BUDAPESTI CORVINUS EGYETEM

CREATING A CONCEPT IMPORTANCE MEASURE FOR DOMAIN KNOWLEDGE IN THE CONTEXT OF LEARNING

DOCTORAL DISSERTATION

Supervisor: Réka Vas, Ph.D.

Christian Weber

Budapest, 2017

Christian Weber

Creating a Concept Importance Measure for Domain Knowledge in the Context of Learning

Department of Information Systems

Supervisor: Réka Vas, Ph.D.

Budapesti Corvinus Egyetem

Doctoral School of Business Informatics

Creating a Concept Importance Measure for Domain Knowledge in the Context of Learning

Doctoral Dissertation

Christian Weber

Supervisor: Réka Vas, Ph.D.

Budapest, 2017

Table of Content

Illustration Index

Table Index

Acknowledgements

I would like to express my special gratitude and appreciation to my supervisor and mentor Dr. Réka Vas. In all stages of the research, in our collaboration and as a lead on the path to master scientific imagination and the excavation called thesis, I treasured your continuous help and support, your steady readiness to approach new ideas while keeping track of the greater whole and your insights and sharp rational eye on details and on their implications alike. Scientific and personal development are two paths following one direction, being never one but being never apart, and I am grateful to be able to learn from a mentor who walked a good share on both paths and were ready to share it. Thank you for being ahead while holding out your hand. Without, this work would be a complete different one and so would be the researcher.

Further I would like to express my special gratitude to Dr. András Gábor, who were a mentor and supervisor for many national and international students throughout his career and from whom I had the great honour to learn within my years in Budapest. His visible and more so often invisible, dedicated, careful, yet strong hand not only shaped and built the environment for this work – he was also never tired of giving a vision in times of uncertainty, giving trust in times of failure and sharing a life of experience and achievement when our hands were just plain empty. It takes much to put the lost on a road, it takes more to teach them how to walk, it costs even more to make them use their heads to look beyond the walk, but it takes all to take that one step back to make the lost learn from their mistakes without letting them get lost again.

I would like to thank my mother, my father and my brother who were always there to support me and walk with me in both – the good and the not so good times. I am eternally grateful for them to believe me and believe in me when I decided to pursue a Ph.D. I would like to thank Dr Madjid Fathi for picking me from the floor in front of a seminar room to believe in me and for being both kind and diligent enough to call up my mother to make her believe that I will make it, to then turn down the phone to make me believe.

Friends. I am more than eternally grateful to you. You that you dared to walk with me. You that share the dark when the lights are out of reach – and let me do the same for you. The ones who didn't stopped with opening a door but gave that tiny push to make it through. The ones who smiled for me without anyone watching. The ones who talked to me when I was deaf. The ones who dare to be wrong and right with me. The science punks. The we believers. The ones who are tempted to believe that: Rain falls gently on everyone and anyone alike. And all the others.

This work was created within the Eduworks Marie Curie Initial Training Network Project (PITN-GA-2013-608311) of the European Commission's 7th Framework Program and I acknowledge their financial but also their scientific support.

1. Introduction and Problem Description

Exiting the financial crisis, the European Union has set a clear path in 2010 for a Union wide change, bound to the Europe 2020 Strategy. Europe should deliver a smart, sustainable, and inclusive growth. A growth which can't be tackled without an educated and capable workforce. As such, two of the five main pillars of the Europe 2020 Strategy to guide targets on a national level, are employment and education.

Following the lead of Europe 2020, the Smart Specialisation Strategy (s3) strives to connect and assist EU countries and regions to develop, implement and review Research and Innovation Strategies on a local level. Smart Specialisation centres around fostering micro, macro and meso level competitive advantages and potentials for excellence, with a strong focus on entrepreneurship and innovation, and recommendations for changes to acquire them in a new knowledge economy.

At the same time, there is an ongoing pressure of globalization and internationalization, blurring the borders between markets and previously segregated job profiles and their requirements. More and more economies are becoming service economies, increasingly incorporating ICT technologies to empower services, which require changes in the profiles and in the variety of workers. Finally, the civil society becomes enclosed and infused with technology, coming nearer to a network society where communities are increasingly virtual and interconnected.

In this frame, change and the ability to cope with changes, becomes an essential trait for the modern worker. **The problem is that the traditional education can't cope with this transformation and in this way, fails to meet the urgency of adaptation.** Well trained workers in a smart economy need a continuous education – coming in short cycles of training and practice, which takes place within their job environment.

The focus is twofold: improving the present work, and developing the ability to learn for future work requirements. The need for job knowledge[1] is here a union of the educational supply and economical requirements and leads to the question on how to keep the job knowledge up-to-date and how to effectively train on the job.

This requires such education that is fit to purpose and time - regarding the

[1] Job knowledge is here seen as the knowledge needed to perform specific job roles and includes conscious expertise and tacit knowledge.

selection and setup of educational material - to finally harness the real potential for a smart growth and a participation in the knowledge economy.

1.1. Background and Focus of the Thesis

Modern workers have to cope with many changes and new flexible requirements in their working environment. Projects may require new skills, co-workers may leave the company – leading to a shift in responsibilities –, or new labour market opportunities motivate to acquire different skill sets. Coping with changes and satisfying the need for personal improvement, creates a high need to know the extent of the personal education, e.g. through continuous self-assessment to monitor the personal progress.

Over the last years, semantic technologies emerged as a new approach to see knowledge as a structured and connected asset, rather than an isolated one. This also had a great effect on the vision about and the handling of organisational knowledge. Based on the technologies, emerging from the last iterations of the web – first developing to an application driven web (web 2.0) and then to a semantic web (web 3.0) – new improved ways to store, access and update knowledge are developed, together with new, proven solutions and best practices.

Information is now interconnected, and as such offers new possibilities to learn and educate what a person needs to know to perform well in different learning situations, as formal education, during their job or – one step further – for their future job and future education. With the blending of learning and semantic web technologies, a new generation of systems emerged, that makes use of interconnected information and semantically enriched knowledge structures to help people to learn what they need to know.

So, to cope with changes and to overcome the limitations of a static, formal education, new educational systems - making use of connected and structured knowledge - could fill the gap and connect *what a person knows* and *what is needed to fit to new requirements,* in situations like: applying for a new job, pursuing a new education or adapting to changes within the job roles. But, as access to information becomes more flexible and the information delivered becomes more extensive and connected, the selection of the right information at the right time is also becoming more and more important.

Single, isolated pieces of information can be turned into a network of information, presenting - when enhanced by the power of semantics - a knowledge structure. This structure – by disclosing the context of the required knowledge - enables a more flexible way of learning. While people are able to judge what are the relevant concepts [2] in such domains that they know well, this becomes a considerably harder task in new domains of knowledge. In situations where a person explores new knowledge the question definitely arises: “*What to learn first?”* So far, an approach to distinguish the relevance of concepts in a given domain to enable learning is missing.

This thesis will explore how to use the information about the structure and the semantic of a knowledge structure in a field of learning, to create a measure for describing the importance of the single concepts. Furthermore, this thesis will examine how an implementation of this new measure in an adaptive system of technology enhanced learning can be realized.

1.2. Detailed Research Questions

- **Research question 1:** How can **the semantic model of a learning domain** be utilised to identify **which knowledge area**(s) **is** (are) **of high importance for learning** in comparison to other knowledge (concepts) within the model of the domain? (*Methodology: [Modelling][Experiment]*)

A semantic model is a conceptual model and an abstraction of a specific part of the real world. It includes additionally semantic information to describe its individuals and the “how” of their relations. In this regard, the model explains in a formal way the semantic of its instances and relations. Many different semantic models of different domains of life exist. In the context of learning, a semantic model can model the learner or the domain of learning.

The semantic model of a domain (domain ontology) – used within this thesis – includes the concepts or knowledge which are needed in the domain. The semantic model explicitly models different relations between concepts like “requirement”, “sub-knowledge” and more specialized relations. Furthermore, the concepts can be different

[2] Within this thesis “concept” and “knowledge” are partially used interchangeable, yet they are applied with a different context: concept is used when addressing the elements within the structure to store modelled knowledge, while knowledge is used in the context of learning and modelling the knowledge to learn.

and belong to different types or classes of individuals, like knowledge area, example or basic concept. The relations and concepts and their formalized types, enable a structured, semantic rich domain model, which in turn enables to structure a learning process.

Throughout the process of learning a learner learns and masters concepts one after another till the domain is mastered. Now the question is – is it possible to use the information about the structure (***How are concepts connected?***) and the information about the semantic of the domain (***What is connected?***) in a systematic way to create a measure to rate which concepts should be mastered first in a domain and therefore should be assessed first and are "more important" to enable a better (or faster, more sustainable or more specialized) learning?

"Importance" in the context of learning or in a colloquial sense can have different meanings as "significant" and "meaningful". In the frame of this thesis and in the context of supporting the assessment of concepts for learning, the importance is interpreted in two dimensions: a quantitative (***How well-connected are concepts in a network?***) and qualitative (***How needed and central is a single concept for connected concepts, based on its underlying semantic?***). Based on these considerations, a working definition of concept importance is "***the degree to which a concept is connected, and central in terms of its semantic, to other concepts in a domain network of concepts***".

- **Research question 2:** How can a **measure, quantifying the importance of concepts** in a semantic model **be utilized, integrated and implemented** in an online assessment solution? (*Methodology: [Build] [Experiment]*)

Measuring or "deriving" the importance of concepts by applying the domain model, enables to select which concepts are valuable to learn first and to explore the given domain. Learning is an incremental process with phases of learning and reflection. To cope with the complexity of domains to learn, a system to support the learning process has to do a selection of concepts should be learned first. One possible solution for the selection, is to assess the state of knowledge of a learner on a selected domain. The gained information can then be used to tailor learning-paths and provide personalized learning recommendations. In this regard, knowing which concepts are important for learning will give an indication what to assess first – to detect the current training need, which then supports to tailor the learning and reflection to focus on concepts which are beneficial to be learned first.

In online, technology enhanced learning solutions, learning can be translated into a continuous cycle of online assessment and tailoring of learning. Assuming a well-defined, sound measure which defines the importance of concepts within a semantic model of the domain – *that is the outcome of Research Question 1* – the next step is to define how the measure can be utilized in supporting and framing the continuous assessment of the training need. Furthermore, following the utilization of the concept importance, how can the measure be implemented into an assessment algorithm for a specific technology enhanced solution for assessment and learning?

1.3. Research Context

In the current research the focus is on work-force education and life-long learning. In this regard, potential modelling approaches of concept importance have to be investigated by taking into consideration both: **learning theories** and an **organizational context**. In contrast, grasping the personal education and development potential, requires the testing of the current state of education through assessment, while adapting to the personal performance and learning potentials, as addressed by **adaptive testing** and **user modelling**. Furthermore, **network analysis** is becoming a strong driver to gain insights about networks of people, information and knowledge and yields the potential to discover deeper the relationship between single knowledge elements.

To create a new measure of concept importance, specifically methodologies from the field of user modelling will be investigated. User modelling captures how to describe users and their knowledge and background through specific representations. It is used in technology enhanced solutions for learning for: content adaptation and representation adaptation, and for deriving learning recommendations. Especially for applications in the field of education, user modelling takes the perspective of describing the user based on the structure of the knowledge domain to learn. In this regard, it provides a suitable view and fundament for a new concept importance measure in the context of learning and education. Especially semantic enhanced domain models will be considered, as they further specify a semantic dimension to the learning domain, which can be exploited to model a concept importance measure.

Accordingly, the new concept importance measure will be created, taking into consideration the principles and methods of learning theories, organisational learning and networks analysis in the context and application of adaptive testing. Figure 1 visualizes the connections of topics which will be discussed in detail in the literature review. As visualized in Figure 1, this thesis will use User Modelling as a starting point to develop the initial ideas to create the concept importance measure. To frame and shape the ideas, the fields of Organisational Learning, Learning Theory and Network Analysis will be considered – background-wise and regarding common solutions and further implications. These areas are considered to ultimately enable to utilize the concept importance measure for Adaptive Testing and Learning Recommendation and – in a labour market context potentially to enable Pre-selection with a system, using the measure.

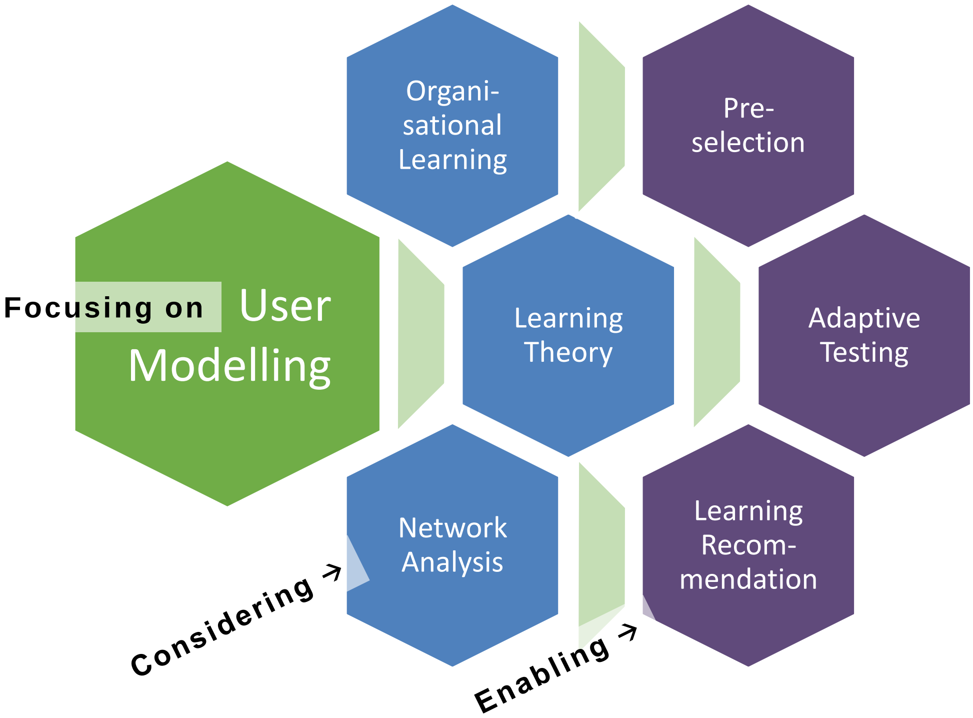


Figure 1: Supporting focus, considerations, and applications of the planed research.

2. Training and Education in the Organisational Context

With the on-going integration of new, computerized devices in our daily lives and especially the development of new technologies accelerating with the push of the Internet, different, previously manual tasks have re-emerged as new technology enhanced versions. Training and education are especially profiting here from the new opportunities in integration and developments in the related fields of science.

Assessing human education, abilities and various aspects of performance always comes with the need for a strong set of methodologies from the field of education. They are supported by enhancements in the field of neuroscience, where more and more processes and relations of human reasoning are explored, enabling a better understanding of learning and integrating insights from the psychology of learning. At the same time, in the organisational context, the impact of a worker’s knowledge is recognized as a central factor in the economic success of human capital, resulting in new disciplines on how to improve and assess individual knowledge.

2.1. Continuous Training

Kuckulenz (Kuckulenz, 2007) introduces continuous training under the aspect of human capital. Human capital goes back to the seminal work of Becker (Becker 1964). Blundell, Dearden, Meghir and Sianesi (Blundell et al., 1999) define the three main components of human capital as:

- *early ability (acquired or innate)*
- *qualifications and knowledge acquired through formal education*
- *and skills, competencies and expertise acquired through training on the job; (or acquired in the migration to a new job)*

Learning, in the sense of human capital, is an investment in personal education, temporarily giving up a part of the current income in favour of an expected higher later income on a personal level and a higher expected productivity of the workers on the organizational level. Blundell, Dearden, Meghir and Sianesi (Blundell et al., 1999) address training, based on the use in empirical studies, as “generally defined in terms of courses designed to help individuals develop skills that might be of use in their job" and exclude formal training in the form of school and post-school education. Furthermore, they make evident a strong correlation between the performance in the two earlier components and the likelihood to engage in additional job related training and the performance on the same training.

Even though the general definition of training excludes regular education as formal training, embarking upon higher education implies a job relevant and job targeting decision and, while formalized, requires self-motivated learning behaviour. In this regard this chapter will address training and education synonymously, implying a process of for the job and on the job training and focus on these respects. Furthermore, change is becoming a constant in daily life, following the pace and requirements of the markets, variations in labour and formal education undergoing constant change. Testing is the important factor to outline the current education and job- and work-horizon, supporting the knowledge worker to self-assess current strengths, potentials and the bottlenecks towards a next stage of education.

For a situation of training and education, computer aided, adaptive tests are one powerful enabler to connect the knowledge within a field of education. Current solutions to support a computer aided test preparation and test execution, neglect the fact that education and testing always takes place within a context, evolving around a problem context, given by the organisation, in line with the performance of the assessed individual.

2.2. Formal, Non-formal and Informal Learning in the Organisational Context

Learning and especially organisational learning is a fractured area with different views, strengthening a variety of differing fundamentals and insights. As such the definitions of types of learning are still an object of discussion. With the target of creating a common understanding, different organisations over time promoted different sets of definitions, while matching in the three major types of learning: formal, non-formal and informal learning. Especially for use in policy and decision making, within and across countries, with the OECD, UNESCO and CEDEFOP, three grand multinational organisations proposed definitions for the three types of learning in line with the development in the literature of organisational learning, as shown below in Table 1.

Table 1: Types of learning for policy making.

	OECD (OECD, 2013)	**UNESCO** (UNESCO, n.d.)**, based on** (Commission, 2001)	**CEDEFOP** (CEDEFOP, 2011)
Formal Learning	Formal learning is always organised and structured, and has learning objectives. From the learner's standpoint, it is always intentional: i.e. the learner's explicit objective is to gain knowledge, skills and/or competences.	Formal learning occurs as a result of experiences in an education or training institution, with structured learning objectives, learning time and support which leads to certification. Formal learning is intentional from the learner's perspective.	Learning that occurs in an organised and structured environment (in an education or training institution or on the job) and is explicitly designated as learning (in terms of objectives, time or resources). Formal learning is intentional from the learner's point of view. It typically leads to validation and certification.
Non-Formal Learning	Mid-way between the first two, non-formal learning is the concept on which there is the least consensus, which is not to say that there is consensus on the other two, simply that the wide variety of approaches in this case makes consensus even more difficult. Nevertheless, for the majority of authors, it seems clear that non-formal learning is rather organised and can have learning objectives.	Non-formal learning is not provided by an education or training institution and typically does not lead to certification. It is, however, structured (in terms of learning objectives, learning time or learning support). Non-formal learning is intentional from the learner's perspective (Werquin, 2007).	Learning which is embedded in planned activities not explicitly designated as learning (in terms of learning objectives, learning time or learning support). Non-formal learning is intentional from the learner's point of view.
Informal Learning	Informal learning is never organised, has no set objective in terms of learning outcomes and is never intentional from the learner's standpoint. Often it is referred to as learning by experience or just as experience.	Informal learning results from daily life activities related to work, family or leisure. It is not structured (in terms of learning objectives, learning time or learning support) and typically does not lead to certification. Informal learning may be intentional but in most cases it is non-intentional (or 'incidental'/random).	Learning resulting from daily activities related to work, family or leisure. It is not organised or structured in terms of objectives, time or learning support. Informal learning is in most cases unintentional from the learner's perspective.

All three definition streams share a common core with a slightly different emphasis, as seen in Table 2, based on the UNESCO definition and taken from (Werquin, 2007). An important factor is the use and reasoning on the learner's intention of learning. Informal learning differentiate here that a situation may have an intention as the acting person has a goal but the main intention is not "learning" but rather learning is an inevitable, unconscious consequence.

Table 2: Differences of learning frames (Werquin, 2007).

	Organised	**Learning Objective**	**Intentional**	**Duration**	**Leads to a Qualification**
Formal Learning	Yes	Yes	Yes	Rather long and/or full-time	Yes[3]
Non-formal Learning	Yes or No	Yes or No	Yes or No	Rather short, or part-time	No[4]
Informal Learning	No	No	No	NA	No

Generalising and especially from an organisational view, formal learning connects to school and university education, offering certified degrees from accredited studies, while non-formal learning means organisational training activities which are "structured" and "organised" in their implementation but not formal in terms of the mediated education. This is independent of the potential that organisations (internal or external) can certify trained skills, with or without additional evaluation.

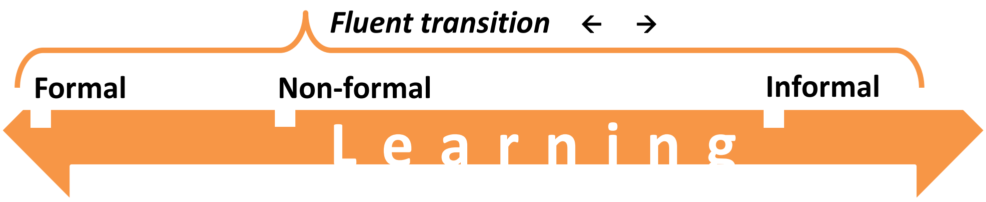


Figure 2: Fluent transition between types of learning.

As the understanding of formal, non-formal and informal education is strongly connected to the cultural background and the spatial environment, the line between the former cannot be drawn in an absolute fashion and adds to the on-going discussion within the field. According with the variety of situations of learning, rather than being an absolute categorization it is a transient scale, as shown in Figure 2, and is always a mix, depending on the specific situation.

3 "Almost always"
4 "Usually no"

Easterby-Smith, Crossnan and Nicolini (Easterby-Smith et al., 2000) capture a part of the areas of on-going discussion under the term of operational learning. They name the starting fields "units or levels of analysis" addressing the question whether organisational learning is the sum of individual learning or needs to be approached differently. Here Garratt argues that a small group of people could have a major influence on strategic decisions within an organisation and as such a small group of senior managers could give a good approximation of the thinking of an organisation (Garratt, 1987).

Fiol and Lyles suggest the organisational structure and procedures directly affects the individual learning (Fiol and Lyles, 1985), while Shrivastava collects that the knowledge of individual decision makers creates policies and procedures for organisations which are then embedded into the organisational structure and the organisational socio-cultural norms (Shrivastava, 1983). Shrivastava summarizes further four main aspects of organisational learning: organisational learning as *"adaptation"*, *"assumption sharing"*, *"developing knowledge of action outcome relationships"* and *"institutionalized experience"*, highlighting in his summary that organisational learning is partially motivated to respond to change.

People learn in a variety of different environments and situations and approach learning intentionally and unintentionally, following conscious and unconscious goals alike. Non-formal training in organisations usually takes place in a fit between the organisational goals and the interests of the workers, e.g. within the framework of seminars and training systems. Non-formal learning approaches are opening the potential for an effective and on-going education within organisations and make use of previous formal education while the teaching offered is driven by the organisational interest, also embedding as such the organisational culture and benefiting from the job relevance for the workers, and overall facilitating further implicit informal-learning processes.

Vaughan and Cameron (Vaughan et al., 2009)(p. 2) gives a set of reasons why assessment in the workplace is yet under-explored in the current literature, placing emphasis on the impact of informal learning, including one emphasis on the understanding of learning:

> *"There are several, interwoven reasons why there is very little literature dealing directly with workplace assessment. Firstly, workplace assessment is closely related to forms of learning that are not recognised or understood as learning, which means that such learning is less likely to be assessed."*

A technology enhanced testing and learning solution could act here as a great benefactor and yield a high potential for more additional informal learning outcomes. It could provide the impact of organisational learning in a timely and spatially independent fashion, based on the need of the organisation and accessible in situations where workers are searching for further education.

Furthermore, a technology enhanced approach can foster the hand-shake between organisational and the workers' learning interests through basing the learning and testing on the processes of the organisations – which connect the structural aspects of the organisation with the job roles and job role competencies of the individual workers. To assess and motivate the learning of these competencies across a wide range of individuals a degree of adaptivity is mandatory for a working and sustainable assessment.

2.3. Short Summary: Training and Education

Table 3: Summary "Training and Education in the Organisational Context".

Aspect	Description
Continuous training need	Fast changing markets and job profiles require flexible, educated workers and a continues training. Training in non-formal education and higher education requires foremost a self-motivated flexible learning behaviour.
Lack of assessment in non-formal/informal training	The assessment in the workplace is yet under-explored in the literature. Technology enhanced learning can here fill the gap, in a timely and spatially independent fashion.
Need for flexible, adaptive solutions	To enable an assessment and learning across different individual and organizational needs, flexible and adaptive systems for technology enhanced learning are needed.

