collected in the CCMs is quite minimalistic. This has the benefit of not requiring too much compilation work for teachers, but as a draw back it does not provide too many insights on the program structure. The question is then: how shall additional information be collected?

Potential extensions

(note: no data is yet available about these alternative representations)

Intuition: CCMs collect only first-order details about program structures.
How can we get second-order ones?

notes

The extensions that we identified so far are the two listed here.

Potential extensions

(note: no data is yet available about these alternative representations)

Intuition: CCMs collect only first-order details about program structures.
How can we get second-order ones?

Alternatives identified up so far:

- Directed CCMs / CCGs
- CC-Submatrices / CC-Subgraphs

notes

A simple extension may be to collect for each course not only what it teaches in terms of concepts, but also what it requires in terms of concepts, along with the importance of both the taught and required things.

Directed CCM

Example

	Course 1		Course 2		Course 3	
	teaches	requires	teaches	requires	teaches	requires
Concept A		1		2		1
Concept B	1			1		2
Concept C	2		1			2
Concept D			2		2	
Concept E					2	

notes

This would indeed enable to have a directed-graph representation of the program, and this is intuitively giving more information and possibilities to detect issues at a program level since it is more connected to the intuition of seeing a program as an information flow. For example, one may use refined graph connectivity indexes that may highlight how some concepts are key for the structure of the information flow. This could then be useful to re-design some courses so that they take more care about these concepts to be learned more deeply.

Directed CCG

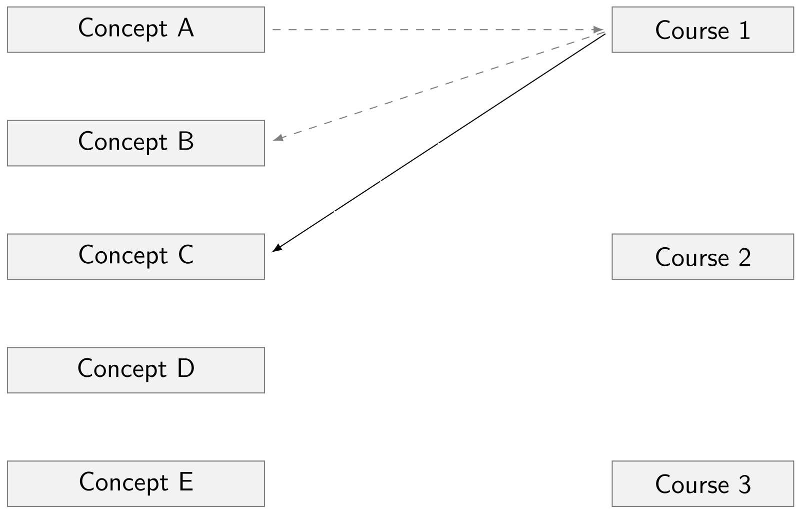
Example

notes

This would indeed enable to have a directed-graph representation of the program, and this is intuitively giving more information and possibilities to detect issues at a program level since it is more connected to the intuition of seeing a program as an information flow. For example, one may use refined graph connectivity indexes that may highlight how some concepts are key for the structure of the information flow. This could then be useful to re-design some courses so that they take more care about these concepts to be learned more deeply.

Directed CCG

Example

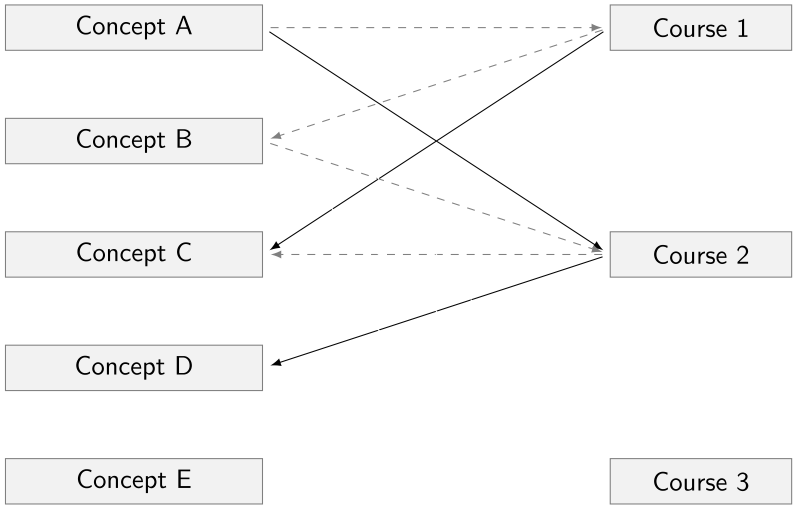


notes

This would indeed enable to have a directed-graph representation of the program, and this is intuitively giving more information and possibilities to detect issues at a program level since it is more connected to the intuition of seeing a program as an information flow. For example, one may use refined graph connectivity indexes that may highlight how some concepts are key for the structure of the information flow. This could then be useful to re-design some courses so that they take more care about these concepts to be learned more deeply.