3. Concepts of Adaptive Testing

In knowledge-driven professions adapting to new fields of knowledge and expertise is a daily necessity, which also covers cases of organisation-internal change. Maintaining and extending here the right set of knowledge, requires personalized development with the potential to follow multiple paths of education in dependency of the personal context. To fully unlock the personal potential in a resourceful way, education has to mediate between the need to educate and the ability to learn. This requires an open and adaptive approach to testing, maintaining a balance between the time to assess the state of the learner and his/her time to learn detected gaps in education

3.1. Assessment and Self-Assessment in the Context of Education

Structured education is still frequently packed into static built curricula with one main path created to master the contents. To enrich this main road of education and to cope and connect to individual learning situations, students have to identify the gap between their own knowledge and the expectations of the specific curriculum. So as students evolve throughout the fields of higher education, they are steadily encountering situations where they have to evaluate their existing knowledge and reason on ways to expand their knowledge. Further students tend to misjudge their personal skills, with their self-prediction being often substantive and systematically flawed (Dunning et al., 2004). A systematic and objective solution for self-assessment could help here to support self-prediction of personal proficiency and to prevent a wrong or biased self-evaluation.

Self-assessment provides feedback to the students about their personal performance in relation to the target of education, which is prolonged by more learning and repeated self-assessment. This feedback cycle is usually integrated into other technology enhanced learning solutions in the frame of a learning management system (LMS) also known as virtual learning environment (VLE), which combines educational services with channels of feedback (Lonn and Teasley, 2009). Or the feedback could be mined, even from high volume, high throughput educational systems like MOOCS (Tabaa and Medouri, 2013). Seidel, Perencevitch and Klett, summarizes here the requirements and distances between the tutor and the student in technology enhanced environments of learning (Seidel et al., 2005).

A survey of the market by the OECD in 2005 indicated that within the tertiary education, systems often focus mainly on management activities (OECD, 2005), neglecting the potential for a full learning support, especially for individual learning. Indeed one of the best ways of education is a direct tutoring of short reoccurring learning-paths, as shown by Bloom (Bloom, 1984) for classic tutoring and Fletcher (Fletcher, 2003) for computer aided systems. Further, as Jonassen shows in his thesis about constructivism, knowledge to assess for any scope of learning is interrelated and associated with past experiences (Jonassen, 2009).

Following Jonassen, knowledge in education could be split into nine types across three categories to capture the human cognitive behaviour. In his discussion, eight out of nine knowledge types underline that knowledge in the scope of learning is interrelated and strongly associated with previous experiences. As such, a supporting solution for self-assessment should grasp and formalize the knowledge to assess in the context of related knowledge.

An important factor in learning is the distance between the expectation of the tutor and the learning performance of the student. Here a short cycle of repeated assessment and learning is a major factor for a better personal learning performance (Roediger, 2008). This aspect directly benefits from the focused concentration on knowledge-areas as the main exchange concept between students and tutors. As even further the close connections between learners and educators via direct tutoring is one major enabler for computer aided systems (Fletcher, 2003), each step towards a more direct interaction through focused concepts is an additional supporter.

In technology enhanced assessment and especially in self-assessment, with its stronger correlation to personal education, the educated knowledge has to be seen in the environment of past experiences and related knowledge to improve situations of learning.

3.2. From Classic to Adaptive Testing

For an adaptive test, in contrast to a classical linear test, the number of test items and the order of questions is only determined during the test itself. The goal is to determine the knowledge level of the test taker as precisely as possible with a number of questions as low as possible. It does not require deep analysis to realize that in contrast, linear tests have always been constructed to meet the requirements of test-givers.

Selecting the adequate examination method or setting up a good test is far from being an easy task. A test that is much too easy or much too difficult is a waste of time. If the test is too easy for the candidate, it is likely to invite unwanted candidate behaviour such as thoughtless mistakes. On the other hand, questions that are much too hard also produce generally unreliable test results, as candidates will give up early and cease to seriously attempt to answer the questions, resorting to guessing, response patterns and other forms of unwanted behaviour.

According to traditional testing methods, all candidates should get absolutely the same questions. This way the results and the performance of candidates can be easily and clearly compared. Adaptive testing breaks away with this approach. Every candidate may receive entirely different question and not even the order of questions is defined in advance. This means, answers of previous questions determine which questions will be asked in the following steps.

3.3. Adaptive Knowledge Testing

An adaptive methodology was first applied in psychology and Alfred Binet worked out the first form of adaptive testing in 1904 (Simon and Binet, 1904). Alfred Binet's goal was to develop intelligence tests which were aimed at diagnosing the individual. By moving away from the target of assessing an assumed homogenous group he could eliminate the issue of fairness, providing the same test framework but assessing it on the basis of personal performance. He then customized the test based on an individual rank, by ordering the items (questions) according to their assumed difficulty. He would then start testing the candidate through continuously estimating the probable level of the candidate's ability, based on the difficulty of the asked items.

The selection of items then represented a subset of the overall set of available question items which he believed to meet the detected level of the candidate. If questions were answered correctly Binet would assess successively harder item subsets and would stop as the candidate failed frequently. Vice versa, if the candidate failed questions, then Binet administered successively easier item subsets, finishing if the candidate succeeded frequently.

For an adaptive test, in contrast to a classical linear test, the number of test questions and their order is determined within the process of testing. The target is to find a trade-off between determining the knowledge level of the test candidate as precisely as possible while administering questions in numbers as low as possible. An adaptive tests favours detecting the capabilities of the test candidates in contrast to linear tests, which are constructed to reflect and meet the requirements of test-givers.

A challenge for this type of adaptive tests is, concurrent to linear tests, to set up a test or set of test items which assess an appropriate level of difficulty. Both an easy test and one that is too complex may fail. If a test is too easy for a candidate it may invite unwanted human behaviour such as careless mistakes or unneeded, yet wrong assumptions. Furthermore, questions and question sets which are much too hard to answer produce generally unreliable test results, influencing candidates to give up early and inviting frustration, leading to guessing, response patterns and additional unwanted and biasing behaviour. Within traditional testing methods, candidates should be assessed by the same set of questions, static in number and assessed completely. This renders the results and the performance of each candidate comparable and reliable for repetition

Binet's procedure has been improved and refined by many authors, like Lord (1980) (Lord, 2012) Henning (1987) (Henning, 1987) and Lewis and Sheehan (1990) (Sheehan and Lewis, 1992). In their approaches, items are stratified by their difficulty level, and subsets of items are formed at each level. The test starts with administering subsets of items and it goes on with moving the difficulty up or down in accordance with the success rate on each subset.

3.4. Computerised Adaptive Testing

Among the different approaches for adaptive testing, computerized adaptive testing is a well explored example on how to realize adaptivity and provides a solid researched ground with a strong link to the initial idea of adaptive testing.

Reckase developed the first computer aided adaptive testing in 1974 (Reckase, 1974), fusing the potentials of IT systems with adaptivity and a user centric strategy to overcome the limitations of linear testing. In case of computer adaptive testing (CAT), it is not the test-givers task anymore to select the items, define the order of further items and evaluate results, as the computer will execute all these tasks. The candidate gets administered questions automatically, with a difficulty level that is accordant to his estimated ability level, with which examination stress is reduced and a suitable level of motivation is assured. But the most important feature is that the time of testing is much shorter, as shown by Welch and Frick (Welch and Frick, 1993).

Linacre (Linacre, 2000) provides a detailed description of the theoretical background, major types and logic of operation in adaptive testing. Several methodological approaches have been developed for CAT which evolved from the principles of psychometric measurement. Nowadays CAT has reached a broad application from local education (Čekerevac Zoran and Petar, 2013) till mobile solutions on smart devices (Triantafillou et al., 2008). From the numerous variations, Item Response Theory (IRT) (Holland and Wainer, 1993) and Knowledge Space Theory (Falmagne et al., 1990) have spread the most.

The main characteristics of CAT – independently from the approach – are:

- The test can be taken at the time most convenient to the examinee; there is no need for mass or group-administered testing, thus saving physical space.
- As each test is tailored to an examinee and no two tests need be identical, which minimizes the possibility of copying.
- Questions are presented on a computer screen, one at a time.
- Once an examine keys in and confirms his answer, he is not able to change it.
- The examinee is not allowed to skip questions, nor is he allowed to return to a question which he has confirmed his answer to previously.
- The examinee must answer the current question in order to proceed onto the next.
- The selection of each question and the decision to stop the test are dynamically controlled by the answers of the examinee (Thissen and Mislevy, 1990).
- Providing the simplest version of adaptive testing, dichotomous items (multiple choice questions) are weighted along a linear scale and answers can only have two values a time (true or false).

Gershon (Gershon, 1992) suggests, that the first item, and perhaps all items, should as a tendency be a bit on the easier side, giving the candidate a feeling of accomplishment in a situation of challenge. If there is a criterion pass-fail level, then a good starting item has a difficulty slightly below that.

In case of tests consisting of polytomous items, answers can have more than two values. Typically, the test-taker would not be able to recognize any difference between a strictly dichotomous question and a polytomous question. The difference is in the scoring. False answers (distractors) may be more correct than others, and so are given greater scores (or credits). Naturally, the correct option is given the greatest score. At the same time, selecting questions for a test from a set of polytomous items requires more care since it is not possible to define the difficulty of all items with intermediate levels which indicate different ability levels.

Since polytomous items are more informative regarding the candidate performance than dichotomous items, polytomous CAT administrations usually comprise fewer items. At the same time writing polytomous items, and developing defensible scoring schemes for them, can be difficult.

A CAT-based system cannot operate adequately without an objective measurement model that enables the evaluation of results. Most of the computer-adaptive testing systems are based on the models of Rasch (Rasch, 1960) and Wright (Wright, 1988).

Almost all ability tests are based on the hypothesis that abilities can be ranked along one single dimension. But in the average, no test is exactly one-dimensional. If candidates are to be ranked relative to each other or relative to some criterion levels of performance, an approximation to unidimensionality has to be achieved. When a new item is entered to the test-bank it has to be decided how to determine the difficulty level of the new item and how to preserve the unidimensionality of the test. In most of the cases the estimation of the difficulty of new items could be reduced to a matter of maintaining a consistent stochastic ordering between the new and the existing items in the bank (Rasch, 1960). Instead of defining the real difficulty level, the difficulty level and ranking of the new item is defined by comparing it to old items in the bank. The primary goal is to maintain a consistent stochastic ordering. Its most spread tool is the Rasch model (Roskam and Jansen, 1984).

The essence of the dichotomous Rasch model is that it presents a simple relationship between the test-takers and the items. Each test-taker is characterized by an ability level expressed as a number along an infinite linear scale of the relevant ability. The local origin of the scale is chosen for convenience. The ability of test-taker n is identified as being Bn units from that local origin. Similarly, each item is characterized by a difficulty level that is expressed as a number along the infinite scale of the relevant ability. The difficulty of an item i is identified as being D_i units from the local origin of the ability scale.

Accordingly, the relationship between test-takers and items is expressed by the Rasch model (Thorndike and Linacre, 2000):

$$log_e(P_{ni}/P_{\mathrm{in}}) = B_n - D_i \tag{1}$$

Where P_{ni} is the probability that test-taker n succeeds on item i, and P_{in} is the probability of failure. Wright (Wright, 1988) suggested here in 1988 an algorithm which is easy to implement, and could be successfully employed at the end of each learning module to keep track of student progress.

The Item Response Theory (IRT) were introduced to CAT by Wainer and Mislevy (Thissen and Mislevy, 1990) and is a statistical framework in which examinees can be described by a set of ability scores that are linking the actual performance on test items and the examinees abilities. In this regard, it is extending the use of the Classical Test Theory (CTT), where characteristics of the test-taker and the test could not be separated.

For a classic, linear test structure, CTT and IRT could perform similar, when used for pre-testing test items, and offer the same quality, with CTT being more invariant than IRT methodologies (Fan, 1998; Lawson, 1991). Nevertheless, for CAT, where tailoring, adaptive test builds are the focus of interest, IRT-based methodologies are mandatory to enable the system to provide test items in an adaptive and personalizing manner (Stage, 2003).

Each item can be associated with one or more of the following parameters – the difficulty level, the discriminatory power and the guessing factor. The difficulty level describes how difficult and complex an item is, the discriminatory power explains how well the test item differentiates students of different proficiency level, while the guessing factor is the probability that a student can answer correctly by simply guessing. Before each test, the test bank must be balanced and checked that there are not over-represented or under-represented knowledge areas. Finally, the values of test item parameters must be defined. The progress of testing under IRT is given in Figure 3.

In line with the observation of a potential bias of questions, Vandenberg and Lance (Vandenberg and Lance, 2000) underlined, that a missing assessment of differences across assumptions may lead to an inaccurate reasoning. In the cultural context, this phenomenon is identified and formalized by Differential Item Functioning (DIF). In case a person with the same underlying ability has a different chance to answer a question correctly because of differences in the cultural background, this question is showing DIF.

Makransky and Glas (Makransky and Glas, 2013) modelled here a solution for Computerized Adaptive Testing, virtualising a question with DIF and attaching a context label to account it differently according to the cultural context, while still assessing the same question.

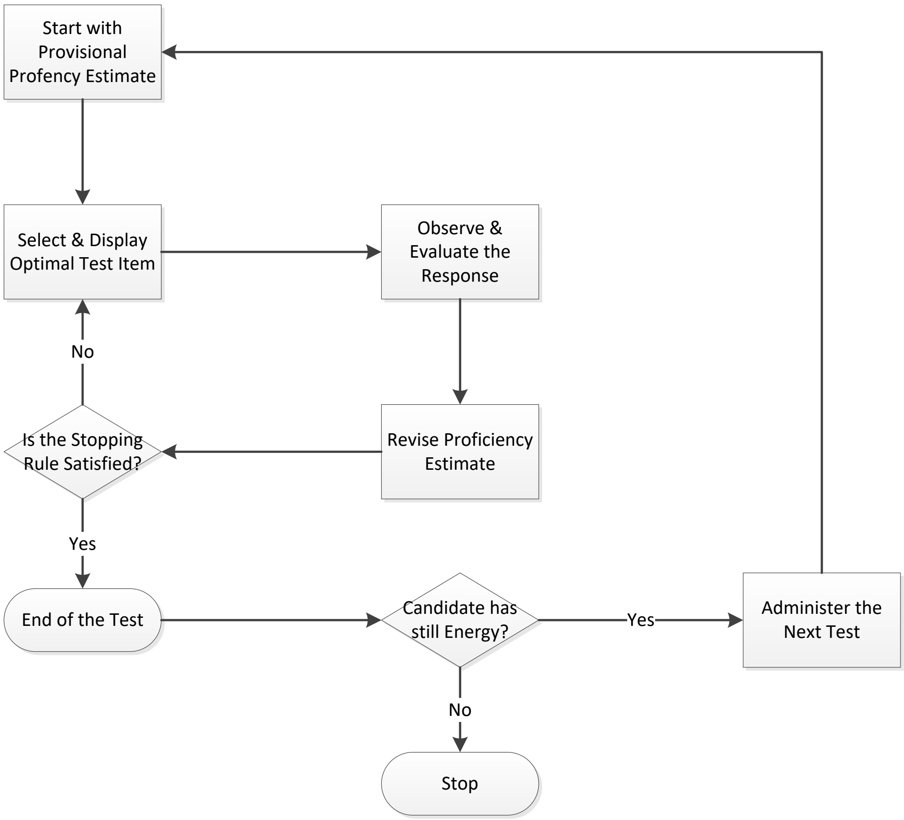


Figure 3: Computerized Adaptive Test algorithm, following the process designed by Thissen and Mislevy (Thissen and Mislevy, 1990).

3.5. Limitations of Computerized Adaptive Testing

CAT focuses on judging the ability of learners in a specific domain of learning, based on the known difficulty of questions, called items, while adapting the length and course of the test to the detected ability of the learner. To offer an adaptive test, a CAT system makes use of a repository of questions with known difficulty measures. Various methods exist to derive the difficulty of "items", while the most prominent is the Item Response Theory (IRT) (Bunderson et al., 1988; Lord, 2012).

According to Welche and Frick, IRT based models need between 200 and 1000 examinees to gather sufficient feedback to estimate the necessary item parameters, which renders systems dependent on stable repositories of items (Welch and Frick, 1993). Welche and Frick also proposes a hybrid extension, using expert systems, to shorten the time needed for estimation, yet, the high number of examinee per item is an obstacle to integrate new items. A different limitation is the validity of items. As items take iterations of examinees to estimate, an early use of items will lead to higher estimation errors and a lesser likelihood for a valid adaptation of the assessment to the user's ability. Makransky develops here an approach for calibration, using phases of random and adaptive assessment to propose a compromise between online item estimation and text length (Makransky and Glas, 2014). Still *the factor "time to use" the CAT-based solution remains a limitation*.

Another limitation is that models as IRT are accounting for only one dimension of ability only, e.g. on mathematical items for algebraic ability. Yet other abilities may be tested with a similar or recombined set of items as the "arithmetic" ability. To overcome the initial limitation, Reckase collects, in his seminal publication "Multidimensional Item Response Theory" (MIRT) (Reckase, 2009), MIRT approaches which can take into consideration different ability sets as a vector of personal characteristics on items and sets of items. While MIRT approaches are well used and well explored, they also introduce new requirements and limitations. To integrate multiple dimensions more complex considerations and higher amounts of examinees are needed. The approach enables a higher flexibility in terms of the number of ability dimensions but it further *limits the flexibility to add new items and consider different dimensions for a test*.

A CAT-based test is focusing specifically on rating the ability of a learner. While this implicitly incorporates the exploration of the field based on the online assessment, the main target is testing and not learning. To learn to further adapt to the learning progress, new approaches are needed. Bunderson, Inouye, and Olsen in 1988 summarizes four generations of computerized educational measurement (Bunderson et al., 1988)(p. 5):

- *"Generation 1. Computerized testing (CT): administering conventional tests by computer."*

- *"Generation 2. Computerized adaptive testing (CAT): tailoring the difficulty or contents of the next piece presented or an aspect of the timing of the next item on the basis of examinees' responses."*

- *"Generation 3. Continuous measurement (CM); using calibrated measures embedded in a curriculum to continuously and unobtrusively estimate dynamic changes in the student's achievement trajectory and profile as a learner."*

- *"Generation 4. Intelligent measurement (1M): producing intelligent scoring, interpretation of individual profiles, and advice to learners and teachers, by means of knowledge bases and inferencing procedures."*

Dynamic changes are part of the expected 3rd generation, to continuously track learners across different parts of a curriculum, independent of individual areas. Generation 4 extends to intelligent scoring, incorporating different sources of information about the learner and introducing inference. While the model "aged" in regard to the taken forecast, the 3rd and 4th generation nevertheless forecasts the current development well and pictures another limitation of CAT-based approaches, that is a lack of interpretation of the user profiles of learners beyond the tracking of abilities, to finally enable a personalized, inference-based ground for recommendation.

To finally enable a new approach, working on a connected knowledge, rather than tracking singular, individual concepts, and being aware of the dependencies between concepts and the individual profile and progress of the learner, rather than tracking single abilities on fixed domains, a different approach is needed. A new approach should be domain and domain structure aware, flexible to changes in the underlying domain and being able to fast adopt to changing and new sets of concepts to learn. Among many available solutions for assessment and learning, especially the STUDIO solution for technology enhanced learning and assessment fits well to the set requirements (Vas, 2007) and can be also applied in an organisational learning context (Gabor and Ko, 2016). The STUDIO system builds upon a complex, flexible and detailed semantic structure – the domain ontology – and makes use of an adaptive assessment test which can be adopted to a range of assessment strategies. Furthermore, the system integrates a rich statistics and visualization interface which enables to extract and visualize the progress of individuals and groups of individuals (Weber and Vas, 2015).

3.6. A Generalized Approach to Adaptive systems

Burgos, Tattersall and Koper define adaptivity in the context of e-learning as *"the ability to modify eLearning lessons using different parameters and a set of pre-defined rules"*, while in contrast *"adaptability is the possibility for learners to personalize an eLearning lesson by themselves"* (Burgos et al., 2007). Further they divide the phenomenon of adaptation into three types:

1. *Interface-based:* Which is also known as adaptive navigation. Elements and options or configurations are adapted in terms of properties as size, position and colour.
2. *Learning flow-based:* The learning process is adapted in terms of the sequence of learning contents. Learning is a dynamic and personalized learning-path, customized to each user based on the known performance for each run of a course or system.
3. *Content-based:* The respective content of resources and activities changes based on the feedback and insights on learning.

Vandewaetere, Desmet, and Clarebout (Vandewaetere et al., 2011) compared 53 well selected publications on the value of learner models for computer-based adaptive learning environments in their study *"The contribution of learner characteristics in the development of computer-based adaptive learning environments"*. One major observation of the study is the great variability in the research domain of adaptive systems. Following in many cases an interdisciplinary approach for research, developed approaches are often not aligned in terms of design methodology and evaluation. Based on this lack and the synaptic character of the study, Vandewaetere, Desmet, and Clarebout extracted a common ground on the development and structure of an adaptive system.

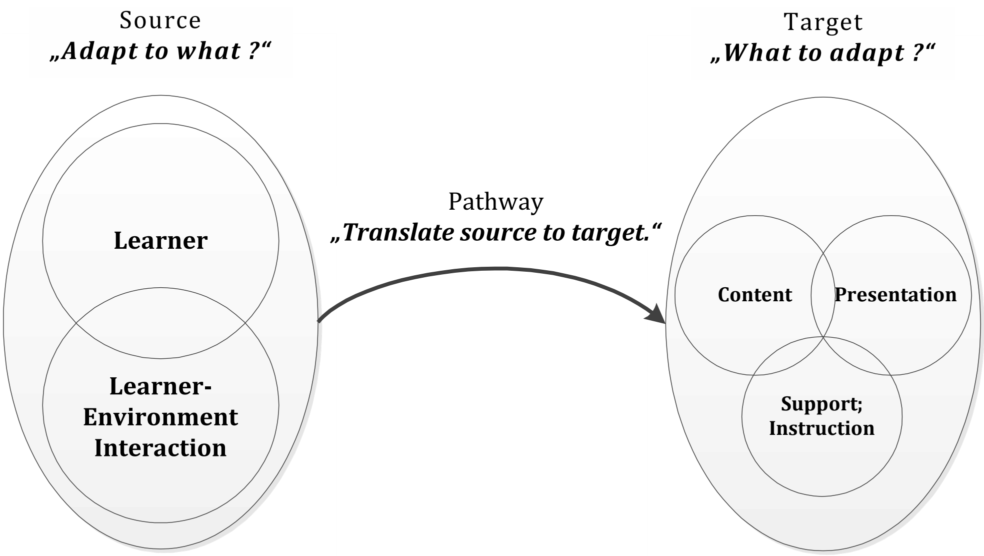


Figure 4: The tripartite structure of adaptive instruction, by Vandewaetere, Desmet, and Clarebout (Vandewaetere et al., 2011).

In their collection, an adaptive instruction can be considered as owning a tripartite nature into which a system is split in terms of functions. The first component captures the source of adaptive instruction (To what will be adapted?), the second component summarizes the target of an adaptive instruction (What will be adapted?). Together the components are connected to each other by pathways of adaptive instruction (How to translate the source into a target?), as the third component of the triple structure. An overview of the concept and building blocks is given in Figure 4.

Following the proposed structure, Vandewaetere, Desmet, and Clarebout, extract three groups of parameters which act as source for adaptation:

- *Cognition*: capturing characteristics of cognition, as working memory capacity, prior knowledge, cognitive style, learning style and learning goals or goal orientation.
- *Affect*: with characteristics of learners, as frustration, confusion, eureka/delight, certainty and frustration, mood and self-efficacy on the boundaries of cognition and affect.
- *Behaviour*: noting characteristics of behavioural styles on computer-based environments, as interaction parameters with a possible need on learner control and help or feedback, the degree of self-regulated learning, or the number of tries per task, received grades and exercises solved.

The target of adaptive instruction can vary strongly with the application of adaption and could be split into three dimensions (Vandewaetere et al., 2011). The most common first dimension captures the learning material and content itself, the accompanying support, display and additional elements which can be adapted to the learner or the interaction with the system, as hints, prompts and agents.

Within the second dimension, time related factors of adaptive instruction are captured as static/constant before the instruction starts, based on earlier information, learner models with long term and general information, then adaption instruction can be based on measures across the use of a system.

The third dimension of the target of adaptive instruction tackles the method which the adaptive instruction is offering and their degree of control as program or instructor-controlled adaptive systems, using learner characteristics evaluated as useful for the application. Further learner controlled methods, which uses shared control and adaptive guidance, can be applied. Finally, in an open learner model (OLM), the user model is open to the user as an additional source of information for learning, shifting the process more into self-responsibility.

Table 4, derived from (Vandewaetere et al., 2011), summarizes the frequency of adaptation targets in adaptive instruction and adaptive learning systems, showing that content, presentation and navigation are the most common targets of adaptation.

The translation of the source into a target of adaptation depends heavily on the type and implementation of an adaptive system. In this regard two different views can be compared: adaptive and intelligent technologies. Vandewaetere, Desmet, and Clarebout summarizes that "*Adaptivity in a system does not make that system intelligent, or the other way around*". As Brusilovsky and Peylo gather – "*Adaptive and intelligent technologies are two different notions*" (Brusilovsky and Peylo, 2003) and as such represent two different approaches. Adaptive systems take into account the information, gathered in a learner model and behave different for different groups of students, while intelligent systems make use of artificial intelligence methodologies to improve the support of the learner – e.g. through machine learning techniques.

Table 4: Frequency of adaptation targets in funded research solutions – publications can address multiple targets (Vandewaetere et al., 2011).

"Target" of adaption	Frequency of a specific adaptation target among 29 analysed researches
Content	
Content	6
Personalized suggestions of content	1
Web-based raw content filtering	1
Lesson contents	1
Appropriate exercises	1
	10
Presentation	
Presentation	4
Presentation style	1
Adaptive presentation	4
	9
Navigation	
Navigation	1
Personalized navigation structure	1
Adaptive navigation	2
Navigational route	1
Navigational control	1
	6
Instruction	
Instruction	4

Instructional strategies	1
	5
Feedback	
Feedback	2
short feedback	1
Appropriate feedback	1
	4
Hints	
Hints	2
Meta-cognitive guidance	1
Suggestion	1
	4
Difficulty	
Difficulty level	1
Level of material	1
	2
Others	
Adaptive group formation	1
Prompts	1
Summaries	1
Assertions	1
Meta-cognitive guidance	1
Adaptive pacing	1
Interactive Help	1
Mastery-decisions	1
Remediation decisions	1
Messaging	1
	10

Limitations can be as diverse as the systems applying adaptive solutions for learning an instruction but following further (Vandewaetere et al., 2011)(p. 124), Vandewaetere, Desmet, and Clarebout address as an universal constraint of current research:

> *“Irrespective of the source of adaptive instruction – the learner, the interaction between learner and environment, or both – most of the developed learner models stay behind the bars of theoretical propositions and only few learner models have been empirically tested.”*

As such, collecting, aggregating and analysing adaptive systems in practical environments and throughout diverse situations and contexts of learning may help to overcome the current deficits of evaluation.

3.7. Learning Management and Intelligent Tutoring Systems

Solutions for technology enhanced learning are usually embedded into the bigger frame of learning management systems (LMS) or virtual learning environments (VLE), which additionally integrate different functionalities supporting the managerial aspects of institutionalized education. On an abstract level, LMSs combine educational services with channels of feedback (Lonn and Teasley, 2009). The different potential technological options and requirements for frameworks of LMSs are summarized by Seidel, Perencevitch and Klett (Seidel et al., 2005).

As part of a complete learning and management approach, specific learner and domain centred solutions exist which support the learner with instruction and access to tailored educational content, including feedback loops for testing the re-adapting the mediated content. In this instructional frame a variety of technology enhanced solutions are applied. The main systems are based on Computerised Adaptive Testing (CAT, Section 3.4), Adaptive Hypermedia Systems (AHS, Section 4) and Intelligent Tutoring Systems (IRT).

Murray summarizes, based on Wenger and Ohlsson, *“Intelligent Tutoring Systems (ITSs) are computer-based instructional systems with models of instructional content that specify what to teach, and teaching strategies that specify how to teach”* (Murray, 1999; Ohlsson, 1986; Wenger, 1987). Murray distinguishes in the frame of ITS: Content Models, managing and providing content in a context for learning and enabling “on-the-fly” content creation; and Instructional Models, modelling instructional strategies to get closer to the benefits of an individual instruction by a teacher.

While all types of technology enhanced systems for learning incorporate the aspect of adaptation to the ability, performance, and/or profile of a learner, especially ITSs focus on reasoning and doing inference on the available information about the learner. In this regard, ITSs focus on filling the role of a tutor, providing an individual teaching experience, making use of a range of adaptation options to adapt to the learner. In line with the allegory of the cave (Jowett, 1901), an automatic tutor can always only be a partial projection of a real tutor, yet, in contrast, Corbett et al. summarizes a clear benefit: *“if such systems [ITSs] realize even half the impact of human tutors, the payoff for society promised to be substantial”* (Corbett et al., 1997).

For the majority of technology enhanced solutions for learning, including ITSs, learning happens in a blended learning environment. Blended learning applies if a system is embedded in a bigger educational frame (curriculum, whole or specific domain) and supports the regular learning. Within this thesis, it is assumed that any solution that is developed within this work to support the learning process will be embedded in a technology enhanced solution for learning. Furthermore, it is assumed that such a solution will work in a blended learning environment, where learners have a basic understanding of the specific domain of learning and use the system to increase their understanding of their individual learning progress and complete their personal knowledge within the given domain.

3.8. Short Summary: Concepts of Adaptive Testing

Table 5: Summary of "Concepts of Adaptive Testing"

Aspect	Description
Need for adaptive testing	In order to enable efficient, personalized learning, learners' knowledge gaps have to be identified individually. At the same time, self-prediction and self-evaluation tend to be flawed. An objective solution for self-assessment can close the gap.
Short assessment and learning cycles	To close the distance between tutoring and the learning performance of a student or worker, short cycles of repeated assessment and learning are important.
Tripartite nature of adaptive systems	An adaptive instruction is based on a user or environment driven source and adapts to a specific target of adaptive instruction. In this regard, an adaptation of a system to learning needs has to be based either on a model of the learner or the domain of learning.

4. State-Of-The-Art of User Modelling

To cope with today's rapidly changing environments, requirements and information overload, there is surely a need for more flexible workers. At the same time – to support the learners and workers of tomorrow – the software they use to learn and work also has to adapt to their specific needs. Technology enhanced solutions for learning or learning support encompass different fields of application, as mobile learning, e-learning, educational games and hybrid solutions. Especially when a system adapts its interface, interaction and behaviour, based on the active interaction with the user and based on information about the user, the model of the user has to be an explicit part of the adaptive solution. One class of user-adaptive systems with a strong focus on individual learning and adaptations are Adaptive Hypermedia (AH) or, more specifically, Adaptive Educational Hypermedia (AEH) systems. Adaptive Hypermedia Systems (AHS) are regarded as a specialization of adaptive systems in education, implementing a specific user model for individual users to adapt hypermedia content as "linked content" to the user's knowledge and goals (Brusilovsky, 1996) [5].

To facilitate adaptive behaviour, a system needs a conceptualisation of the user, captured as the *user model*. The user model is a representation of distinctive features and information about an individual user. This provides the base for adapting a system to a specific user or groups of users, by adjusting the systems behaviour, the look and selection, and the rate and granularity of information. If e.g. a system offers information about specific processes within an organisation, the system can use information about the users' job-role and position to prioritise and personalise search results to offer the most relevant selection (Jameson, 2003; Micarelli et al., 2007) adapted to the user. When a user is reading deeper into the search results, the system can further adapt links to other information sources (Brusilovsky, 2003; Weber and Specht, 1997) or prepare relevant learning resources to learn about the current reading (Dagger et al., 2002).

4.1. User Models: Modelling "What"

User models are a flexible approach to capture features of a user. Specific realisations of a user models can vary in terms of structure, representation, content and complexity. According to Brusilovsky and Millán, user models can be differentiated into: "*models*

[5] Especially in an educational context, user models are also addressed as student models.

that represent features of the user as an individual [… and …] models that represent the current context of the user's work" (Brusilovsky and Millán, 2007). They further distinguish that, the former type of models applies to all adaptive web systems, while contextual models are of importance in changing environments of learning as in mobile and ubiquitous systems, where the working and learning context is subject to change.

Chrysafiadi and Virvou (Chrysafiadi and Virvou, 2013) summarize Self (Self, 1990), who argues *"that student modeling is a process devoted to represent several cognitive issues such as analyzing the student's performance, isolating the underlying misconceptions, representing students' goals and plans, identifying prior and acquired knowledge, maintaining an episodic memory, and describing personality characteristics."*.

Three basic questions are gathered by Chrysafiadi and Virvou for student modelling: *"what to model, how and why"* (Chrysafiadi and Virvou, 2013). Furthermore, with a focus on the user as an individual of a designed model, Brusilovsky and Millán defines five distinctive features of a user model, concerning the questions "*what to model*": "the user's knowledge, interests, goals, background, and individual traits". These different features and their relevance are described in more detail in the following sections.

4.1.1. User's Knowledge

The user's knowledge is of central importance in user-modelling, and in many application cases it is the only feature modelled. The knowledge is observed as the user's knowledge on a specific topic –which is taught or supported by the system as the domain of learning. A characteristic of a knowledge centred user model is that the user's knowledge can change between two points in time. It can increase in terms of learning and decrease in terms of forgetting. As a consequence, the modelling of knowledge in an adaptive system comes with the need to recognise changes in the user's knowledge and update the user model to reflect the detected knowledge changes.

The most basic concept to create a user knowledge model is the *scalar model*. It estimates the level of the existing user knowledge of the provided learning material as a single value on a given scale. E.g. quantitative as a number in an interval [0,5] or qualitative as good, average, bad. In this regard, they behave similar to stereotype models (discussed in detail in Section 4.2).

The common application of a scalar model is the classification of users into different user classes. The model estimates the detected level of knowledge on a defined scale, acting as the range of classes. A scalar model can then be used to classify users into simple classes, fitting their detected knowledge level, and presents content accordingly. A system can then create – based on the detected class – an adaptive behaviour, as e.g. an adaptive presentation which visualizes content fitting to the user's class through filtering and rearranging parts of the presented material.

The model can be used to create full page or partial outputs for adaptive hypermedia (Encarnação, 1997; Premlatha and Geetha, 2015) or feed stronger structured concepts like intelligent documents (Ahonen et al., 1996), where content instances are constructed on demand from different structured resources. Scalar models focus on the user's knowledge and make use of self-evaluation and external testing for the knowledge estimation.

Yet, according to Brusilovsky and Millán (Brusilovsky and Millán, 2007), the main shortcoming of the scalar model is the low precision. User knowledge, captured on a reasonably sized domain, can differ in various parts of the domain and may not be sufficiently labelled by a single dimension scalar assignment. A user may have mastered the knowledge area of classification in the domain of data mining but yet being a beginner in the field of clustering. As Brusilovsky and Millán summarize, *"a scalar model effectively averages the user knowledge of the domain"*.

An improvement on the modelling of user knowledge is given by *structural models*. Structural models are based on the assumption that domain knowledge is clustered into independent fragments which together composes the domain. Existing models can be differentiated across two aspects: the type of the represented knowledge – *procedural* or *declarative* – and the *concept* of the comparison of the user's knowledge against an expert's level of knowledge in an area – addressed as domain model, expert model or "ideal student model" (Brusilovsky and Millán, 2007).

A prominent example for a declarative model is the overlay model, which represents the user's knowledge as a subset of an expert based *domain model*, while the estimation of the user's knowledge is stored for every fragment of knowledge in the domain model. In contrast, procedural models could be modelled as *Learning Networks* (LN), created initially by Koper and Tattersall (Koper and Tattersall, 2004). Learning networks incorporate an expert-based learning network domain which represents a graph of available learning events, called activity nodes. Users are then mapped onto the learning network domain and create a learning track while moving from one learning event to another.

In a LN, the development of a user from one user profile to another can be modelled and tracked as e.g. the transition of one role to another (students, teacher, researcher, practitioner), based on the single events leading to the role change. In this regard LN can be a powerful tool to also share and recommend a learning track to other groups or individuals of learners. The caveat is that the emerging comprehensive networks are not well suited to consolidate a common knowledge domain to efficiently map the progress of the user onto a complete view of the domain.

4.1.2. User's Interest

The focus on the users' interests is tightly connected to user profiles used for adaptive information retrieval and for filtering systems in environments where large volumes of information are available (Morita and Shinoda, 1994) and also used in web-based recommender systems. In the case of recommender systems, the implementation and application considerations have to be further distinguished for e-commerce recommender systems and for systems in technology enhanced, where aspects of both, informal and formal learning, have to be taken into consideration (Drachsler et al., 2009).

Thanks to e-commerce applications, the use of user interest models has increased over the last decade and is now applied as widely as user knowledge based models. Commercial examples are here Amazon, YouTube, Netflix, Yahoo, Tripadvisor, Last.fm, and IMDb (Ricci et al., 2011). The user's interest has importance in environments where the access to information is primarily interest-driven. The initial models implemented a weighed vector of keywords coding the interests and were dominant in early information retrieval and filtering systems (Gauch et al., 2007).

Concept-level models are an extension to the keyword-based approach, modelling the user's interests and their specific connections. Here the interests are modelled as a weighted overlay of a concept-level domain model. It is analogue to the modelling approach of the overlay model (described in detail in Section 4.3). On a content-rich domain model, an additional concept overlay can be used to model the user's interest as distinct aspects, overlaying the domain model. According to Brusilosky and Millán (Brusilovsky and Millán, 2007), "*semantic links in the domain model allow different* *kinds of interest propagation to compensate for sparsity, a standard problem of large overlay models."*

To enable a multi-concept indexing, where selected documents are connected throughout different domains, semantic models – or semantic extensions to existing models – can be used to connect multiple sources. With this approach the closed-world approach of systems and their domain models could be overcome, to offer e.g. personalised news feeds, based on the user's interest and using an additional ontology – which is then used to model the relation between concepts of different domains (Conlan et al., 2006; Frasincar et al., 2011).

4.1.3. User's Goals

A user goal or user task is a representation of the purpose of a user's engagement with the system. A user's goal can come in different shapes and abstraction levels: the target of the current work, a need for specific information, a learning goal or combinations of these goals. Tracking the goal in an adaptive system means to detect what the user wants to achieve. As Brusilovsky and Millán gathers, this comes with an additional challenge (Brusilovsky and Millán, 2007)(p. 10):

> "The user's goal is the most changeable user feature: it almost always changes from session to session and can often change several times within one session of work."

Initial research on user goals are addressing user goals in regard to work goals with the target of adapting interfaces and offering intelligent help systems (Benyon and Murray, 1988; Fischer, 2001; Kaplan et al., 1993). Other studies investigate the relation to learning goals and the context of instructional planning and sequencing systems (Brusilovsky, 1992; Lester et al., 1999). Applications in the context of user goals can be hypertext-based help systems and adaptive information access systems, which benefit from an identification of user goals to adapt their access to the available information (Encarnação, 1997; Francisco-Revilla and Shipman III, 2000).

A user's goal is often implemented using a goal catalogue approach, which behaves similar to an overlay knowledge model (described in more detail in Section 4.3), capturing the user's goals as an overlay to a catalogue of goals. The adaptive system can then recognise goals based on the goal catalogue, while the catalogue could be implemented as a set of goals or as a more complex goal or task hierarchy. Defined goals in goal catalogues are derived from experience or are extracted from adaptive interfaces and instructional planning (Brusilovsky and Millán, 2007).

A goal hierarchy is more stable since goals are continuously decomposed into sub-goals and short-time goals. In a basic implementation, an adaptive system considers only one single goal as active at a time. The user modelling then has to recognise or map which goal is the current active goal. Based on the selected goal, the adaptive system can then adapt its behaviour to the active goal, e.g. through adaptation rules. One limitation of user goals, is that deriving user goals from a detection process is not precise in general. This is especially true in environments where different information sources may capture the current goals or only large scale logs of interactions deliver the appropriate feedback to derive a goal (Lu et al., 2013) by using methods of data mining.

4.1.4. User’s Background

The *user’s background* describes previous experiences collected outside of the *core domain* of a specific application. Backgrounds can encompass granular experiences and general experience labels as job roles and responsibilities, working experiences, specialized process experiences and abstract backgrounds as specific views on the knowledge domain. A background is considered as a stable set of features, which doesn’t change or change only slowly over time.

A user’s background is a well-used input for content adaptation and can be used to capture background information as the profession or known job roles to adapt content in that context. The user’s background is either filled with known information about the user or is detected and derived from the user’s interaction with a system. E.g. in a medical system the information about the profession could be applied to scale the level of knowledge, the presentation of knowledge and the complexity of the used language to the detected levels of a given target group such as doctors and nurses (Beaumont, 1994).

Brusilovsky and Millán (Brusilovsky and Millán, 2007) outline a similarity between the perception and modelling of the user’s background and the user’s knowledge: *“By its nature, user background is similar to the user’s knowledge of the subject [as] i.e., it is also mostly a measure of knowledge beyond the core domain area”*. However, the use and representation differs in application.

4.1.5. User’s Individual Traits: Cognitive and Learning Styles

A user’s individual traits are features which define the user as a specific individual. Individual traits are a collective name for a number of concepts and theories. They gather aspects as: personality traits, cognitive styles, cognitive factors and learning styles. Individual traits are stable features which do not change or change only over an extended period of time. Individual traits are a valuable tool to adapt the system to the specific needs of individuals or groups of users and can be used to differentiate users into groups with similar needs.

Individual traits are collected directly from the user by well-designed psychological tests, essentially creating a specialised psychological profile. Each psychological test used in this regard, is specialised and custom prepared for one specific theory. An alternate methodology is the use of machine learning based methodologies, detecting the individual traits of a user, based on static information about the user and collected information from the utilization of the system by the user (Garcia et al., 2007), e.g. captured as event-logs and click-streams (Gündüz and Özsu, 2003).

As cognitive styles and learning styles share a number of features, concepts and considerations they are interchanged in some research. Learning styles focus on human learning, while cognitive styles focus on the actual processing of information. In comparison to cognitive styles, learning styles are a more active topic and their implementation in adaptive systems is also more frequent, following the implicit potential of an impact on the quality of learning. Akbulut and Cardak conducted a study in 2012, comparing 70 studies of peer-reviewed articles in international databases on the use of learning styles in adaptive educational hypermedia (Akbulut and Cardak, 2012). Out of the analysed studies 71.4% addressed learning styles (50% Felder-Silverman, 8.6% Kolb, VARK 7.1%, Honey and Mumford 5.7%) and only 17.1% cognitive styles. Still – especially for research considering learning styles in general – the notion of learning styles could sometimes be taken as a proxy for cognitive styles.

The similarity is underlined by the observation of Riding & Cheema (Riding and Cheema, 1991)(p. 194) who suggest that:

> “'Learning style', seems to have emerged as a more common term or a replacement term for cognitive style in the 1970's. *[..T]he impression that is formulated in the*

usage of these terms is that those working under the umbrella of 'learning style', take cognitive style into consideration, but would probably describe themselves as interested in more practical, educational or training applications and are thus more 'action-orientated', while the term cognitive style has been reserved for theoretical, academic descriptions."

Riding and Rayner defines cognitive style as *"an individual's preferred habitual approach to organising and representing information"* (Riding and Rayner, 2013). Triantafillou, Pomportsis, Demetriadis and Georgiadou extend this definition in the following way: "*Cognitive style is usually described as a personality dimension that influences attitudes, values and social interaction. It refers to the preferred way an individual processes information.*" (Triantafillou et al., 2004).

As for learning styles, the definitions of cognitive styles and the emphasis of different aspects can differ across researchers and their focus of consideration. The main views can be categorised into the dimensions of field-dependent/independent, impulsive/reflective, conceptual/inferential, thematic/relational and analytic/global, as summarized by Liu and Ginther (Liu and Ginther, 1999).

Ausburn and Ausburn collect three specific aspects of cognitive styles (Ausburn and Ausburn, 1978), which can also characterize learning styles:

- *Stability: "[...] generality and stability over time and across tasks."* The majority of cognitive styles are behaving consistent across tasks for individuals.
- *Relationship to ability: "[...] minimal relationship with traditional measures of general ability."* Cognitive styles correlate to traditional measures as IQ tests but the correlation cannot account for all the variance in general ability.
- *Relationship to learning tasks: "[...] the most important characteristic of cognitive styles, at least to the field of educational technology and instructional design, is their relationship to a number of specific characteristics, abilities, and learning activities. "*

Schmeck distinguishes two overall types of learning styles – a global-holistic/field dependent/right brained and a focused-detailed/field independent/left brained and asserts that "*although both styles are equally good for problem solving, each style is likely to be* *associated with greater efficiency in specific tasks. The most effective problem solvers should exercise strategies connected with both aforementioned styles*" (Schmeck, 1983).

Coffield, Moseley, Hall and Ecclestone evaluated and compared in their seminal report *"Learning styles and pedagogy in post 16 learning: a systematic and critical review"* in detail 13 of the most influential models of learning styles (Coffield et al., 2004a). The selection of 13 comes out of a pool of more than 70 theories. This shows the strong appeal of the field of learning styles but it also pictures a fractured research area with a high number of different, partially overlapping, approaches.

Learning styles and cognitive styles are used in a variety of applications in the field of adaptive educational hypermedia. Akbulut and Cardak collect in their study on *"Adaptive educational hypermedia accommodating learning styles: A content analysis of publications from 2000 to 2011"* (Akbulut and Cardak, 2012), dynamic and holistic student modelling approaches under support of learning styles application with variables as: knowledge/competency level (Triantafillou et al., 2003), cognitive traits (Graf and Kinshuk, 2010), multiple intelligences (Maria Zajac, 2009), learning goals (Del Corso et al., 2001), cognitive styles (Prieto and Garcia, 2006), learning modalities (Rumetshofer and Wöß, 2003), spatial ability (Wang et al., 2006) and learning behaviour (Tseng et al., 2008).

Learning styles are a well-used and well researched support for adaptive educational hypermedia but especially in the recent years, researchers isolated limitations to the application of learning styles. One major critique is that across all present studies the general impact of learning styles on the learning performance is weak or missing at all. Willingham, Hughes and Dobolyi use a constructed example to visualize the common outcome of analysis: assuming a fitting and mixing auditive learners and auditive material and visual learners and visual materials constellation a correlation between the learning style and the material could be indeed proven but for a specific case still one single representation style would still create a "best" result for both groups (Willingham et al., 2015). Another common critique is that *"Although there are a multitude of inventories and models for assessing learning styles, most are not reliable"* (Coffield et al., 2004b).

Robotham widens the discussion on learning styles and proposes, instead of making a forced learning style/performance link, to promote a more self-directed learner (Robotham, 1999):

> *"Using existing inventories of learning styles, individuals are simply allocated to a narrow range of categories, containing a limited number of learning activities to which they are, in theory, best suited. [...] Higher education teaching should seek to move beyond the enhancement of performance [...] and consider the development of foundation skills, such as self-directed learning. An able self-directed learner may [...] choose to use a particular learning style that is relatively narrow in nature, but they are consciously taking that decision, in view of their perception of the needs of a particular situation."*

While this view shares the criticism of an unreflective use of learning styles, it also highlights a profit for adaptive systems: Adaptive systems can provide a conscious choice of learning styles for the content- and presentation-adaptation to the user. The choice may be supported by methodologies to detect learning styles and offer choices, but the choice should be transparent to the user. In this regard, an active choice of learning styles can complete of the user model, considering the change of the selected learning style over time.

4.2. Stereotype Models

A basic user model is a simple feature-based model. It takes the concept of a user and attaches different specific features of individual users as knowledge, goals, interests or learning styles. While a user uses an adaptive system, and interacts with the presented content, the features of a user change. In this regard, to use a feature-based model, the system has to track and detect the changes of an individual users for individual user profiles, which also have to be handled individual for the adaptation of the system. An alternative approach here the stereotype modelling which groups users and user features for a stronger framing of the possible adaptations.

Stereotype models go back to the work of Rich (Rich, 1979). Stereotype models aim to cluster the user of a system into groups, which are called stereotypes. These stereotypes are then used to handle each user the same way within a specific stereotype class but differently across classes. In contrast to a feature-based model a user is then represented by one stereotype.

Within a stereotype users share a set of features which composes the stereotype as a class. But for modelling and adaptation only the information of the stereotype is used and an adaptation algorithm only has to handle a single information in contrast to sets of features. In this regard the goal of stereotype modelling is to create a mapping from a specific constellation of features to a stereotype class, representing the common, shared knowledge and features of users in that class. If a user changes the personal constellation of features it doesn't require a remodelling of the user but a reclassification to another stereotype, as shown in Figure 5, based on (Hnida et al., 2014).