Directed CCG

Example

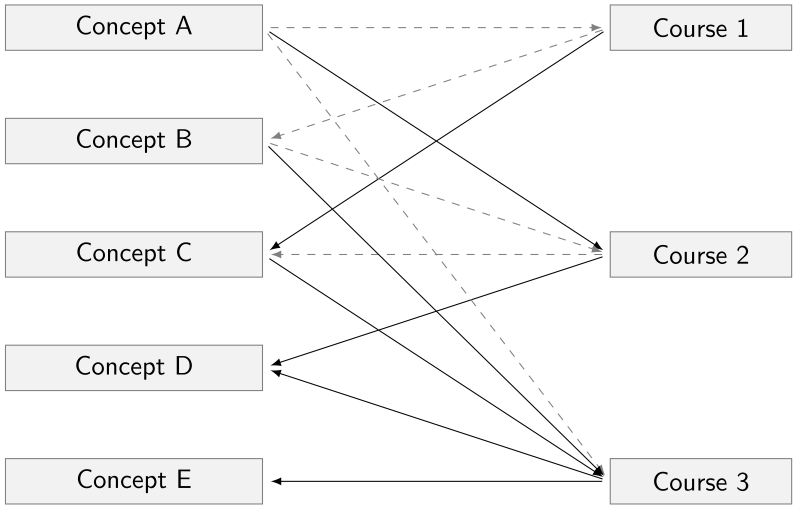


notes

This would indeed enable to have a directed-graph representation of the program, and this is intuitively giving more information and possibilities to detect issues at a program level since it is more connected to the intuition of seeing a program as an information flow. For example, one may use refined graph connectivity indexes that may highlight how some concepts are key for the structure of the information flow. This could then be useful to re-design some courses so that they take more care about these concepts to be learned more deeply.

Directed CCG

Example



notes

A further extension is to consider to make a CCM for each individual course. Here a strategy may be to try to connect the flow of information *within* the course by saying which concepts are needed to be known to master each one of the taught concepts. It would also be useful to register the importance of each taught concept (in terms of “teaching time” or other metrics) within the specific course. Note that this requires to fill up much more information than before, but would also give a complete description of the information flow within the program, and thus enable deeper analysis and corrective actions through opportune combination of the various CC-subgraphs relative to the various courses.

Courses-concepts Submatrices (CCSM)

Example

		course prerequisites			taught concepts		
	teaching time	vector spaces	matrix-vector multiplication	fund. th. of lin. algebra	eigenvalues & eigenvectors	characteristic polynomials	Jordan forms
eigenvalues & eigenvectors	45%	2	2				
characteristic polynomials	20%			2	1		
Jordan forms	35%		1		2	2	

notes

There is a problem, though, that we are starting to study now. How can we decide among which tool to use at a program level? Every type of CCM / CCG comes with different levels of information details / filling requirements. Teachers tend to be busy, so asking too much may lead to people refusing to use these tools. How to balance costs and benefits requires estimating both of them, and this is something we are studying now.

Big issue

what are the trade-offs
"efforts for collecting"
vs.
"usefulness of the information"?

notes

Thus our next steps in this project are the following ones.

Next steps

- investigate the trade-offs *"efforts for collecting"* vs. *"usefulness of the information"* for the various types of CCMs / CCGs
- extend the current set of questions
- involve more teachers and other organizations
- continue the full-scale tests
- develop data-visualisation and data-analysis tools

notes

And this is how people has been perceiving our project. The paper for this conference is available at the following link: `http://staff.www.ltu.se/~damvar/Publications/Fjallstrom%20et%20al.%20-%202018%20-%20Developing%20Concept%20Inventory%20Tests%20for%20Electrical%20Engineering%20(CITE)%20extractable%20information,%20early%20results,%20and%20learned%20lessons.pdf`

Bonus slide: summary of the opinions from the stakeholders

- the project is perceived as useful for everybody (students, teachers and boards)
- the tests are preferred online, at the beginning or before a learning period, 1 or 2 hours per test, 2 or 4 times per year
- it is very important for students is to see their progressions (much less to compare themselves with others)
- students slightly appreciate using reports for future job applications
- students do not like at all having the test count as an academic record