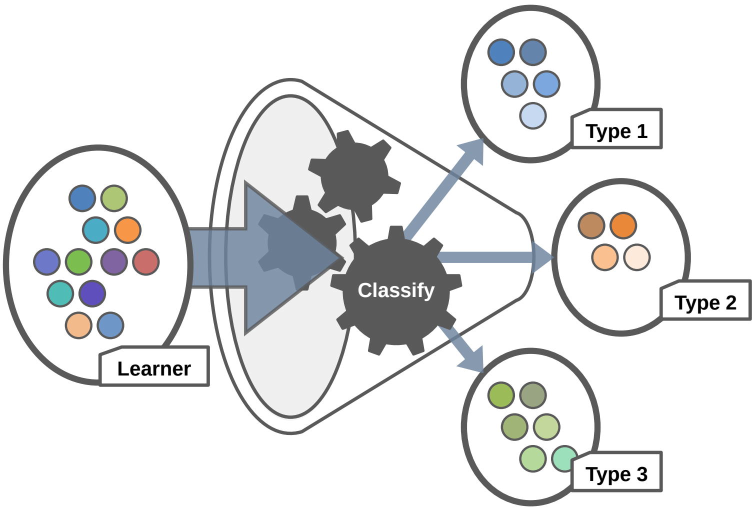


Figure 5: Classifying learner into stereotypes, based on the figure of Hnida, Idrissi and Bennani (Hnida et al., 2014).

A user is assigned to a stereotype when the detected user's characteristic fit to the features in a respective stereotype. To derive a class, following Kay (Kay, 2000), a stereotype M integrates a set of triggering conditions $\{tM_i\}$, where each $\{tM_i\}$ is a Boolean expression based on a feature or a function over a set of features of the user model, and a set of retraction conditions $\{rM_\mathrm{j}\}$. A stereotype M is active when Equation (2) becomes true and a stereotype is deactivated if Equation (3) becomes true if any of the respective conditions become true. In this regard the assignment of stereotypes shares similarities with an inference process (Rich, 1979). Usually one stereotype is implemented as a default class in which users are added before a stereotype triggers. This default class can also act as a fallback.

$$if \; \exists \; i, \qquad tM_i = true \; \rightarrow active(M) \tag{2}$$

$$if \; \exists \; j, \qquad rM_j = true \; \rightarrow not \; active(M) \tag{3}$$

Stereotype models are widely used across adaptive and personalized tutoring systems and are especially suitable to solve the problem of initialization of user models in cases where only limited information is available but additional user information can be inferred from the assigned stereotype. In this regard, stereotypes can be used in combination with other modelling approaches, where stereotypes are used to initialize the user model with known features then to switch to another methodology more suitable for the continues update of the user model or to improve stereotype models by integrating machine learning methodologies (Tsiriga and Virvou, 2004).

Even though stereotypes are a solid and appealing approach, they come with limitations. Stereotypes are pre-modelled – either by a human expert or based on machine learning methodologies and therefore inflexible and have to be explicitly updated or rebuilt in case of overall changes (Tsiriga and Virvou, 2002). Furthermore, stereotypes need a pre-classification to identify classes and for the stereotypes, while the process could be error prune and time consuming.

4.3. The Overlay Model

Brussilovsky and Millán summarize overlay knowledge modelling as the representation of *"an individual user's knowledge as a subset of the domain model that resembles expert knowledge of the subject"* (Brusilovsky and Millán, 2007). In this regard it resembles, in the context of education, a mapping of the known user's knowledge as a subset (Martins et al., 2008) onto a specific domain of expertise which is the object of learning of a user within an adaptive system.

The model assumes that the student can have incomplete but correct knowledge of the target domain. The difference in the set of the student and the expert knowledge is assumed to be a lack of skills and knowledge of the specific learner (Bontcheva and Wilks, 2005), which has to be overcome through learning to finally master a domain. Overlay models, based on their nature, tend to be designed expert-centric and expert-driven and are commonly directly engineered by experts or with the support of an expert rather than being extracted and modelled automatized from other representations.

The initial concept of overlay models goes back to the work of Stansfield, Carr and Golstein (Stansfield et al., 1976) and their seminal ITS implementation of the Wumpus advisor I (WUSOR). The solution was targeted at graduate and undergraduate students to advise in cases of students ordering "bad" moves within the rule based game Wumpus. In case of bad moves, the system uses the overlay model and a set of rules to explain the flaws of the move, together with the rational of alternate moves, teaching the students concepts of reasoning, probability and uncertainty. The Wumpus advisor was one of a number of initial Intelligent Tutoring Systems (ITS) which shared a strong connection to expert systems and used different models, which were similar to the overlay model but different in their specific designs and implementations, as e.g. BUGGY (Sacerdoti, 1977), the Leeds Modelling System (LMS) (Sleeman and Smith, 1981) and GUIDON (Clancey, 1987).

Overlay models can be realized in different ways and have in common that the student's known knowledge is represented as an overlay on the expected knowledge of a specific domain. The simplest, structure-less overlay is a set or vector of independent concepts the user has mastered, which enables a fine-grained tracking of the user's knowledge. Yet, concepts are seen as independent from each other and for situations with a low number of observations and a large amount of concepts, only the observed fraction of concepts are predicted (Brusilovsky and Millán, 2007). One way to overcome this limitation is the introduction of relations between concepts. Additionally, an overlay model can be further detailed by attaching qualitative, numeric and uncertainty-based labels, called "weights", to concepts in the student overlay of the model to represent the user's mastering of single concepts more granular.

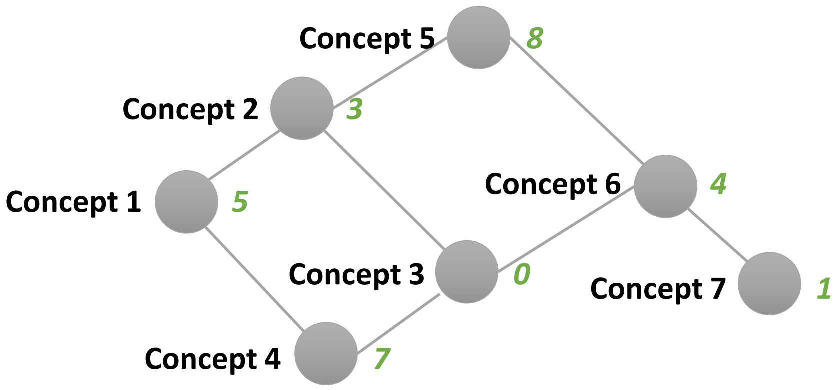


Figure 6: A domain model of the domain to learn, including a numeric overlay.

An advanced overlay model can be created by introducing a domain model, integrating relations between the modelled concepts. A domain overlay model can be visualized as a network of connected concepts, as presented on Figure 6. The network visualizes the domain of learning and a numeric "overlay" to the domain pictures the user's known knowledge on the specific concepts. Concepts are shown in relation to each other and numeric labels on concepts show the detected degree to which concepts are mastered by the user.

4.4. Semantically Enhanced Overlay Models

A further extension to a domain overlay model is the introduction of different types of relationships within a domain model, which is motivated by semantic networks, known as a *network model* (Brusilovsky and Vassileva, 2003).

A major profit of a semantic network model is that the modelling of more complex domains becomes possible. A system can identify related concepts which are relevant to master connected concepts and semantic relation types can be used to model relations as "prequisites" and other relations to express dependency. Based on the semantic extension, users can receive a more detailed and complex feedback on their learning needs. Ontologies can have different applications in the context of user models and adaptive systems as: providing a general user model ontology (Heckmann et al., 2005), modelling the learning environment and desired competencies (Hnida et al., 2014), modelling the user's personal and academic information (Panagiotopoulos et al., 2012) and user model interoperability (Brusilovsky et al., 2005).

According to Panagiotopoulos (Panagiotopoulos et al., 2012), ontologies have been used widely for user modelling as 1) they provide a formal representation of abstract concepts and can be reused and extended for many tasks and 2) enable the extraction of new knowledge through inference on the content of the ontology. Specifically, in the context of user models and adaptive systems, Sosnovsky and Dicheva (Sosnovsky and Dicheva, 2010), clusters the application of ontologies in two main directions: 1) to model the structure of the domain, while elements of the ontology represent characteristics of the user and 2) to explicitly structure the user profile in terms of dimensions and states. While the later direction shows similarities to stereotype models, the earlier direction correlates to the idea of the overlay model.

A semantic enhanced overlay model provides the profit of an interchangeable domain representation of the domain to learn, while providing the potential for an adaptive system to reuse the structure and semantic of the model to map the user's learning progress and structure the recommendations for learning. Sosnovsky and Dicheva (Sosnovsky and Dicheva, 2010) highlight here that the *"[r]epresentation of a domain model as an ontology and modelling a user's knowledge, interests, needs, or goals as a weighted overlay on the top of it, enables the usage of standard representation formats and publicly available inference engines, as well as an access to a vast pool of technologies for ontology mapping, querying, learning.."*.

The idea of a semantic enhanced overlay model can be further extended by the "projection approach" to overlay models, developed by Brusilovsky (Brusilovsky, 1994). In his approach, the user overlay – which captures the state of the user and is projected onto the domain of learning – is presented within a user-centred architecture. In the projection approach the internal overlay model is combined with a set of rules (How to use and how to transform the model?) to adopt or “project” the model for different use-cases within the system. The concept is to provide one application centric user model which then can be exploited differently by different components of an adaptive system, as e.g. the learning manager or the visualization interface, to offer adaptation for a learning environment.

Components which offer different functionalities for adaptation, also can have different requirements towards their input to finally facilitate the adaptation. E.g., while the user-model may define degrees to which the user mastered a concept, components may only need the information if something is mastered or not. The overlay projection uses at the base the classic overlay approach to represent the user’s progress as numeric overlay onto the structured domain model, e.g. the progress on a specific topic in the numeric range of 0-5. To account for the different component needs, Brusilovsky applies a threshold technique as a projection and “projects” a numeric range of the degree of mastered knowledge to a binary concept of known and unknown, tailoring the projection to the information need of the component. Which projection to apply, when and how is decided by a set of rules, called a *projector*, while one projector is used for one specific component or scope. Projections can be used to change the use of the user model online and on-demand without changing the user model itself.

In a similar fashion, also a semantic model could be “projected” to the specific needs of a component. To do so, the semantic of single entities and their relations can be translated to a numeric value, which e.g. can be used – through applying projection rules – to modify the degree to which a user mastered a concept. E.g. if – for a given knowledge element in a semantic enhanced domain model – relations exist that mark other connected knowledge elements as requirements, a fitting projector (rule) can define that the selected knowledge element can only be fully “known” if the knowledge elements which are connected through “required by” relations are also “known”. This way, the structure and semantic of the domain model further specifies the requirements of the associated learning process and help to better scale the behaviour of components.

4.5. User Model-based Recommender Systems

A user model is always connected to a specific purpose, recommending on the right education, on the right set of adaptation rules for the interface and content of a system but also recommending on products in a more industrial context. The practical implementation of recommendations and the theoretical foundation to derive recommendations is given by the area of Recommender Systems. Drachsler, Hummel and Koper collect in regard to user models (Drachsler et al., 2009)(p. 1):

> *"The increasing use of Recommender Systems (RSs) that support users in finding their way through the possibilities on offer on the WWW is obvious. Many online companies like amazon.com, netflix.com, drugstore.com, or ebay.com (Linden et al., 2003; Schafer et al., 1999) are using a RS to direct the attention of their costumers to other products in their collection. The general purpose of recommender systems is to pre-select information a user might be interested in (Adomavicius and Tuzhilin, 2005)."*

The same is valid for applications in education and learning, building on the same set of user models as e.g. user interests (Pazzani and Billsus, 2007), with the difference of a stronger focus on the user's knowledge and goals. According to Drachsler, Hummel and Koper a recommendation system in the scope of technology enhanced learning has "*the goal to support learners in their competence development in order to achieve a specific learning goal*". A learning goal is here connected to a specific concept, or domain of concepts which has to be mastered on a specific level. Specially in the context of "learning networks" – as networks of learners, where leaner can share learning activities in different roles (teachers, learner or knowledge providers) – Drachsler, Hummel and Koper suggest two different types of recommendations: to "*structure Learning Activities in a pedagogical way*" and to "*suggest emerging learning paths to learners*". A further overview of learning networks in technology enhanced learning is given by Manouselis (Manouselis et al., 2011).

Furthermore, Anderson distinguishes the specific feedback of recommender systems in terms of the power of the underlying model, in the context of expert systems (Anderson, 1988):

- **Black Box Models:** makes use of methods of reasoning on a domain, without using an explicitly model of the domain, generating a correct input-output behaviour without being able to deliver instruction.
- **Glass Box Models:** makes use of an expert modelled or derived knowledge representation, which enables reasoning and an explanation of the results. Anderson emphasis: "*the expert system [..] is going to be more amendable to*

tutoring than a black box model because a major component of this expert system is an articulate, humanlike representation of the knowledge underlying expertise in the domain".

- **Cognitive Models:** makes use of findings within the field of cognitive science to not only represent and reason but also to access knowledge in a humanlike way.

Pazzani and Billsus collect in the context of product recommendations that: *"[a]lthough there are different approaches to learning a model of the user's interest with content-based recommendation, no content-based recommendation system can give good recommendations if the content does not contain enough information to distinguish items..."*. In the context of technology enhanced learning, items can be interpreted as content or adaptation goals, fitting to the specific user's need. Content which cannot be distinguished based on the content itself or based on the approach to apply the user model onto the content and other adaptation goals, cannot be recommended in a valuable way to the user. Furthermore, users who cannot be differentiated based on their user model will not receive different tailored recommendations from a system. So only by being consciousness on the limits of the modelling a valuable and working recommendation is possible. One aspect in this sense is the need to be conscious about the "How" of the support for learning and – consequently – how a system should provide the specific learning support to enable a continuous and sustainable learning.

4.6. Short Summary: User Modelling

Table 6: Summary of "User Modelling"

Aspect	Description
Adapting to "what"	User models offer a conceptualization of users, modelling users' features relevant in the adaption of adapting system behaviour. Knowledge can be presented as a feature of the user or as a model of the domain to master.
Student overlay for the domain model	Overlay models are matching user knowledge to a network representation of a targeted expert's knowledge and monitor the progress of assessment and learning by labelling the concepts in the model with numeric values, tracking the performance.
Need for a semantic domain model	A semantic network model enables to express more complex domains and derive a better learning feedback. Especially ontologies are a major improvement as a domain model (since ontologies are structured, interchangeable, and model the meaning of individual concepts and their relations).

5. Learning Theory Considerations

Throughout all situations of life education stays as a stable companion. This seems to stay true with new and changing requirements coming with the new technology enhanced society. But as the environment in which knowledge and processes to gain knowledge are needed is changing, so are the requirements for education. While the discussion on how to provide the best education is diverse and in continues motion, one general question persists as a steady consideration: How should education be provided, e.g. top-down, starting from the general and highlighting correlations or bottom-up, teaching the specifics first and introducing the topic's fundamentals. And further – how should this be reflected in supporting technologies for education.

In a class room situation for a top-down approach a teacher will try to give a general overview first, introducing the big picture paired with an overall motivation, content- and outcome-wise, showing the correlation between the aspects of the particular field and immersing the students in a way which triggers the personal motivation to learn and master an area. Following the example of De Grauwe (Grauwe, 2010) in the context of Macroeconomics, addressing the top-down approach from a system point of view, a top-down system is a system which respective agents understand the fully. They are capable to see the system as a blueprint in which they can optimize their actions. An alternative understanding would be to see the system as a building which can be represented by its blueprint which is understood by its architect.

Contrasting to a macro-first top-down approach, a bottom-up way of teaching will tackle the details of a specific topic area first to piecewise develop the topic towards the understanding of the whole area. The content focus is on teaching and mastering facts and rules, important to understand the detailed parts of the overall picture as building blocks which create the whole. A bottom-up approach is instructor-driven and targets to break down the complexity of an area and simplify the learning process through mastering the details first, making use of memorizing and repetition. Following further the example of De Grauwe taken from a system perspective, in a bottom-up system, individuals understand only parts of the system. The system then works by applying simple rules on the individual level, together creating the whole, which this way resembles the behaviour of natural systems.

Strongly interwoven with the consideration of a top-down and bottom-up education is the comparison of behaviourism- and constructivism-based learning. A bottom-up education is here favoured by the concept of behaviourism, which is based on studies on animal behaviour and the response to rewards and punishment, while top-down education aligns with the concept of constructivism, which understands learning as a process of connecting new knowledge to previous learned knowledge. As both pairs of concepts are related, contrasting top-down against bottom-up teaching and behaviourism against constructivism is in many situations a proxy for each other.

5.1. Behaviourism and Top-down Education

As summarized by Ertmer and Newby (Ertmer and Newby, 2013), behaviourism makes use of the concept of stimulus and response. Learning, following behaviourism, occurs when a learner gives an adequate response to a presented stimulus. E.g. while showing a learner a specific math problem, the problem represents the stimulus, while the fitting answer of the learner is the response. The key question of behaviourism is then how to strengthen and sustain the association between the stimuli and a successful response. Furthermore, the long-term goal is to foster positive responses by adding reinforcements to positive responses.

The proof of the positive effects of positive and negative reinforcements is going back to the experimental work of Skinner (Skinner, 1974). Following the theory, the learner is characterized as reactive to the conditions of learning rather than active, weighting the environment higher than the inert and active motivation of the learner.

Teaching in this framework takes a strong emphasis on preparing and controlling the arrangement of stimuli and the consequences of given responses. Furthermore, the learner is continuously assessed to recognize where to start the instruction and to detect which reinforcement actions are effective for a specific learner. For transferring learned knowledge to new situations, learners are expected to generalize situations, with features shared or similar to previous learned behaviour.

Organizing teaching in the frame of behaviourism, emphasis strategies which improve the linking between stimulus and response with methodologies like reinforcement and practice. Learning happens through making use of recalling facts, generalization, association of explanations and performing/repeating learned procedures. Furthermore, and simultaneously limiting, it is recognized that behaviourism isn't suitable to explain higher level skills and processing.

5.2. Constructivism and Bottom-up Education

Furthermore summarized by Ertmer and Newby (Ertmer and Newby, 2013), constructivism "is a function of how the individual creates meaning from his or her own experiences". Constructivism envisions the mind as a filter, which filters the world to create its own reality. In this regard the mind is conceived as the source of the derived meaning. The knower constructs a reality or interprets it, based on his or her perception (Jonassen, 1991). Following Jonassen, the knowledge is constructed as a result is based on previous experience, the mental structures, and beliefs a person uses to interpret objects and events. E.g. in a class room situation a teacher would introduce the general problem to solve and give the question of methodology to the learners for reflection and construction of their own methodologies in favour of connecting to their previous experience to only then advancing to the detailed methodologies.

The concept of constructivism, as collected by Perkins (Perkins, 1992), goes back to the seminal work of Piaget (Piaget, 1954) under the influence of cognitive psychology, guided by researchers as Bruner and Neisser and Goodman. In contrast to the view of behaviourism, constructivism takes the view that the knowledge of a learner is mind dependent and has to be mapped onto a learner. It receives ongoing attention since then, and is used to observe and reason on learning as an experience driven function, creating meaning based on new and previous experience.

Teaching in the frame of constructivism takes an emphasis on practical involving the learner in situations which are embedded into a meaningful context. Understanding is strongly connected to the number of experiences collected in the context of the target education, where the learner develops ideas to master the situations. Learning and the transfer of knowledge always takes place in a context in the view of constructivism and the different contexts offer different links to the knowledge to learn.

To organize teaching in the scope of constructivism the teacher focuses on telling the story of a task, rather than setting the structure for the learning of the task. Bednar, Cunningham and Duffy argue in this regard that *"information cannot be remembered as independent, abstract entities"* (Bednar et al., 1991). Learning in the frame of constructivism doesn't happen by learning and following a set of rules and strategies. The learning and effective use of knowledge comes results from using the acquired tools in real world situations.

5.3. Technology Driven Ways to Learning Theories

In contrast to the clear separation of learning theories for a promotion of learning, rather than teaching following a specific theory, the conscious selection of a theory by the teacher for a specific learning situation is crucial. The set of selectable learning theories is wider than the top-down and bottom-up or behaviourism and constructivism and includes further theories as cognitivism and recently connectivism and offer alternative views on the theories with the concepts of objectivism, pragmatism and interpretism. Resulting, the question arises if a learning theory is regarded as the leading theory and if none is leading which theory to select at a given time – moreover considering new developments in learning environments like the rising impact of technology and especially information systems.

An alternate view on the question of a selection of learning theories for learning environment is given by one of the most prominent citations for the use of constructivism, taken from the seminal work of Bednar, Cunningham and Duffy (Bednar et al., 1991):

> *“Instructional design and development must be based upon some theory of learning and/or cognition; effective design is possible only if the developer has reflexive awareness of the theoretical basis underlying the design.”*

While the citation is taken as a strong emphasis for constructivism by the constructivism community, Duffy corrects this image in an interview (“Interview with Thomas Duffy,” 2000):

> *“It seems quite logical to me that no one is going to design instruction that they think will be ineffective--or counterproductive to learning. If they are acting in a way that they believe will promote learning, then they must have some notion of what learning is all about. […] Besides acknowledging this particular point, we were simply arguing that it would be worthwhile for designers (and educators in general) to become more aware and better able to articulate their views.”*

Duffy underlines here, in contrast to a default learning theory selection, the importance of the awareness of the designer in the process of instructional design, rather than the selection of one and only one learning theory.

In this regard the awareness of the concepts and tools for teaching maybe sufficient, but a thorough and conscious instructional design is mandatory for the learning success. A careful and purpose- and situation-fit selection and modification of a learning theory outweighs the defaulting to one singular learning theory. Siemens emphasises in his seminal work about connectivism (Siemens, 2005) the definition of learning of Driscoll (Driscoll, 2005) *“a persistent change in human performance or performance potential” that “must come about as a result of the learner’s experience and interaction with the world”*. As Siemens notes this definition *“encompasses many of the attributes commonly associated with behaviorism, cognitivism, and constructivism”* (Siemens, 2005).

Seeing learning and learning theories from this general perspective even a hybrid solution for organizing learning environments can be feasible and effective if designed carefully. For an insight into the possible methodologies and tools a second look onto the aspects of behaviourism and constructivism are valuable. A comparison of behaviourism and constructivism is given in Table 7, based on the classification by Schunk (Schunk, 1991) and the elaboration and extension of Ertmer and Newby (Ertmer and Newby, 2013).

Table 7: Comparison of behaviourism and constructivism, based on the classification by Schunk (Schunk 1991) and Ertmer and Newby (Ertmer and Newby 1993).

	Behaviourism	**Constructivism**
How does learning occur?	Learning occurs when a proper response is given following the presentation of a specific environmental stimulus.	Learning is the process of creating meaning from experience. Learners build an interpretation of the world based on individual experiences and interaction.
Which factors influence learning?	The arrangement of stimuli and consequences is the most important factor. Environmental conditions are emphasized before learner related aspects.	Learner and environmental factors are important and the interaction between both creates knowledge. Learning happens in a situational context and it is important in which situation the knowledge is used.
What is the role of memory?	The role of memory is neglected by the behaviourism.	Constructivism isn’t focused on memorizing particular facts. Understanding is developed based on continues use of knowledge in specific situations. Learning needs the factors activity, concept and cultural context.
How does transfer occur?	Transfer to an application is the result of generalization.	Transfer is reached by involving the learner in authentic tasks linked to and within meaningful contexts.
What types of learning are	Strategies of learning are important which strengthen	Learning is always observed in a specific context and in conjunction

best explained by the theory?	the stimulus-response associations. The learner performs in situations of learning involving discrimination, generalization, association and chaining.	with a specific content. Constructivism is most effective for advanced knowledge acquisition while initial acquisition is better supported by behavioural or cognitive approaches.
What basic assumptions / principles of this theory are relevant to instructional design?	• producing observable and measurable outcomes in students • pre-assessment of students to determine where instruction should begin • mastering early steps before progressing to more complex levels • use of reinforcement to impact performance • use of cues, shaping and practice to ensure a strong stimulus-response association	• emphasis on the identification of the context in which the skills will be learned and subsequently applied • emphasis on learner control and the capability of the learner to manipulate information • need for information to be presented in a variety of different • ways • supporting the use of problem solving skills that allow learners to go "beyond the information given" • assessment focused on transfer of knowledge and skills
How should instruction be structured to facilitate learning?	The instruction is structured around presenting the target stimulus and providing opportunities for the learner to practice to make the proper response.	Instruction has to be designed to show students how to construct knowledge, give multiple perspectives for specific problems. The designer of the instruction has to instruct how to create meaning and how to monitor, evaluate and update constructions. Further the designer has to design experiences for the learner with relevant contexts in which a task can be experienced.

The comparison of Table 7 underlines the different trends of behaviourism and constructivism learning theories: behaviourism as the top-down, decomposing, fact oriented learning theory which is focused on stimulus/response pairs and constructivism as a bottom-up, generalizing, context oriented theory which is focused on linking experiences to new situations.

5.4. Connectivism as a Network Driven Theory for Learning

Cognitivism, in contrast to constructivism and behaviourism, is a more technology and network oriented theory. It focuses on knowledge as symbolic mental constructs within the learner's mind, while learning stores the symbolic representations to the learner's memory.

Yet – as Siemens addresses (Siemens, 2005) – the majority of learning theories conclude that learning occurs inside a person only and fail to address learning outside of learners as technology based learning. Additionally, existing theories do not tackle personal and organizational learning within organizations and neglect to assert the value of what is being learned. Seizing these limitations, Siemens collects seven questions not yet addressed by current learning theories, especially taking into account the shift to a new technology enhanced society (Siemens, 2005)(p. 3):

- *How are learning theories impacted when knowledge is no longer acquired in the linear manner?*
- *What adjustments need to made with learning theories when technology performs many of the cognitive operations previously performed by learners (information storage and retrieval).*
- *How can we continue to stay current in a rapidly evolving information ecology?*
- *How do learning theories address moments where performance is needed in the absence of complete understanding?*
- *What is the impact of networks and complexity theories on learning?*
- *What is the impact of chaos as a complex pattern recognition process on learning?*
- *With increased recognition of interconnections in differing fields of knowledge, how are systems and ecology theories perceived in light of learning tasks?*

Considering these questions, Siemens proposes a new learning theory – connectivism (Siemens, 2005). In the new connectivistic vision, within the digital age, learning cannot solely rely on personal experience any more but rather is derived as a competence from forming connections. Facing the speed and need of the technology enhanced society, the learner cannot experience every situation and borrows the experience from other people as their collected knowledge.

Connectivism strengthens the view that learning is motivated by connectivity, connecting experiences but also external information, residing in external, potentially interconnected, sources. Learning occurs in environments with shifting core elements – potentially outside of the learner's control – connecting specialized information sets and focusing on connections while connections which offer the learner to learn more are more important than the current state of knowledge. Connectivism fosters the understanding that decisions are based on changing foundations and stress the importance of the ability to differentiate between important and unimportant information. Siemens formulates in this regard eight principles of connectivism (Siemens, 2005)(p. 5):

1. *Learning and knowledge rests in diversity of opinions.*
2. *Learning is a process of connecting specialized nodes or information sources.*
3. *Learning may reside in non-human appliances.*
4. *Capacity to know more is more critical than what is currently known.*
5. *Nurturing and maintaining connections is needed to facilitate continual learning.*
6. *Ability to see connections between fields, ideas, and concepts is a core skill.*
7. *Currency (accurate, up-to-date knowledge) is the intent of all connectivist learning activities.*
8. *Decision-making is itself a learning process. Choosing what to learn and the meaning of incoming information is seen through the lens of a shifting reality. [...]*

Following the concept of connectivism a considerable amount of new publications started to emerge in the breach between technology enhanced learning and information systems, in line with the work of Downes *"An Introduction to Connective Knowledge"* (Downes, 2008). Connectivism appeals especially in situations of informal learning and technology enhanced working environments, where learning in constantly changing situations and requirements cannot be explained any more by traditional learning theories. Yet the new approach comes with limitations.

Connectivism received a variety of critiques, especially in more traditional publications, as the concept of connectivism is being weakly rooted in existing literature. Furthermore, the initial publication is incomplete and still weak regarding the criteria of learning, as categorized by the classification system of learning and instructional design by Schunk (Schunk, 1991). Additionally, Van Plon Verhagen criticizes that the concept of connectivism has a curriculum level focus, instead of being an instructional level theory, since it tackles "what is learned" and "why" and not "how learning takes place" (Van Plon Verhagen, 2006), as needed for a complete learning theory.

Independent of existing critiques, connectivism offers a new concept of learning and in this regard, presents how a blend of concepts of existing learning theories is possible. It is showing in parts similarities to other parts of the traditional theory, e.g. by acknowledging connected experiences similar to the view of constructivism.

In this vision, parts of different learning theories can be used for an instructional design with different viewing angles, if a conscious, consistent and sustainable conception is used for the long-term design of a specific learning environment. A technology enhanced environment is sure to have to be considered as an influence in the conception of learning but it could also act as a game changer using the technology directly for learning and furthermore for testing, implementing a well-designed process as a middle way to the current learning theories. But what is needed for a technology enhanced and enabled learning environment?

Testing has to be part of a technology enhanced process of learning. Reflecting on the subject of teaching, Bransford, Franks, Vye and Sherwood summarize in a short manner *"wisdom can't be told"*. But the question is not if wisdom can't be told - as experience can be transferred by communication - but if a learner can then make use of it. Even when explicitly telling a learner *"a piece of knowledge"* and assessing the repetition of the delivered information, it doesn't prove to be guaranty that a learner then possesses the ability to apply the information in new situations. Only testing can reveal what is mastered yet.

In contrast to testing, Perkins arguments, *"learners commonly know far more in a passive sense than they ever muster in realistic contexts of application. Although "there* *for the quiz," the knowledge otherwise disappears into the attic of unused memories"* (Perkins, 1992). It needs a connection between what has to be learned, what is tested and what is still to be learned based on the testing. In this regard, repetition may not be the sole creator of knowledge but it enables to re-enter a feedback loop of learning and testing to extend what is known and strengthen what is mastered.

Following the ideas of Siemens and Drownes, what is learned is always connected to internal or external sources of information and has to be considered in a situational context and the context of connected information. Siemens concludes in this regard *"The pipe is more important than the content within the pipe"*. There is no proof for this assertion yet but the connection to information is increasingly important and can on the long-term proof to be more crucial than single information accessed through a connection in a specific situation. A technology enhanced environment therefore has to be aware of connected information, while the instructional design should be aware of the application "situation", reflecting the situations for which the learning is targeted and the technological environment in which the learning occurs.

A new middle way on the use of learning theories, working in a technology enhanced environment and implemented as a technology driven process for learning is possible but has to satisfy specific design criteria. Considering the view of connectivism and taking into account the previous considerations on learning, testing and the increasing importance of connectivity of information, a technology enhanced and connectivity aware process for learning should satisfy five design criteria:

Table 8: Essential criteria for a technology enhanced and connectivity-aware process of learning.

Criteria	Description
[What connections?]	Implement awareness for connected information and render the connections transparent to the learner.
[What situations?]	Consider and account for the context of the variety of target situations of the knowledge to learn.
[What is known?]	Assess what the learner knows and what the learner has mastered.
[What to learn?]	Implement a feedback mechanism between the assessment of knowledge and the learning of knowledge.
[How to learn?]	The process of learning has to be designed conscious of existing learning theories, consistent in their selection and composition and sustainable in the conception of the process.

5.5. Short Summary: Learning Theory

Table 9: Summary of "Learning Theory Considerations".

Aspect	Description
Instructional theory	Learning theories express a concept of learning and a theory of instruction. From the range of theories, most applied are behaviourism - fact oriented and bottom-up, and constructivism - generalising and top-down.
Learning through connections	Connectivism brakes with the classic learning theories and propose a network oriented theory, expressing the connectivity of knowledge to other knowledge as important and the access to knowledge across knowledge networks as crucial for learning.
New technology centred way of learning	A technology centred way of learning can be based on connectivism, valuing: What kind of connections link different knowledge elements? What is already known by the learner? How to select what to learn next?

6. Network Analysis

Networks – or graphs – are an important and omnipresent approach to structure and visualize aspects of the daily life, modelling concepts like processes, knowledge or networks of people and their relationships. Each graph represents a collection of objects, connected through edges, modelling relationships between objects. Different naming concepts exist for the components of graphs, across different disciplines – in the following objects are referred to as nodes and edges are referred to as relations. A graph can have different topologies, while the main types can be distinguished – in an abstract sense – into directed and undirected graphs. In a directed graph, relations are directed, allowing the transition from one node to another only in pre-defined directions, while a relation can have two directions. In contrast, in an undirected graph, relations have no specified direction. Figure 7 visualizes a simple graph with numbered nodes, where a) shows a directed and b) shows an undirected version of the same graph.

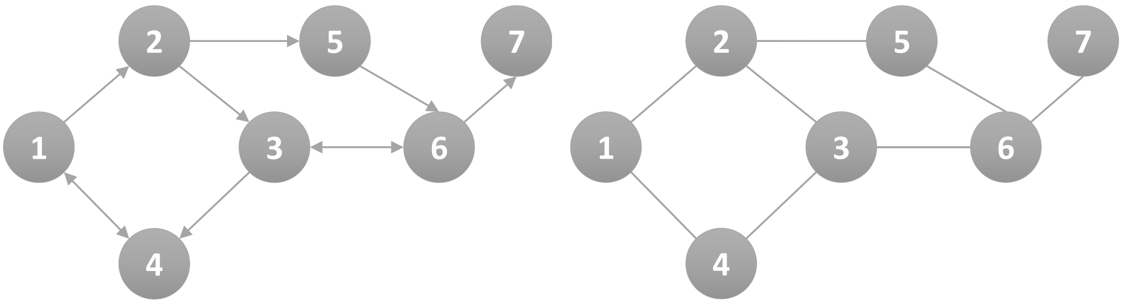


Figure 7: Simple graph, shown as a) a directed graph and b) an undirected graph.

Graphs are used in a variety of disciplines, as mathematics (graph theory), social science (social networks and social networks analysis), biology (food-networks, neural networks), computer and engineering sciences (information-based networks as the internet or telephone networks) for tasks of visualization and to reason based on the modelled content and relationships. “Network” and “graph” are used synonymous, while the specific use differs across disciplines (Barabási and Pósfai, 2016).

In the recent years, technology enhanced social networks turned into one prominent application for networks and their analysis. Indeed, the first steps within the field goes back to the 1930s, studying problems of sociometry and group dynamics, investigating dynamics in, essentially, (non-technological) social networks (Scott and Carrington, 2011). Social network analysis (SNA) investigates the relationship between people and how information flows within a network – based on context, relationships and communication patterns.

6.1. Centrality Measures

Networks, or “graphs” are used in variety of different disciplines to visualize, explain and reason on systems behaviour, and can be explained full or in parts by entities and their relations. Newman gathers that “*connections in a social network affect how people learn, form opinions, and gather news, as well as affecting other less obvious phenomena, such as the spread of disease*” (Newman, 2010). But to reason on and through networks, tools and methods are needed to capture the appearance and interaction of a network. As Newman captures here further, *“Unless we know something about the structure of these networks, we cannot hope to understand fully how the corresponding systems work.”*

Graphs can have different structures, being undirected or directed and enhanced with weights and contextual information. In contrast to purely structural measures, centrality measures strafe to attach an importance to the connectivity of nodes, which is conceptually near to the expression of connection importance in the frame of connectivism. A node could have inbound and outbound connections. For the calculation of the centrality of nodes, the graph, except for the Katz-centrality, can be assumed in a simplification as undirected. The equations for centrality in the next sections are in line with the formalization of (Newman, 2010).

6.1.1. Degree-Centrality

The most straightforward indicator for centrality is the degree centrality. Degree-centrality captures the number of edges connected to a vertex or node k.

For a node *i* with a neighbourhood of *n* direct connected nodes, the degree-centrality for the complete graph can be captured by an adjacency matrix as:

$$k_i = \sum_{j=1}^{n} A_{ij} \tag{4}$$

6.1.2. Eigenvector-Centrality

As an extension to the idea of the degree-centrality, the eigenvector-centrality also considers the centrality of other connected nodes. In this regard the eigenvector-centrality accounts for the centrality-based importance of other nodes. So even for neighbourhoods with a low number of connected nodes, a connected node will be judged as important if itself is connected to other important nodes.

Calculating the eigenvector-centrality is an iterative process, where each iteration improves the centrality till the graph converges. The Eigenvector-centrality x_i of all nodes i is calculated with:

$$x_i^t = \sum_j A_{ij}\, x_j \tag{5}$$

6.1.3. Katz-Centrality

The Katz-centrality extends the eigenvector-centrality in this way that it introduces parameters to modify the centrality for a given graph. The parameter α scales the centrality, while the parameter β defines a baseline, so that e.g. for any value of β greater than zero, the centrality for a node will be also always greater than zero and add additional importance to the surrounding nodes. The Katz-centrality is defined as:

$$x_i^t = \alpha \sum_j A_{ij}\, x_j + \beta \tag{6}$$

6.1.4. Betweenness-Centrality

The Betweenness-centrality measures how often a given node is part of shortest paths between nodes and models the flow of information within a network. As pointed out by (Newman, 2010), nodes with a high degree of betweennes may indicate a high control of a single node in a network over the flow. For a node i on the shortest path between s and t, while n_{st}^i is 1 if the node is on the shortest path, the betweennes is given by:

$$x_i = \sum_{st} n_{st}^i \tag{7}$$

6.1.5. Closeness-Centrality

For measuring how “close” a node is to other nodes in the network, the closeness-centrality measures for a node *i* the mean across all shortest paths to all other nodes *j* within the graph, with:

$$\ell_i = \frac{1}{n} \sum_j d_{ij} \tag{8}$$

6.2. From Social Networks to Concept Graphs and Domain Ontologies

Social network analysis investigates the relationship between people and how information flows within networks of people – based on their context, relationship, and communication. Taking people into account as bearer of knowledge, the interaction within social networks can be generalized as an interaction of knowledge holders.

Abstracting the concept, each person in the network possesses a different set of knowledge that could be broken down into a set of related sub-knowledge areas. In this regard graphs are suitable vessels to model the knowledge of people and how these knowledge areas are related to each other. In other words, graphs can well represent experts' knowledge that is crucial in knowledge intense occupations to efficiently support applications of learning, training and active knowledge transfer. The overlay model, addressed in Section 4.3, is an example for a graph of knowledge, modelling knowledge on a general level, where the knowledge of a virtual expert on a domain is modelled and on a personal level, where the knowledge of a learner is projected on and described by the expert's modelled structure.

Similar to the spatial relationships between individuals, the relation of knowledge can have a quality. Specially – to enable a sustainable modelling and management of expert's knowledge – requirements, concerning the level of complexity and meta-structure that define the possible compositions of the different knowledge elements, have to be defined. A prominent example from the field of biology are here taxonomies of animals (Huxley and others, 1940), modelling a hierarchy of animals to express a classification of animals into different species and sub-species, where the quality or semantic of the relations in the graph is "specialization". A more complex graph-based methodology to address concepts and their relations with different semantics, expressing individual concepts and qualities of connections as elements of the graph, are ontologies (Hebeler et al., 2009).

It is no coincident that Google named its semantical enhanced, graph-based search-engine knowledge graph to enhance search queries and the query result visualization, when introducing it in 2012 (Singhal, 2012). Visually the knowledge graph prepares additional contextual information to a specific search result, using texts images and links to further explore a search result. This information can be used by users and machines to differentiate, detail and resolve – comparable to the approach of (Hoffart et al., 2014) –

ambiguous search terms as Casablanca, which could name different entities as a city, a film or a restaurant. For each search, the knowledge network in the background will locate the search term within a semantic enhanced concept graph and extracts possible contexts based on connected concepts and attached features or tags. A similar approach is used for Microsoft’s Satori knowledge graph solution, to enhance search results of the Bing search engine and empower the extended feedback of the Cortana dialog system and other integrated services.

In the context of a user focused concept graph, personal, learned knowledge can be considered in the context of other knowledge and provide additional insights. Insights, not only on the knowledge which were already mastered but also on knowledge which is related and how it provides a context to the current state of knowledge. Social networks analysis can here provide measures for graphs of personal knowledge and how they are structurally related. Based on this measures the model of each learner can be extended with qualities or contexts. E.g. how central a specific concept is or how well it enables access to new knowledge areas within the graph.

6.3. Short Summary: Network Analysis

Table 10: Summary of “Network Analysis”.

Aspect	Description
Network analysis as an enabler for learning	Network analysis methods can analyse "social" networks of people but also information networks and as such yield the potential to analyse concept or knowledge networks for learning.
Suitability of network centrality measures	Centrality measures transform the structural aspects of networks into numbers as connectivity, shortest paths and clusters. As such they can act as a blueprint to utilize the structure to reason on the importance of concepts in networks for assessment (in line with the view of connectivism).
Semantic information as a source	Given a semantic model of the domain to learn, the aspects of the modelled semantic could be an additional source of information. Together with the structural information of a knowledge network, a design of the importance of concepts for learning may become possible.

7. Research Questions and Research Planning

This research will build on two main pillars: 1) the modelling and development of a concept importance measure and 2) the algorithmic and architectural implementation of a knowledge assessment solution applying the concept importance measure. Discovery and reasoning will be of an explorative nature throughout the whole research. Research questions (which are detailed in Section 1.2) and the planned research methodology are described in the flowing.

- **Research question 1:** How can **the semantic model of a learning domain** be utilised to identify **which knowledge area**(s) **is** (are) **of high importance for learning** in comparison to other knowledge (concepts) within the model of the domain? (Methodology: [Modelling][Experiment])
- **Research question 2:** How can a **measure, quantifying the importance of concepts** in a semantic model **be utilized, integrated and implemented** in an online assessment solution? (Methodology: [Build])

7.1. Research Methodology

Within the doctoral school of business informatics, approaches of the disciplines of social science and computer science are applicable and are used in different situations in dependency of the background of a specific research and the need of the application. The area of the planned thesis is set into the broader context of labour market considering aspects as the learning and mastering of knowledge for new job roles or positions, or the learning and the acquisition of new knowledge in informal settings of learning.

Yet the methodologies which will be used to address the problem are technology driven – as the field of user modelling and adaptive systems – and insights, derived from the use of domain knowledge, are represented semantically and structural as a domain ontology. In this regard, a computer science related exploratory approach to the research methodology is more suitable to address the expected findings and to integrate the derived insights into the existing stream of research.

Following Amaral et al. (Amaral et al., 2011), research methodologies in the field of computer science can be divided into five methodologies:

- ***Formal*** **–** proving facts about algorithms and systems.

- ***Experimental*** **–** evaluating new solutions for problems.
- ***Build*** **–** building a physical or software artefact.
- ***Process*** – understanding processes in computer sciences which are used to accomplish specific tasks.
- ***Model*** **–** defining an abstract model for a real system or application.

Based on this overview the next sections will shed a more detailed light on the collected methodologies, based on the summary of Amaral et al. (Amaral et al., 2011).

7.1.1. Formal Methodology

A formal methodology is used to prove facts about algorithms and systems, while focusing on: a formal specification of a software component to enable an automatic verification of an implementation of the component; on the time or space complexity of a specific algorithm or the quality or correctness of a solution, generated by the same algorithm.

A formal methodology is a foremost used in theoretical computer science. It is formal and mathematical and is concerned with the fundamental limitations of modelling and abstraction. Furthermore, a formal methodology can be used to derive statements about the computability and complexity of algorithms.

7.1.2. Experimental Methodology

Experimental methodologies tend to be split into two phases – an exploratory phase in which measures are identified which help to identify the relevant questions of a research and a second evaluation phase which attempts to answer the identified questions.

In exploratory phase within the frame of experimentation, the researcher has to answer the questions: what question is the experiment meant to answer, what variables affect the result and what is outside of the control of the researcher, what measures will be taken into account for the variance of the used variables and what results are statistically significant. The reporting phase after the conduction of the experiment summarizes what has been learned from the experiment, together with a careful selection of different representations of the results to underscore the specific points which were tackled. The elaboration of the results is followed by a discussion to provide deeper insights into the collected and analysed data or “explain” the gathered results.

7.1.3. Build Methodology

This methodology is fundamentally about the building of a software system, documenting the planning, the composition, and the final process of building. Aspects to consider in the process of building are:

- Software Design – how to split the solution into components for implementation and testing, how to handle the design complexity.
- Component Reuse – are and what components are available and what are the implications of their use.
- Programming Language – carefully consider the profit of knowing higher level concepts of one language against the suitability of a different language to learn, which may be more adequate to build the specific system in terms of run-time speed, expressiveness, and reliability.
- Software Testing – testing modules throughout the development time and consider an automated approach to testing.

The final "built" should be compared against existing solutions, if applicable, or additional claims as speed, space requirements, and other measurable factors should be backed by suitable measures and statistics.

7.1.4. Process Methodology

A process methodology is human focused in terms of tackling how humans construct and how they interact with software systems.

Focusing on software systems, the methodology investigates how systems are designed and built initially and evolutionary. Research in this line of work may identify design and implementation patterns and strategies, trends, and the codification of best-practices.

Interfaces, their construction, and the human interaction is investigated in the field of Human Computer Interaction (HCI) and involves beside the design further considerations about the consciousness of users, design best practices and interaction patterns.

In the context of cognitive modelling, human cognitive processes are hypothesised and mirrored into a practical solution which is either mirroring the processes or implemented to support specific cognitive processes. In all cases, empirical data will be collected and analysed to evaluate theorized trends and to investigate if the specific solution can be abstracted and generalized to other solutions.

7.1.5. Model Methodology

Amaral et al. (Amaral et al., 2011) defines modelling as *"[...] the purposeful abstraction of a real or a planned system with the objective of reducing it to a limited, but representative, set of components and interactions that allow the qualitative and quantitative description of its properties"*. A modelling approach is driven by the application for which it is planned and the targeted research for which the modelling is conducted and therefore can lead to multiple correct results.

In a scientific context, a model is built to capture and account for important aspects of a target system at the cost of less important aspects. How to decide which aspects are important and which are not is part of the modelling strategy. Modelling is considered as an evolving process which is focussed on a selected sub-discipline of research.

8. A Concept Importance Measure for Domain Knowledge in the Context of Learning

In the context of the planned research work the task of designing a concept importance measures for semantically enhanced structured knowledge conforms to a modelling approach. It abstracts the existing conceptualisation of a knowledge structure to a representative measure and in this regards models the concept of relevance for the specific application of learning. To verify the presented model the methodology of experiment is suitable. In this regards the used methodology will be a hybrid of modelling and experimenting, with a stronger focus on the model and the question of how to utilize and reflect it in a practical computer science motivated solution. The utilization of the intended measure is, following the definition of the research methodology, fitting to the “build” methodology. The final implementation is a part of a greater software framework. In favour of the complexity of testing and the learning and assessment focus of this work, the evaluation of the implementation will be an integrated part of the evaluated field studies.

The goal of this thesis is twofold, as tackled in the detailed research questions in Section 1.2: First, to create a measure which reflects the importance of knowledge areas – or concepts – in a semantic model of a learning domain and second, to implement a solution which uses the concept importance measure for online assessment with the eventual goal of learning – as knowing the knowledge gaps, means knowing what to learn.

For these goals, the literature background work of training and education in the organisational context in Section 2 provides a narrative, addressing that training and finally learning has a context and a motivation, which will be reflected in the solutions which are supporting the process as e.g. a semantic model of the domain to learn, or an assessment solution, using the model to provide further learning support. The literature background of adaptive testing in Section 3, extends these ideas by going deeper into the potentials of adaptive assessment, which will support the latter utilization of a concept importance measure. Section 5, then considers the common learning theories, which can frame and base the idea of a concept importance measure but also inform about the important aspects for the specific design – especially the theory of connectivism.

To finally address and design the importance of concepts in a semantic model of a learning domain, the introduced centrality measures in Section 6 – in line with the perspective of the learning theory of connectivism – can be a rational starting point. For persons holding specific knowledge in a network of persons – in contrast to a direct semantic domain model of the knowledge within a domain – Joksimović notes about centrality measures that (Joksimović et al., 2016)(p. 2):

> “Despite the prevailing, and largely unchallenged, understanding that occupying a higher social centrality leads to a higher academic performance, research findings are inconclusive about which centrality measure (or combination of measures) is the most significant predictor of academic achievement”.

In this regard a singular “one measure” solution of network analysis may yield limited results for isolating educational potentials based on a domain ontology (semantic domain model). Even though Joksimović’s statement addresses the connection between social connections and their impact on academic performance, rather than making a statement about the topology of a contextualized map of knowledge, it gives a first indication that to reason on a domain ontology may need a more complex approach, which integrates a mixture of measures and inputs.

Using, and blending different information about a domain model and being aware of the context in which parts of a network are relevant, may provide a better basement for a concept importance measure than using centrality measures. A view which is backed by the further survey of Joksimović, gathering, concerning to the connection between centrality and academic performance, that “*several recent studies have revealed somewhat contradictory results, indicating that the predictive power of social centrality measures highly depends on the context that frames students’ interactions*”– and in this regard, looking closer on the context of concepts in a domain ontology model for learning and the semantic of the model itself, can provide the right inputs for a concept importance measure.

Reflecting on the final implementation side – in the perspective of user modelling and adaptive systems of Section 4, a domain model can be used to extend or – similar to an overlay-model – replace the user model and add additional context. When a user model is utilized to adapt the learning experience to the progress of the learner, as for an online assessment and learning solution, a system moves through several phases to derive an adaption: extracting information, filtering, aggregation, selecting fitting features and concepts, and, finally, deriving a decision on how to adapt. To make a flexible use of different sources of information for the adaptation process in this work, an approach is needed to address the information which can be gathered based on the domain model.

The literature analysis sheds light on a comprehensive fundament of topics in the context of user modelling and adaption, learning and network theories, which can support the creation of a concept importance measure on a theoretical base and inform the implementation side of assessment and learning [6]. Yet the analysis has also shown that no current single theory provides a well-developed, "of-the-shelf" starting point to rate the importance of concepts for learning. While this underlines the value of this work in terms of "closing the gap", it also contributes to the challenge of defining a measure. Table 11 gathers the lessons learned from the literature study, and highlights their use, relevance and impact for defining and finally utilizing a concept importance measure, in line with the summaries of the findings of the literature sections, gathered in Table 3, Table 5, Table 6, Table 9 and Table 10.

Table 11: Combining the background and lessons learned of the literature study to a foundation for the concept importance measure.

Focus	**Description**	
Creating a concept importance measure	***Training and Education in the Organisational Context***	**Continuous training need –** Training in non-formal education and higher education requires foremost a self-motivated flexible learning behaviour. Which can be supported by a concept importance measure for learning.
	Concepts of Adaptive Testing	**Need for adaptive testing –** To enable efficient, personalized learning, learners' knowledge gaps have to be identified individually and can be explained further by a measure.
	User Modelling	**Need for a semantic domain model –** A domain ontology model (semantic network) enables to express more complex domains and derive a better learning feedback.
	Learning Theory Considerations	**Learning through connections –** Connectivism propose a network oriented theory, expressing the connectivity of knowledge to other knowledge as important and the access to knowledge across knowledge networks as crucial for learning.
	Network Analysis	**Semantic information as a source –** Given a semantic model of the domain to learn, the aspects of the modelled semantic could be an additional source of information – together with the structural

[6] A user modelling and adaptive testing focused literature overview and statement-article addressing STUDIO, is published in (Weber et al., 2016)

		information.
Integrating a concept importance measure	***Training and Education in the Organisational Context***	**Need for flexible, adaptive solutions –** To enable an assessment and learning across different individual and organizational needs, flexible and adaptive systems for technology enhanced learning are needed.
	Concepts of Adaptive Testing	**Short assessment and learning cycles –** To close the distance between tutoring and the learning performance of a student or worker, short cycles of repeated assessment and learning are important.
	User Modelling	**Adapting to "what" –** Knowledge can be presented as a feature of the user or as a model of the domain to master.
	Learning Theory Considerations	**New technology centred way of learning –** A technology centred way of learning can be based on connectivism, valuing: What kind of connections link different knowledge elements? What is already known by the learner? How to select what to learn next?
	Network Analysis	**Suitability of network centrality measures –** Centrality measures transform the structural aspects of networks into numbers. As such they can act as a blueprint to utilize the structure to reason on the importance of concepts for assessment.

In the first phase of the following research a concept importance measure has to be designed for semantically enhanced and structured knowledge. This task conforms to the modelling approach. It abstracts the existing conceptualisation of a knowledge structure to a representative measure and in this regard, models the concept of relevance or "importance" for the specific application of learning. To verify the presented model, the methodology of experiment is suitable. In this regard the used methodology will be a hybrid of modelling and experimenting, with a stronger focus on the model and the question of how to utilize and reflect it in a practical computer science motivated solution.

To accommodate the model and enable the experiment, a software prototype and a data extraction framework is needed. Both will be implemented and will conform in an applied perspective with the build methodology but parts of the build methodology will be omitted to keep the focus on the modelling task and its relevance to technology enhanced education, as well as conducting experiments, empowered by the built solution. The stages of both researching and implementing the Concept Importance measure are gathered in the following sections.

8.1. Application System Description and Exploration

The realization of this research will be embedded into a technology enhanced learning system, which also will support the experiments accompanying the sub-chapters. The experiments will be conducted in a blended learning environment, which supports the seminar work and studies of bachelor students. The results and continuous system feedback will support the development of the new approach for measuring the importance of concepts.

For the development and implementation of the concept importance measure, the well elaborated STUDIO system for adaptive assessment and learning is selected (Vas, 2016). STUDIO integrates a sound, comprehensive, semantically enriched knowledge structure, which fits to the requirements of the first parts of the literature study, summarized in Table 3, Table 5 and Table 7. The system models the domain related knowledge as an ontology, offering the needed structure and semantics to conduct the research. It will act as a test-bed and a source of domain knowledge. Furthermore, the system provides feedback in the form of reporting, visually exploring the assessment results, the general progress and fitting learning material. The planned algorithmic extension, implementing a concept-importance aware assessment, will be integrated into the STUDIO system. The system will further host the latter experiments and data collection.

To evaluate available data sources and potential factors for the concept importance measure, the system functions and data collection potentials will be explored. The following sections will picture the idea of the STUDIO solution, the existing implemented assessment and the used algorithms, the data organisation, and the scientific relevance of the system.

8.1.1. STUDIO – an Ontology-based System for Assessment and Learning in the Context of Blended Learning

The basic concept of STUDIO is to model the focused education as an interrelated knowledge structure, which divides the education into sub-areas and knowledge items to know. The managed structure formalizes the relation between knowledge areas as a learning context and models the requirements to master specific parts of the education.

This structure is used to create and support knowledge tests for students. Through this combination of assessment and knowledge structure, the student gains the freedom to explore not only single knowledge items but the education in the context of related knowledge areas, while the embedded requirements are used to map the modelled knowledge against the expected educational outcome.

The assessment-system is designed to be accompanied by phases of learning within the system, where the student gets access to learning material, based on and supported by the test feedback. This combined approach offers a unique self-assessment to the students, where the backing knowledge context is used to adapt the assessment in dependency of the test performance of the student.

A major strength of STUDIO is the domain ontology and the domain tailoring. Learners can use STUDIO to examine their knowledge in a process of self-assessment and self-learning. As pictured in Figure 8, to set-up the learning environment, the tutor initiates the assessment domain by selecting those knowledge elements from the domain ontology which express the current domain of learning best. These selected elements are then – in an automated process – complemented by knowledge elements from the ontology. This process works based on the ontology structure and completes the desired sub-domain. The finished sub-domain is then used for assessment and learning within the system.

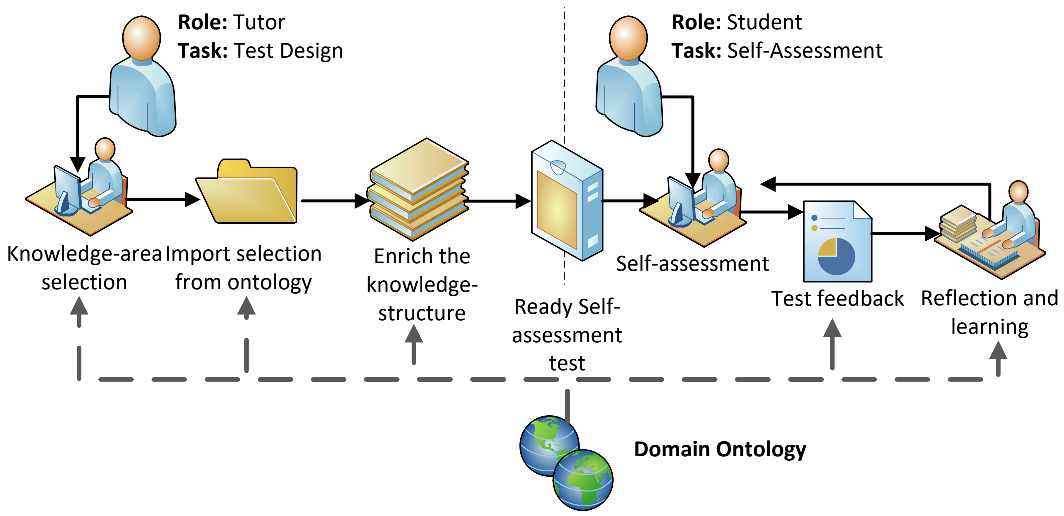


Figure 8: The tailoring, assessment, and reflection cycle of STUDIO.

STUDIO is divided into three main components – the Domain Ontology, the Knowledge Repository, and the Knowledge Retrieval Engine. The Knowledge Retrieval Engine interprets the Domain Ontology, in terms of structure and semantic and adapts the adaptive self-assessment test with questions from the Knowledge Repository to the user and the Domain Ontology. Based on the outcomes of the self-assessment, the Knowledge Engine tailors learning material from the Knowledge Repository to the learning need of the user. The learning need is set into a context by the structure of the domain knowledge within the Domain Ontology. An overview of the architecture is given in Figure 9, while the components are described in detail in the following sections.

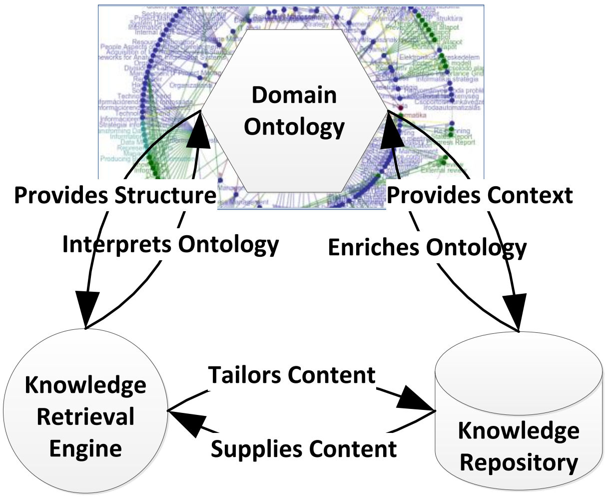


Figure 9: The Three Core Components of STUDIO and their Interaction.

Before any regular examination students may use STUDIO to assess their knowledge on their own. It is the tutor’s responsibility to set the course of self-assessment test in STUDIO system by selecting knowledge areas and sub-knowledge areas which are relevant for the target education from the domain ontology. Then the frame will be automatically completed with elements from the ontology which detail the selected knowledge areas and with elements which are modelled as required for the already selected elements.

As the system stores assessment questions for each knowledge element, STUDIO will – based on the selection – automatically prepare an assessment test, grounded on the defined selection and the domain ontology. The resulting knowledge-test is afterwards accessible as a self-assessment test for the student, who explores the backed knowledge structure, which pictures the expected learning outcome, in cycles of testing, reflection and learning. The process of test definition and assessment is shown in Figure 8.

The knowledge within a domain can be represented in different technical and conceptual ways. Solutions can be differentiated into user model focused – modelling knowledge in the context of the user/learner – and into domain model focused – focusing on abstracting a world model to see knowledge as an extract of the real world. STUDIO makes use of the later concept and models the education as an interrelated knowledge structure, the domain ontology.

8.1.2. The STUDIO Assessment and Learning Cycle

STUDIO follows a straightforward, yet powerful cyclic approach for learning, as shown in Figure 10:

1) **Assessment:** Based on the domain ontology the learner receives questions for the domain in an adaptive self-assessment test. While the learner receives questions, his or her answers determine: what question from which knowledge area to receive next, and how long the assessment will last. In this adaptive process the system follows the network structure of the knowledge elements within the domain ontology and explores, based on the user interaction and the assessment-performance, the network of domain knowledge.

2) **Reflection:** If the user fails considerably often, the assessment will stop and the learner will see a visualization of the domain in an interactive learning interface. Within the interface the domain visualization will colour-code the achieved assessment result with the main colours of 1) green, for correctly answered and accepted knowledge elements 2) red, for incorrectly answered elements, and 3) grey, for elements which are not explored yet. A 4th) colour is orange and highlights elements which were answered correctly but haven’t had “enough” dependent knowledge elements answered correctly.

3) **Learning:** Based on the assessment results, the learner receives access to learning material for each assessed knowledge element. The learning material elaborates the background knowledge, needed to master each knowledge element and its questions. The access is granted for all knowledge elements which were part of the assessment. This way the learning is tailored to the learner, based on the assessment and the domain ontology.

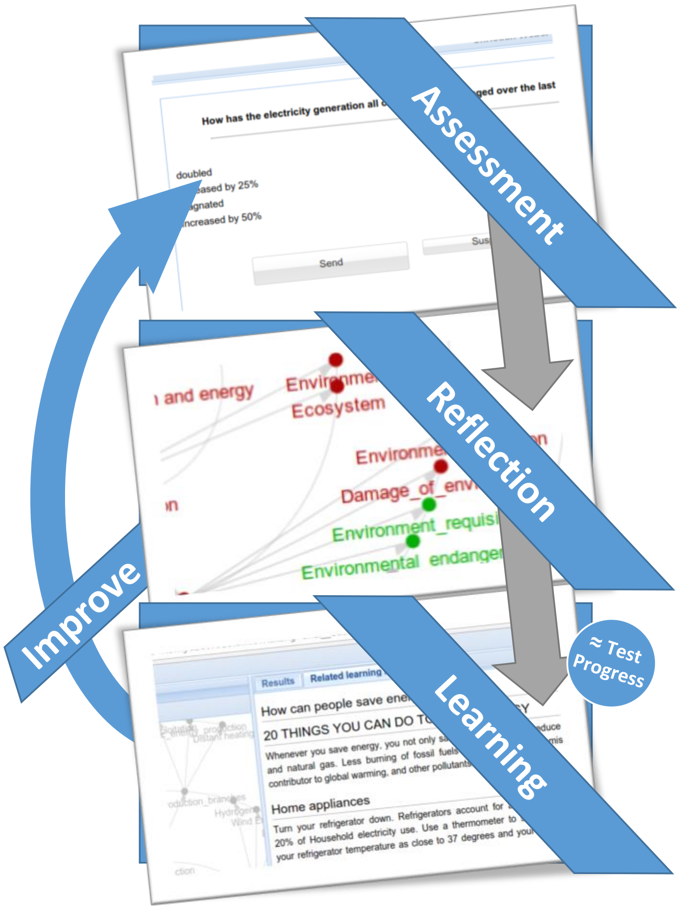


Figure 10: The STUDIO Adaption and Knowledge Discovery Cycle.

This process is designed to be accessed repeatedly by the learner to 1) frame the his/her knowledge, 2) then offer the right selection of learning material, and 3) continuously explore and unlock more material and "learn" the domain and its comprising knowledge elements. Through this cycle and its guided "learner domain ontology interaction" the system offers an adaptive tailoring of the domain's complexity to the current understanding of the learner. The result is an adaptive assessment which "discovers" the learner's knowledge.

8.1.3. The Domain Ontology

The STUDIO system is based on an educational ontology, presented and explained in detail by Vas in (Vas, 2007). Domain ontology is a frequently used term in the field of semantic technologies and underlines the storage and conceptualization of domain knowledge. Domain ontologies are applied in a range of different projects and solutions (Dahab et al., 2008; Missikoff et al., 2002; Wu and Hsu, 2002) and could address a variety of domains with different characteristics in their creation, structure and granularity, depending on the aim and the modelling person (Gavrilova and Leshcheva, 2015).

A specialization in terms of the domain is the educational domain ontology which is a domain ontology adapted to the area and concepts of education. It could target to model different frames of the education as the curriculum or more granular aspects relevant especially for the task of learning and course creation (Nodenot et al., 2004; Psyché et al., 2005; Sosnovsky and Gavrilova, 2006), or it describe the design, use and retrieval of learning materials till creating courses (Bouzeghoub et al., 2003), as well as directly the learner within the education (Chen and Mizoguchi, 2004) as a variant of a user model.

Within the area of educational ontologies, domain ontologies tend to model specific details of the education, in an attempt to model the selected field as complete as possible. This enables a comprehensive view on the field but it can come at the cost of generality, with the risk to be inflexible to handle changes. Other concepts model the education across different ontologies, matching concepts like the learner, the education and the course description, introducing a broad horizon but with additional overhead to combine modelled insights and reason on new instances.

The appeal of the STUDIO educational ontology is the number and focus of the main classes and their relationships between each other. The knowledge to learn is the main connecting concept in the core of education. It enables a great flexibility to be resourceful for different education related questions. An example is here the business process management application of PROKEX, which maps process requirements against knowledge areas to create assessment test which reflect the requirements of attached processes (Neusch and Gábor, 2014).

An important factor in learning is the distance between the progress expectation of the tutor and the real learning performance of the student. Here a short cycle of repeated assessment and learning is a major factor for a better personal learning performance based on feedback (Roediger, 2008). This aspect directly benefits from the focus on knowledge-areas as the main exchange concept between students and tutors. Furthermore, the close connections between learners and educators via direct tutoring is one major enabler for computer aided systems (Fletcher, 2003), and each step towards a more direct interaction through a focused support is an additional improvement.

The class structure of the STUDIO domain ontology fuses the idea of interrelated knowledge with a model of the basic types of educational concepts, as involved in situations of individual learning. Figure 11 visualizes the class concepts as different types of knowledge elements and different relation types, which are used to model the dependencies between different knowledge elements in the process of learning.

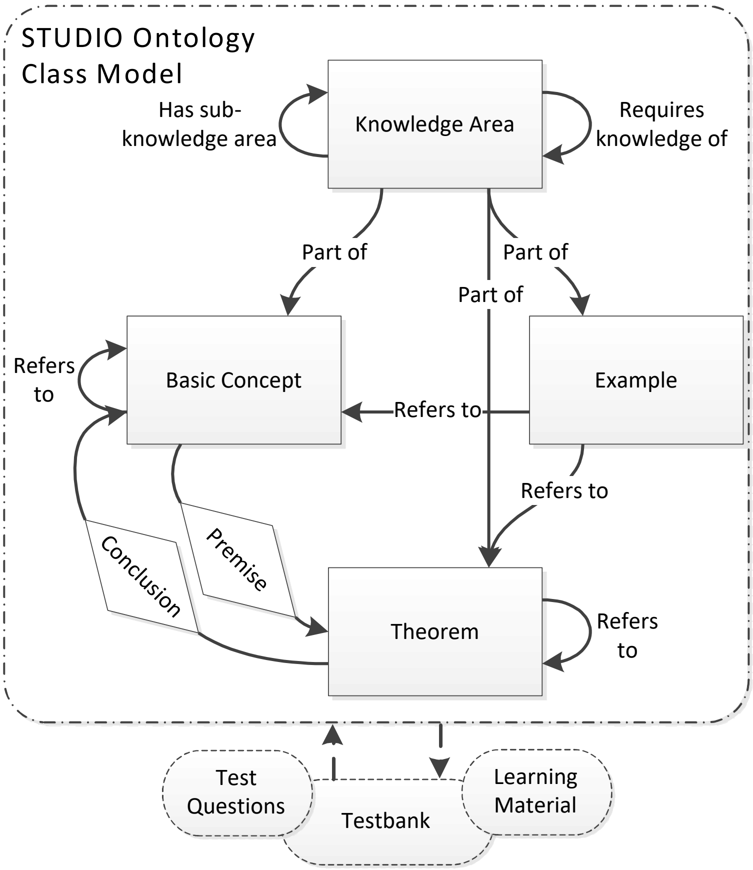


Figure 11: Model of the Educational Ontology (Weber and Vas, 2015).

The "Knowledge Area" is the super-class and core-concept of the ontology. The ontology defines two qualities of main relations between knowledge areas: Knowledge areas could be sub-knowledge areas of other knowledge with the "has_sub-knowledge_area" relation, or be a requirement for other knowledge areas with the "requires_knowledge_of" relation. A knowledge area may relate to multiple connected knowledge areas, as requirements or sub-areas. The "requires_knowledge_of" relation defines that a node is required to complete the knowledge of a parent knowledge area. This strict concept models a requirement dependency between fields of knowledge in education, while the "has_sub-knowledge_area" relation models a hierarchy. The earlier "requires_knowledge_of" relation yields the potential to assess perquisites of learning, analogue to the basic idea of perquisites within knowledge spaces, developed by Falmagne (Falmagne et al., 1990).

Education is a structured process which can split the knowledge to learn into different sub-aspects of learning. Knowledge areas in the ontology are extended by another specialized sub-layer of knowledge element types in order to more effectively support the modelling of education and testing requirements. Figure 11 visualizes the sub-elements and their relations. By splitting knowledge into sub-concepts for the modelling and assessment, the coherence and correlation of self-assessment questions could be expressed more efficiently (with the potential of a more detailed learning feedback).

The knowledge element type of “Theorem” expresses in a condensed and structured way the fundamental insights within knowledge areas. They connect and explain the “Basic concepts” of the modelled knowledge and set them in relation to the environment of learning with “Examples”. Multiple theorems can be “part_of” a Knowledge area. Each theorem may define multiple Basic Concepts as a “premise” or “conclusion”, to structure how the elements of the knowledge area are related. Examples enhance this parts as a strong anchor for practical self-assessment questions and they “refer_to” theorems and basic concepts as a “part_of” one or more knowledge areas.

8.1.4. The Knowledge Repository

In the frame of Knowledge-Based Systems (KBS), the system architecture is composed of: an inference engine, implementing a core set of inference rules for the scope of reasoning; a knowledge base, storing rules and facts; and a user interface. In the scope of web-based and web-technology driven systems, e.g. in adaptive hypermedia systems (AHS), the knowledge base is seen as a more simplified knowledge repository (Leondes, 2010), storing only facts and descriptive content.

The two overall goals of the STUDIO system are 1) to give a framework with what the domain ontology can be built, and 2) to implement a dynamic self-assessment, based on the domain structure. The repository contains questions and learning materials which are associated with knowledge elements of the ontology. Different types and implementations of questions and learning material are possible. There may be one or more questions associated with every ontology class instance.

In STUDIO, multiple choice questions (MCQ) are used, with four possible answers, whereof always only one is correct. Furthermore, the knowledge repository stores the learning materials in the form of mixed-media Wiki pages (hypertext or hypermedia). Each learning material comprises all the factual knowledge needed to pass the associated knowledge element within the ontology. These hypermedia objects are providing support for not just textual content but also for multimedia content. Through the integrated learning material and the dynamic self-assessment, STUDIO provides a unique facility for self-improvement and learning to the users.

8.1.5. The Knowledge Retrieval Engine

One part of the flexibility of STUDIO rests on the use of “Concept Groups”. Concept groups are tailored sets that contain only a certain, selected part of the ontology. Concept Groups enable to adapt in a dynamic the domain context of learning by including only those elements of the ontology which best fit to the learning needs. The knowledge retrieval engine can be considered as a specialized inference engine which implements testing algorithms and evaluation strategies. Testing algorithms use the tailored concept groups and the semantic of the ontology to determine in which order the knowledge elements should be assessed from the user.

8.1.6. The Drill-down Test- and Evaluation Algorithm

The creation of each new self-assessment test begins with the interaction of the tutor with the system. The rational of this initial stage is to base the test on the expertise of a real-world expert in terms of selection and based on a tutor driven estimation of a fitting granularity between the knowledge area levels for the learner. Within the STUDIO frame this expert driven selection is completed by a process driven extraction of knowledge areas, relevant for the organisational processes.

To create a regular self-assessment test through STUDIO, the tutor has to select the relevant knowledge areas and connect them to concept groups which together are creating a tree of groups. Within the system this happens based on the support of the process model. The resulting tree pictures then a sub-ontology of the main domain ontology. For each concept group the system will import related knowledge elements from the domain ontology and complete the test frame. This extraction step completes the frame with all knowledge areas and relations from the domain ontology which are necessary to connect the already selected knowledge areas, based on the concept groups.

The output of the extraction is a cached directed graph representation of the modelled assessment domain. By definition, the top element of the topmost concept group will be set as the start-element and root of the tree shaped graph. The start-element or start-concept acts as a fix-point from which the top-down assessment algorithm will start and to which the bottom-up algorithm will explore the graph to. As such the start-element is the centre from where the imported knowledge structure is interpreted for testing.

In the next stage the selected assessment algorithm will start and move through the knowledge structure, based on the internal navigation rules while online administering the questions connected to each knowledge area the algorithm selects. To explore the knowledge structure for both internal test algorithms, the system makes use of one central assumption, depicted in Table 12:

Table 12:Necessary assumption for traversing the knowledge structure for assessment.

Assumption	Description
Ordering	All knowledge areas relate to “part_of” and “requires_knowledge_of” relations. Resulting, every path, starting with a start-element, will develop on average from general concepts to detailed concepts. As an implication, any methodology to select concept groups for the test definition has to be designed in a way that it selects and orders following concept groups accordingly to also lead from general at the top to more detailed groups at the bottom of the structure.

Each created assessment-test structure defines a sub-ontology of the source domain ontology. This extracted blueprint of the test starts with the start-element of the highest defined concept group. To load and complete the knowledge structure for the self-assessment, the system follows in a cycle the following steps to load the structure, based on the test definition described in the previous section:

1. **Load:** Knowledge-elements are connected through relations. Each relation type between two knowledge-elements has one unique direction, fixing the extracted tree as a directed graph. The system will load all relations between two knowledge-elements, which start with the start-element and ends on another knowledge-element. This creates a two-level structure where the start-node is a parent-element and all related, loaded elements are stored as child-elements.
2. **Select:** The algorithm then successively selects each child-element of the start-element and defines it as a start-element in its own process.
3. **Stop:** When no more knowledge-elements for a parent-element could be loaded, the sub-process stops.

4. **Repeat:** The system then repeats the first steps till all knowledge-elements are loaded into the created tree-structure. When all sub-processes have stopped, the knowledge structure has finished loading.

The overall process flow of the extraction is shown in Figure 12.

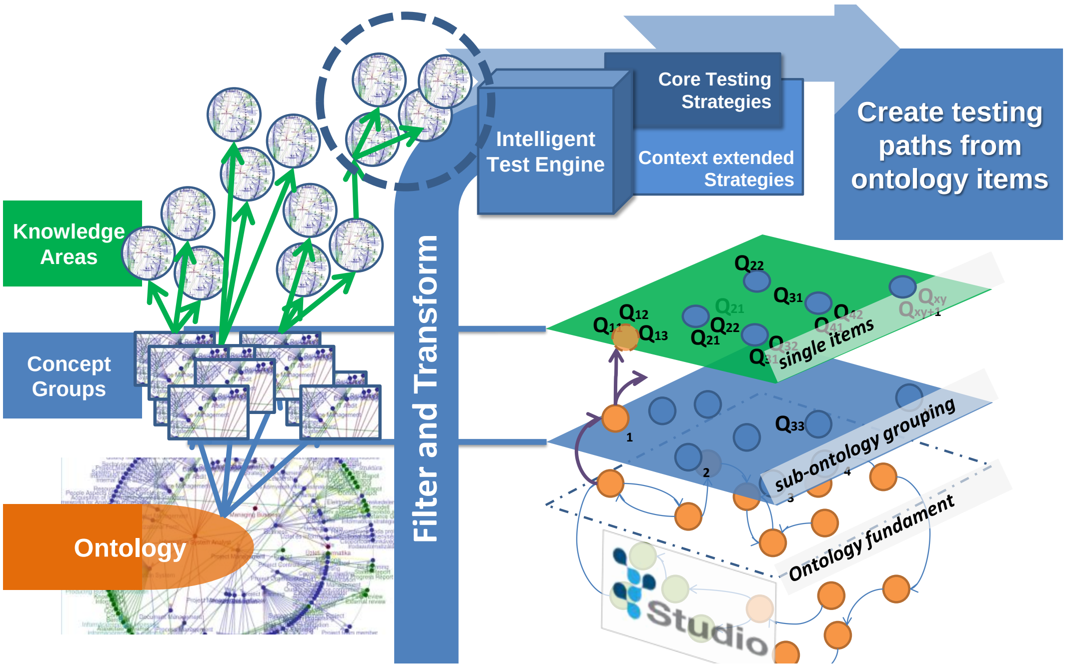


Figure 12: Overall process flow of the concept extraction from domain ontology, based on concept groups.

To ensure a working top-down assessment, the system has to fulfil the assumption, that missing knowledge on an early stage of the knowledge structure hierarchy disables the learner to answer questions for knowledge areas on more detailed levels. This assumption is summarized in Table 13 and follows the taken process view on the assessment that the knowledge about the high-level processes, which correspond to higher levels in the knowledge structure based on the process extraction, is needed to perform acceptable on job roles and tasks, modelled deeper in the knowledge structure.

Table 13: Necessary assumption for traversing the knowledge structure in a top-down setup for assessment.

Assumption	Description
Top-down knowledge dependency	If a test-taker fails on more general concepts the system will assume that he or she will also fail on more detailed concepts. Further, when a sufficient number of detailing concepts failed, the parent knowledge will not be necessarily covered and will be derived as failed, too.

Based on the first assumption, defined in the previous section, the deeper a knowledge-element is within the tree, the more detailed is the concept of the element, creating a hierarchy going from general concepts to specialized concepts while moving down the tree structure. Out of the directed one-directional nature of the defined relations, this loading-process provides a directed tree of the knowledge structure to the test algorithm. An example visualization of tree is shown in Figure 13.

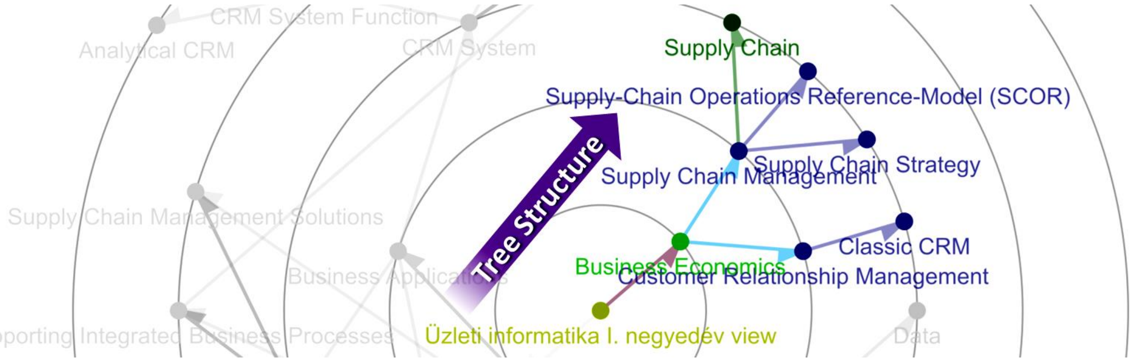


Figure 13: Excerpt from the sub-ontology visualization, showing the desired tree structure.

Adapting to the tree shaped knowledge structure, the top-down assessment triggers the following steps to assess the represented knowledge, based on the questions connected to each knowledge-element:

1. Beginning with the start-element, the test algorithm activates the child knowledge-areas of the start element.

2. The top-down algorithm then selects the first child-knowledge area and extracts a random question out of the question storage connected to the knowledge-element and assesses the answer from the learner through the testing interface.

3. When the learner fails the assessed question, the algorithm marks the element as failed, else it is marked as passed.

4. The system then selects the next non-failed knowledge area, accessible directly or through passed nodes from the start-element, promotes it as a parent node and queries a random question from it to repeat the process afterwards.

Following the test algorithm in a cycle, the system dives top-down the knowledge structure and continuously triggers more questions depending on the learner's answers and on the designed and extracted model of the relevant education. In this regards the STUDIO system adapts the test on the fly to the performance of the learner, captured through the feedback on the test questions. This triggers to follow or not follow more knowledge elements on the same branch of the knowledge tree. As this assessment is working on the structure, created based on the extracted organisational processes, the system adapts the testing of the domain knowledge to the assessed knowledge of the learner, providing as such an adaptive solution for the self-assessment.

This process of mapping the learners performance to the conceptualisation of the domain ontology, resembles the concept of overlay based student modelling (Chrysafiadi and Virvou, 2013). While the learner continues to use the self-assessment through phases of testing and learning, to evaluate the personal knowledge, he or she will dive down further into the knowledge structure and continuously explore more and more detailed areas of the target education.

As the approach is sufficient to assess the alignment of learners to the defined organisational processes, this approach is limited in regards to the speed of exploration of single knowledge elements. The algorithm for assessment stops following a branch of the knowledge structure as soon as the learner fails on a knowledge-element. Especially failing on top level elements near to the original start-element, results in the early exclusion of a complete sub-areas of the overall knowledge domain in the testing process. This can lead to iterations where after a short series of failed answers the assessment will stop on the first level of knowledge-elements, giving no specific feedback on learning and missing the opportunity to assess single task related knowledge, which may draw a picture of the current capabilities of a learner.

This correlates with the desired behaviour fields of testing as computerized adaptive testing (CAT) (Linacre, 2000; Welch and Frick, 1993), but it neglects the possibility that learners could yield knowledge of more detailed concepts, while not having fully mastered general concepts or yet miss the assets to understand the questions of higher level concepts.

8.2. Preparing Data Collection: Integration of an Event-tracking Framework

To enable a latter integration of the concept importance measure and to track and enable the planned experiments within the STUDIO system, a stocktaking and survey of the already tracked and the potentially trackable assessment and domain ontology related variables within STUDIO is needed. The STUDIO system integrates a statistical module and collects the choices and results collected and presented by the assessment and learning module. However, a solution to capture more granular events or component spanning or component independent events is yet missing. To improve the understanding of the assessment and learning process in STUDIO, a new data-collection framework has to be built and integrated. To realize a component-independent tracking of new variables, the system extension will be designed in a way that variables are collected as events, storing – beside the variable itself – information about the time and the location of the collection.

The new framework will enable a deeper understanding of the causalities of the system use and the utilized models through a more flexible and more granular data collection in the later experiment phases. One challenge is to integrate the new event tracking solution into the existing system. A second challenge will be to design the storage and collection in a way that it is reusable in terms of the purpose and unambiguous in terms of the data labelling across multiple data sources. An overview of the existing assessment concept in the system and an outlook on the data collected through the new framework is given in (Weber and Vas, 2015). Furthermore, this extension of the system enables the detection and tracking of student behaviour in the STUDIO system (Weber et al., 2015).

8.2.1. The STUDIO System Architecture

Figure 14 gives an overview of the architecture and the process of STUDIO. To model the respective areas of education for assessment, an expert in the role of a tutor selects a set of concepts of the overall domain, describing best the target education area. This selection of concepts, named “concept group”, describes a test model, which is used to extract the concepts from the domain ontology. The system provides for each knowledge area: knowledge-elements and relations to other knowledge areas from the “Domain Ontology” and from the “Question and Learning Material Repository” for each knowledge-element a pool of questions and learning material. This provides then the knowledge source for further assessment and learning and represents a knowledge structure of the selected test model, as a content enriched sub-ontology of the depicted educational area.

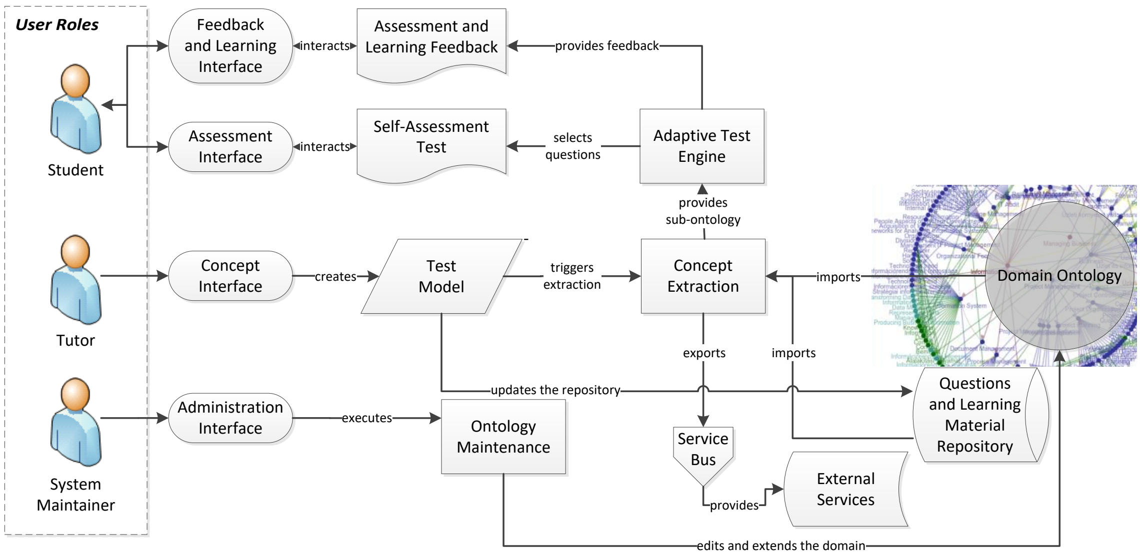


Figure 14: Architecture and process of adaptive self-assessment in STUDIO.

This final sub-ontology is engineered by the concept extraction, implementing the test model based on the previous extraction. Using the resulting knowledge structure, the adaptive test engine will then start the self-assessment from the question pools attached to the single knowledge-elements. Each extracted question gets visualized on the assessment interface with dichotomous answers from which the student selects the single correct choice to continue the test. The adaptive test engine will continuously trigger new questions, collecting responses, and exploring the backed knowledge structure, following the correlated knowledge-elements. When the test finishes the Feedback and Learning Interface is triggered, which visualizes the gathered results and correlations and gives access to the related learning material. Together this provides an evaluation of the personal perception, as a natural self-assessment experience.

8.2.2. An Event-based Perspective on Adaptive Testing

Adaptive testing enables to adapt to the level of knowledge of a test taker in a specific area. It deploys a strategy of testing, to select test questions in relation to the performance of the test-taker. Within an adaptive test, the interaction with the system can be considered as a sequence of events. E.g. answers to issued questions are acting as events to the question selection algorithm, while the algorithm derives new questions to assess, based on the feedback and the information about the state of the test. In this regards the general frame of CEP fits to the process of adaptive testing but introduces additional concepts to connect as well contexts and multiple lines of a knowledge areas as events with states for a new concept of adaptive testing. The focus of this part is not to implement a complex event processing solution but to collect considerations on event processing and shed light on more implications of the planned STUDIO system integration of the later concept importance based system extension.

Complex event processing (CEP), is an approach to software systems, which is based on the idea of events, incorporating logics to filter, transform, and detect patterns in events as they occur (Etzion and Niblett, 2010). It is a monitoring approach, triggering events over time, while observing the system states to derive insights through reasoning. Storing information on the test and activating and completing contexts as pre-modeled events enables to reason on the flow of a test, resembling the monitoring of a running technical system. In the vision of CEP, the approach of adaptive testing can be considered as a systematic, event enhanced, exploration of knowledge paths and areas.

The term of complex even processing goes back to the 2002's publication of Luckham, gathering and defining the concept and methodology as a new standard (Luckham, 2002), while the notion of event processing is used for a much longer time, throughout production near environments where time critical decisions are needed. Fitting, Etzion and Niblett define event processing in a general sense as “any form of computing that performs operations on events” (Etzion and Niblett, 2010). Extending, Luckham describes the complex event as “an event that could only happen if lots of other events happened” (Luckham, 2002), highlighting the dependency of events as one type of complexity.

In the context of an adaptive test, all test takers are operating in their own local environment, which is rich on additional information. The time to answer, the behavior in front of the test or the progress on the test are available, yet unused. Set into a context, they are suitable to be identified into events, as an input for further events and reasoning. Envisioning this set of events as input to the test taker's test environment and feedback to the testing instance in the fashion of a stream of events, it becomes reasonable to use the CEP as a metaphor to consider new solutions.

Figure 15 shows a CEP-based vision on the STUDIO system, interpreting the system operation as a stream of events. The test engine operates the test and sends new questions as personalized events to the user environments, which assess the questions and returns feedback about the given answer into the stream of events. In this frame the user acts as an implicit event consumer and producer. Together with the answer, additional potential information can be gathered at the test client site to be integrated into the event stream. With the methodologies proposed in (Weber and Vas, 2014) and (Weber et al., 2015) (see also Appendix 12.3), the potential arises to reason on the cultural context of a person or handle special qualities of questions which correlate to specific contexts of a job or a specific domain environment.

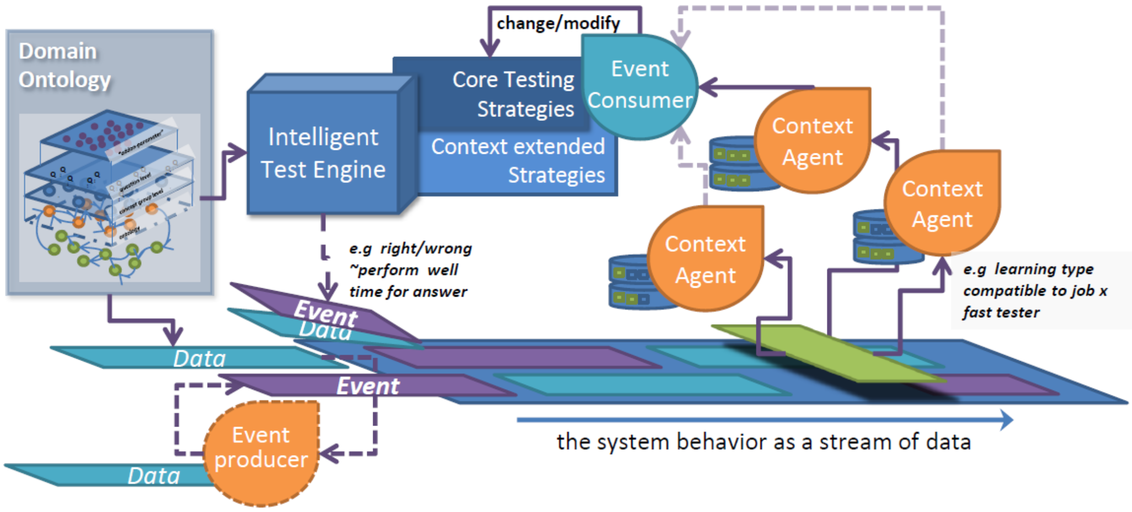


Figure 15: The STUDIO adaptive testing from a CEP perspective.

In CEP perspective, the test engine acts as a specialized event consumer, which is gathering the triggered events. The variety of questions, connected to the ontological representation of the domain knowledge, resembles an interconnected repository of assessment sources. As such it enables the freedom for different compositions of tests.

CEP is applicable to a wide range of problems, where in dependency of the field, other methodologies could act as alternatives, as within embedded systems, time-triggered systems (Obermaisser, 2011), or constraint programming (Rossi et al., 2006) as a programming paradigm resolving constraints. To select on the use and ability of CEP to cope with the selected problem of an adaptive testing approach, a set of requirements can be collected. Fitting to the requirements of the market (Chandy et al., 2011), defines a set of common requirements for CEP solutions, which should be met to make a sufficient use of CEP. A fit of the requirements for CEP to the frame of the adaptive test approach is listed in Table 14. The scenario is here a multi classroom self-assessment test to assess the existing knowledge for a range of different subjects and different domains of the central domain ontology at the start of a new school year.

Table 14: CEP requirements assessment for a CEP enhanced adaptive test.

Requirements	Suitability / Description
Event driven nature	***Medium/High:*** The initial test procedure is a straightforward feedback circle which becomes highly event dependent and event aware with the impact of detected contexts, fulfilling over time while opening additional complexity potentials.
Event rates	***Medium/High:*** The rate of events varies with the amount of parallel test takers. Yet the basic event rate could increase highly with the unlocking of additional relevant contexts for an adopted test environment.
Application complexity	***High:*** The domain-ontology-bound, interconnected knowledge structure, directly impacts the complexity of reasoning to derive changes within the system through the intelligent test engine. Further, the test environment adoption adds complexity with the availability of additional selected relevant contexts.
Timeliness	***High:*** A user is ready to accept system response times in the scale of seconds. Yet the derived changes impact the overall system behavior as well as the selection of personal queries. If a certain group dynamic is detected, the system behavior has to change below a range of seconds for all active tests, to prevent the sending of questions out of an expired testing behavior.
Recommended for event processing	***Yes:*** The system aligns with the requirement expectations of CEP. Further, the dependent triggering of events comes with the potential of a high combinatory impact to increase the target complexity.
Event driven nature	***Medium/High:*** The initial test procedure is a straightforward feedback circle which becomes highly event dependent and event aware with the impact of detected contexts, fulfilling over time while opening additional complexity potentials.

8.2.3. Extending the System Architecture with an Event Tracking Framework

The STUDIO system has been used, extended and evaluated in a number of European and nationally funded research projects, including applications in business process management and innovation-transfer (Arru, 2014), medical education (Khobreh et al., 2013) and job market competency matching (Castello et al., 2014). The system is used and evaluated on a regular basis and in different domains. More detailed examples are: for the integration of learning styles into adaptive learning systems to offer valuable advice and instructions to teachers and students (Truong, 2015), and the studies reported within this thesis, summarized in Section 8.4.6 and Table 30.

For each running test, STUDIO implements a statistic module which collects basic quantitative data about the number of assigned questions, how often tests are taken and how many learners take which test and when. This is completed by more qualitative data, collecting which questions and knowledge elements the students passed or failed. The data collected by the STUDIO statistic module tracks the assessment of the students and stores their reached performance. To conclude further on the mechanisms and impacts of STUDIO and to store how the learner interact with the system, an event-based logging system is developed and integrated, which collects the interaction with the system and keeps detailed information about the interaction as sequences of events.

The event-tracking framework is backed by a database-solution – the event tracking database – and targets to track the interaction of the learner with the system. Considering the STUDIO architecture, introduced in Section 8.2.1, the event tracking database follows the learner interaction and connects to the "Feedback and Learning Interface" and the "Assessment Interface", as pictured in Figure 16. Following and storing defined events, then finally enables to utilize the learner's behavior

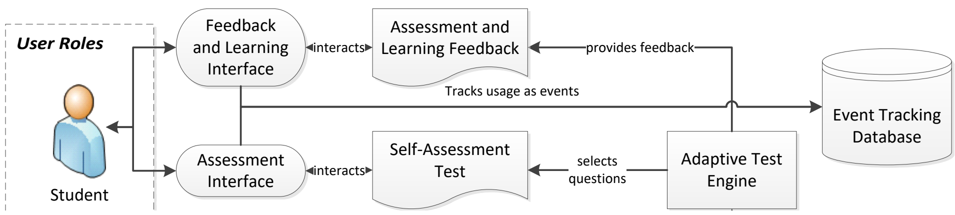


Figure 16: Extension of the STUDIO architecture. Events are tracked from the assessment and learning perspectives.

Each event stores information about the system in 7 dimensions, as described in Table 15 below. To be able to store a wide range of different events, the event template has to be constructed in a way that it can cope with a range of data formats. The solution is to define multi-purpose data fields, which store values freely, leading to a "String value storage" and "Numerical value storage" attribute. To finally keep the tracking consistent, the events have to be used in a concise manner across different interfaces and potentially different systems.

To link the event to the STUDIO system, the "Item reference" can store the unique ID of elements of the domain ontology as the question ID or a concept ID. The "Session identifier" is created once per access to STUDIO and is linked to the user and can be used to split users and their interaction alike. The events will model a sequence and include furthermore the detailed event time and a unique ID to sort and store specific events. Every event is linked to a location and a type of event to better grasp the interaction of the learner with the system and to group by locations and event-types. E.g. a learner reviews a passed concept in the result overview. In this case the event will be stored, linking the event-type “CHECK_RESULT” and the location “ONT_NODE” for the concept and storing the domain ontology related ID of the concept.

Table 15: Event blueprint to store events concerning system interaction.

Attribute	Description
Event-ID	The unique ID of the event.
Event description code	Which type of event and what factors are relevant.
Location code	On which part of the assessment-process or interface the event has occurred.
Session identifier	Each access of the system is one session for one user.
Numerical value storage	Multi-purpose field, filled depending on the event type.
String value storage	Multi-purpose field, filled depending on the event type.
Event-time	The time of the start of the event.
Item reference	A unique reference code, identifying the correlated item within the ontology. E.g. a question or a knowledge-element ID.

All events are stored in order of their occurrence, so if no explicit end event is defined, the next event for the same session and the same user is acting as the implicit end date. Extending the existing storage of information within STUDIO, the event logging system stores events with event definitions, as shown in Table 16 below:

Table 16: STUDIO events-descriptions to identify tracked events (excerpt).

Event type	Description
START_TEST	Marks the start of a test.
END_TEST	Marks the end of a test.
OPEN_WELCOME_LM	The user opened the welcome page.
OPEN_LM	The student opened the learning material tab on the test interface.
OPEN_LM_BLOCK	The student opened a learning material block on the test interface.
RATE_LM	The student rated the learning material.
CHECK_RESULT	The student clicked on a concept to view the result page for the concept.
FINISH_TEST	The test has been finished.
SUSPEND_TEST	The user suspended the test to continue later.
RESUME_TEST	The user has restarted a previously suspended test.
SELECT_TEST_ALGO-RITHM	The algorithm used to assess the learner is selected by the system.
TEST_ALGORITHM_-EVENT	The behavior of the current test algorithm changes, e.g. entering another stage of testing like “random testing”.
ASK_TESTQUESTION	Sends out a test question to the learner to answer.

STUDIO_LOGOUT	The user logs out of the STUDIO system.

To store the events, the system implements a logging database, splitting the concepts of the logging to a star-schema for efficient extraction, transformation and loading. The logging system is modular and easy to extend with new concepts and easy to attach to potential event positions within the STUDIO runtime. Together with the existing logging of the assessment evaluation feedback, this new extension tracks the exploration of the sub-ontology within the assessment and enriches the feedback data with context information of the student’s behavior on the system. The database scheme is shown in Figure 17.

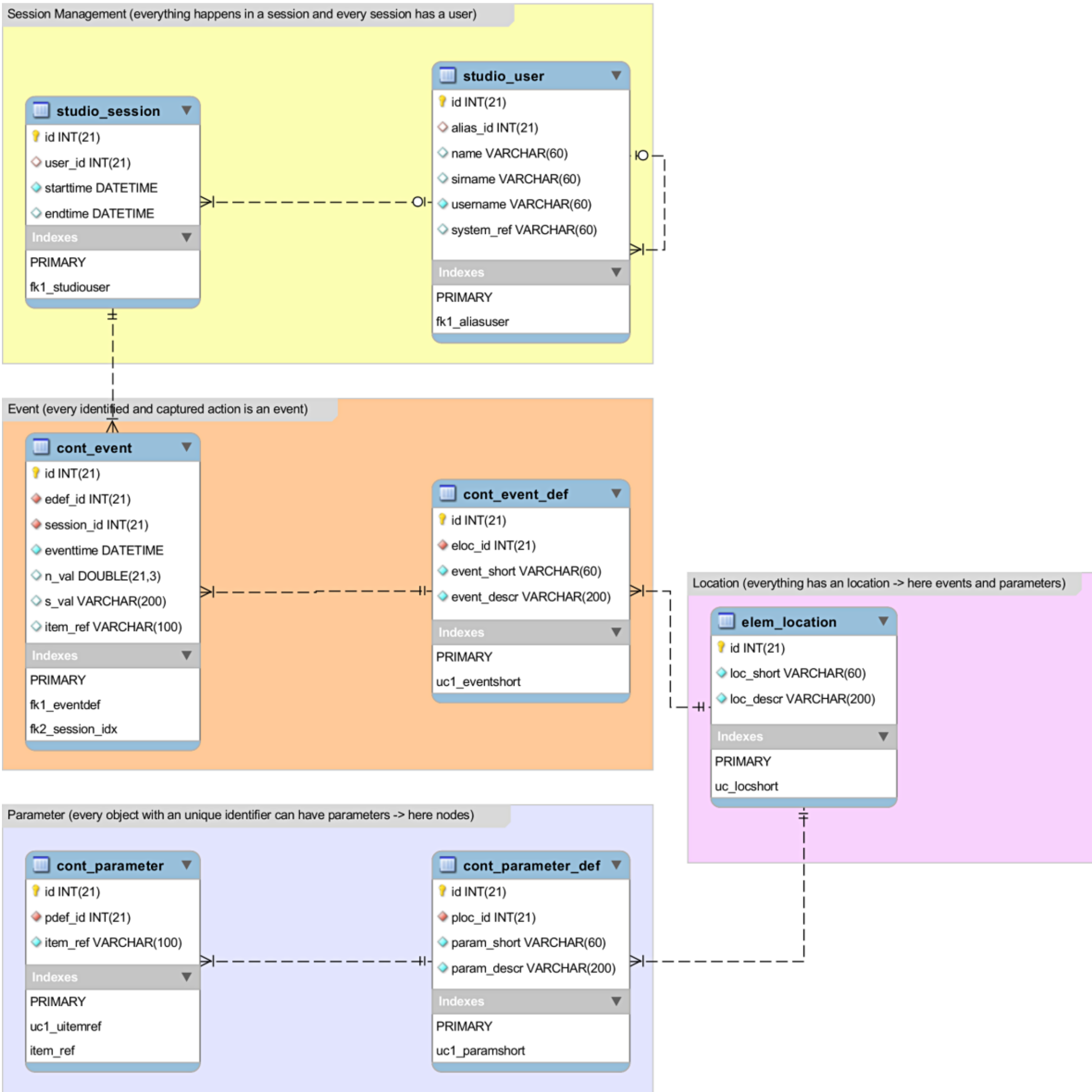


Figure 17: STUDIO star schema based event database.

The database scheme is implemented with the concept that all detailed information about the location, the type of events and the session is stored in separate tables, outsorcing the reptitious information and saving resources. The STUDIO system makes use of the database structure and manages the different events and locations. When initializing the database access, STUDIO caches all event and location definitions available within the database and synchronizes them with an internal versioned definition list. If a definition is missing, then the system will insert them into the database to be available for future events. This way multiple instances of STUDIO can define and use different events, depending on the current scope of the assessment (and assessment algorithms) and the used domain. Once the state of the system and the database is synchronized every insertion of an event is done without checking the definition tables and events can be inserted with a lower resource usage.

Additionally, the database stores parameters with a similar logic to events. Parameters can be “attached” to any known element type of the domain ontology, as questions or concepts and to any part of the interface. Each parameter is also stored with a location and a parameter definition record – following the logic of the event definition table and visualized in the database model in Figure 17. The “parameters” can be considered as flexible labels which can be used to further contextualize the system and the system interaction. E.g. can parameters be attached as bias-labels to questions to store information about a known bias, as explored in (Weber and Vas, 2014) and used in (Weber et al., 2015) and summarized in Section 12.3.

8.3. Design of a Concept Importance Measure for a Domain Ontology

Every learning process integrates a strategy for learning: natural and unconscious as part of the learner; or explicitly, given by a tutor who bases the interaction on personal experience or the explicit knowledge of learning theories. The major learning theories differ in the exploration on how learning occurs and in the explanation on how learning should be organized for a continuous and structured learning approach. Yet learning theories agree on knowledge being interconnected and being learning learned in the presence of other knowledge or experiences.

Embracing the narrative of knowledge being interconnected in the context of learning – as in the STUDIO domain ontology based approach to learning – each specific domain to learn and master can be understood as a network. A network perspective on learning can especially help to understand the complexity and connectivity of domains for learning.

Figure 18 shows the STUDIO-based visualization of the Management Information Systems domain. The visualization resembles a network structure and is based on the underlying domain ontology. Connections between concepts express that concepts are related and relevant to each other and strong and weak connected areas can be identified.

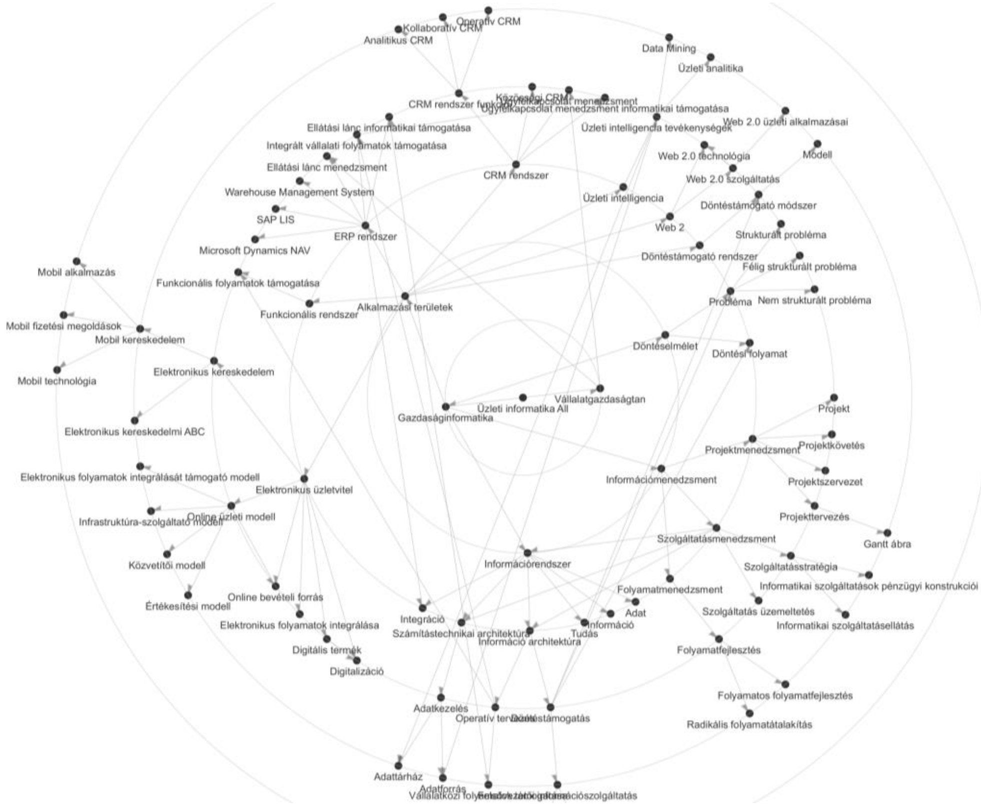


Figure 18: STUDIO based network representation of the Management Information Systems concept group.

Furthermore, the structure introduces a hierarchy. The overall topic centre in the middle of the network is the starting point and connects directly to generalizing topics in the first inner level. From there the structure spreads through directed relations to more detailed, factual concepts in the outer regions of the network. The hierarchical nature of the structure in Figure 18 is based on a semantic hierarchy, which is grounded on the directed relations of the domain ontology (the extraction of a structure/concept group from the domain ontology is covered in Section 8.1.1, 8.1.3 and 12.3.1).

The most frequent type of (directed) relation within the ontology is the relation *'Has sub-knowledge area'*, through which each extract of the ontology – in average – follows a "general knowledge to detail knowledge" flow, starting from the central concept and ending in the outer leaves of the structure. While this hierarchy resonates with specific ideas within learning theories as behaviourism (bottom-up learning) and constructivism (top-down learning), the structural hierarchy itself includes no statement about a specific way to learn based on the structure.

Keeping the observational point of view of this section, four main aspects of the semantic enhanced network structure for learning in STUDIO can be isolated. The main aspects are summarized in Table 17: connectivity, (semantic) complexity and a semantic-based hierarchy:

Table 17: Visual observations and aspects of concept maps, extracted from the STUDIO domain ontology.

Observation	Description
Connectivity	The concepts/knowledge to learn is represented interconnected. Interconnected concepts are connected to express they are related and relevant to each other. How they are related is expressed by the attached semantic. Concepts can differ in how many connections they have towards other concepts (outbound degree) and how many other concepts connect to them (inbound degree).
(Semantic) complexity	Concepts in the STUDIO ontology have different numbers of connections (the degree), concepts have different semantic types and are connected by relations with different types. The class definition of the domain ontology is given in Figure 11. While the complexity in terms of learning is not explicitly given, it is observable based on the structural and semantical complexity.
Semantic hierarchy	The presented structure is based on the STUDIO ontology and each concept and relation implements a semantic description, given by the formalization of the STUDIO domain ontology. The STUDIO ontology defines a variety of relation types which express a “detailing” relation as *‘Has sub-knowledge area’* and *‘Part of’*. As relations are directed and extracts from the domain ontology start with a source/root element, the more distant a concept is from the root element the more detailed concepts are based on the semantic o the relations (in average, as equal relations exist and reverse pointing relations are possible).

Figure 19 shows a topological map of a fictional lake. Taking into the account the observations, collected in Table 17, and assuming the observations yield a potential to link to a theory of learning, a concept network could be reimagined as a similar topological map of the learning domain. Deep areas would capture dense and complex areas to learn, while light areas are capturing less dense areas to learn, which are also being potentially less central for mastering a domain. Knowing such a quality and connecting a learning theory, the structure and semantic of the domain would be well-suitable to recommend on the final process of learning. This specifically is the set goal for defining a measure for the importance of concepts and creating a learning strategy which follows the logic of the measure. Both will be approached in the following parts of this thesis.

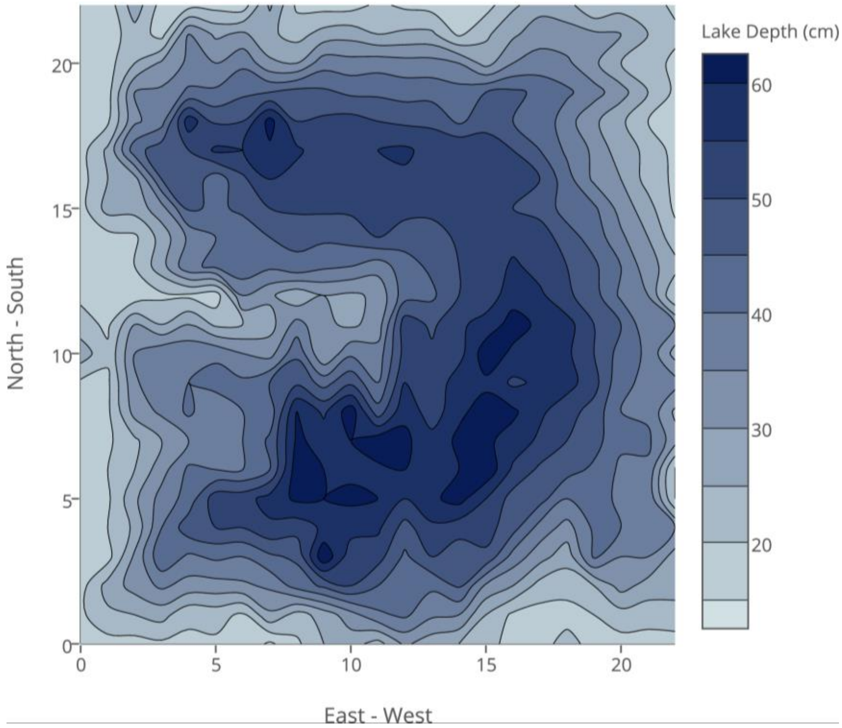


Figure 19: 2013 "classroom lake" contour data plot (Plotly, 2016).

More specific, in the focus of this thesis is the question if a strategy – and more specific, a measure to evaluate the importance of concepts – can be derived from a semantic enhanced structure of the concepts "to learn". The previous chapters helped to highlight discoveries in different areas, in the context of structured knowledge and learning:

1) The domain of **user modelling** provides methods to keep a steady track of the individual background and development of the learner (Section 4);

2) The contextual impact of **learning theories** (Section 0) and a special **organisational learning** focus (Section 2) act as frame for learning and recommend educational strategies to learn;

3) **Adaptive assessment** systems and their relevance for learning were addressed (Section 3). The software solution STUDIO, as a structure- and semantic-driven adaptive approach to support learning (Section 8.1) were introduced as a structure aware, ontology-aware solution for learning and as a platform for implementing an importance measure;

4) Finally, graphs and **network analysis** offer well-explored methods to express and analyse the connections of networks of people and information (Section 6);

Yet, no field or theory indicates or suggests so far a method to measure and differentiate the importance of individual concepts in networks of concepts for the task of a structured learning approach. An exception is the learning theory of connectivism (Section 5.3).

8.3.1. Connectivism as an Interpretation for the Importance of Concepts for Learning

As explored in Section 5.3, connectivism is a learning theory which is approaching learning in a technology and information driven society. It focuses on learning in environments where information is interconnected explicitly, as e.g. the world wide web (WWW) or other non-linear, hypertext-based (Cicconi, 1999) linked sources of information. Siemens formulated eight principles for connectivism (Siemens, 2005) in the first publication of the approach. Revisiting the work of Siemens – while takin into account the concept map perspective of STUDIO – three main pillars can be further extracted from the concept of connectivism. The pillars are shown in Table 18:

Table 18: The three pillars of a connectivistic concept importance for learning.

1) Learning to Form Connections is becoming a Core Competence	**2) Learning = Experience + External Information**	**3) Best Connections = Highest Importance for Learning**
*"Within the digital age **learning** cannot rely on personal experience any more but rather **is derived as a competence from forming connections."***	*"Connectivism strengthens the view **that learning is motivated by connectivity, connecting experiences but also external information,** residing in external, potentially interconnected, sources […]"*	*"[...] while **connections which offer the learner to learn more are more important** than the current state of knowledge."*

These statements are especially relevant in the context of designing a concept importance measure and in contrast to the existing major learning theories. The first pillar underlines the importance of forming connections to new pieces of information as a competence on its own. The forming of connections can be interpreted as a conscious exploration of connected information and resonates well with the circular approach of STUDIO of assessment, result reflection and learning, where the assessment tailors the recommended connections to the current performance of the learner.

The second pillar emphasize that learning is not only connecting existing experiences but also connections to related and relevant information as an external on-demand memory. While the claim of the theory is still open for more backing literature, it undermines the open learning circle of STUDIO which highlights future connections for learning in the result reflection phase. In this regard the explicit availability and the connectivity of known concepts or information is a valuable enabler to connect to new information. Most critical for capturing the importance of concepts in concept networks is the third pillar.

The third pillar emphasizes strongly that connections which enables to learn more (present or in future learning phases) are more important than the current state of knowledge (the part of the network which is already part of the experience). Reflecting the statement on the STUDIO learning cycle it can be concluded that: exploring a concept which is a good connector for future knowledge would be more important to assess – even if it would be failed first by the learner – than exploring a less well connected concept which is expected to be known by the learner, as the earlier concept would enable a better future exploration of other concepts. In this regard, connectivism favours the long-term potential of exploration – through better connected concepts – over a better short-term performance.

The strongest focus of connectivism is the connectivity as it enables a better access to further knowledge. A better access to other knowledge can come with a different quality and complexity: semantic meta-information on concepts and relations can help to evaluate the meaning of “enabling to learn more”, also complex measures of networks, as structural measures, can play a role in putting a concept of quality on connections and concepts. In this regard, to create a model of concept importance, different aspects and attributes of the knowledge structure and its semantics have to be extracted and compared. Here measures, described in the literature of networks and especially network centrality measures, can complete the picture of available inputs and potential methodologies.

Above the identification of features, a challenge will be to fuse available features to one concise measure – taking into consideration attributes and semantics of the structure, and potentially known information about the user. In this regard, it is valuable to collect, group and potentially transform the semantic of the domain ontology into aspects relevant for the identified pillars of a connectivism aware solution for learning, as e.g. the “need for specific knowledge” and the “degree of detail”. This collection can be completed by hybrid considerations on centrality measures as e.g. the degree of concepts. A suitable candidate for integrating different aspects into one measure can be simple utility functions, as used in the field of decision theory.

The output of the design task in this Section 8.3 is a sound concept importance measure. There are different challenges to overcome for the stage of modelling. In the early phases of the development a balance for the measure has to be created between the frequency and range of changes of single aspects within the measure and the resulting changes of the final value of the concept importance measure. This includes explicitly the investigation of the used domain model in terms of the structure of the captured domain knowledge and the semantics of concepts and their relations. The measure as such will be general in the context of the modelling but partially tailored to the focused environment of STUDIO.

The measure will be adopted for assessment and learning in STUDIO in the integration study in Section 8.4, where concepts with a higher importance should be tested earlier than others, to on one hand test a learner on the core knowledge of a domain and, on the other hand, provide the learner with recommendations to follow a personalized learning-path with a good access to more knowledge areas. This initial model will then be integrated into the developed algorithmic framework to facilitate the importance based selection of concepts to assess for the path-based assessment.

Finally, an experimental study within STUDIO in conditions, similar to the previous experiments, will be conducted to collect data about the impact of the new concept importance based assessment strategy. Furthermore, the collected feedback will be used to explore how to refine the final concept importance model.

8.3.2. Pre-Study: Analyzing Connectivity and its Influence on the Assessment Performance of Students

The rationale of this pre-study is to investigate if there is a relationship between the connectivity of a graph structure of concepts of a given domain of learning, and the performance of learners who learn and are assessed based on the same concept structure. If a connection is evident, it will underline the meaning of connectivism for practical solutions for assessment and learning.

To conduct the investigation, the connectivity will be measured by centrality measures. Centrality measures are going back to social science research to measure the relationship in networks of people (Newman, 2010) and continuously find application in new types of networks as project portfolio networks (Wolf, 2015), text analysis (Erkan and Radev, 2004) or protein networks (Jeong et al., 2001). For this study a selection of centrality measures is based on the selection of Section 6.1 and will be used and compared within the analysis.

8.3.2.1. The Pre-Study Design

The data of this study is based on the use of the STUDIO system in a higher education learning environment. The data were collected throughout a bachelor's level course on Management Information Systems at the Corvinus University of Budapest. Within the course, 247 students accessed the system throughout a period of 20 days to get prepared for the final exam. Throughout the period, students passed altogether through 1161 assessment and learning cycles, receiving - with repetition - 73654 questions out of a pool of 165 distinct questions, for overall 61 different concepts, which are the concepts or "nodes" within the underlying domain graph.

The result of each test is depending on the knowledge of the student and on the bottom-up evaluation logic, evaluating the performance on the fly and selecting in an adaptive way the next question to assess. As such, measuring the overall performance may not be fully independent of the rationale of evaluation. To break through this dependency this study will focus on a local performance. To do so a two-stage experiment will be conducted.

Within the first stage the focus is on a local "node neighborhood" performance, where only direct connected neighbors of a node are taken into account. Here the performance measure indicates how many surrounding nodes are passed if the node in focus is passed. In the second stage the focus is narrowed and for a given passed node only the node which is asked next is investigated and therefore a more detailed and direct impact is measured as a "next node" performance.

The nodes which are selected for investigation are selected based on their centrality and both experiments are repeated for each centrality measure. Both experiments include a final step where all passed nodes in the dataset are selected, to analyze the neighborhood and the next node performance. The idea is that if the centrality measures are good candidates to select nodes which predict the passing of surrounding nodes based on their centrality, then the average pass-rate of neighbouring nodes or the next node has to be significantly different from the average pass-rate of the complete set of nodes.

Table 19: Used centrality measures across 61 nodes.

Centrality	Min	Max	Mean	Std. Dev.	Var.
Katz-centrality	0.05	0.21	0.12	0.0367	0.001
Degree-centrality	1	7	2.33	1.578	2.491
Betweenness-centrality	0.00	0.57	0.07	0.1138	0.013
Eigenvector-centrality	0.01	0.39	0.09	0.0899	0.008
Closeness-centrality	0.16	0.32	0.21	0.0385	0.001

To prepare the dataset an extraction and transformation algorithm were developed in the Python programming language (v3.5.1) (Van Rossum and Drake Jr, 1995). For each analysis, one initial dataset was created, where – for every taken assessment test and assessed node – the direct connected nodes were calculated and stored if they were passed or not, while for the second experiment it was specifically stored if the next assessed node were passed or not.

For both sets all five centrality measures were calculated and integrated, using the graph specialized Python library NetworkX (v1.11) (Schult and Swart, 2008). A short summary of the calculated measures is given in Table 19 and an example for the calculated Katz-centrality (KC) and degree-centrality (DC) is visible in Figure 20.

Figure 20: Detail of a graph, extracted from the domain ontology, showing the tree-structure, and including the Katz- (KC) and degree-centrality measure (DC) as part of the node label.

8.3.2.2. Experiment I: The "Neighbourhood" Performance

The first experiment targets to collect the direct neighbourhood of selected nodes, to compare if it is passed considerably higher when the selected central node is passed. The initially selected nodes are selected based on their calculated centrality measure. For comparability, these central nodes are called “CNodes”.

To account for the intuition “the higher the centrality, the higher the local neighbourhood performance” the dataset of the experiment were filtered to all cases where “CNodes” were passed. Then the data were grouped for each respective centrality measure into classes with the same centrality value. These classes were sorted by the centrality measure and, from the highest to the lowest class, centrality value classes were added into the analysis dataset, till the new dataset accounted for 20% of the passed “CNodes”.

This selection was repeated for each single centrality measure and five analysis datasets were created to which a sixth set were added which includes all passed “CNodes”. This sixth group is the comparison group. Table 20 shows the resulting borders for each centrality measure to split the dataset of the experiment into an 80%/20% ratio. For each of the created centrality-datasets a basic descriptive statistic was conducted and the mean pass-rate were derived for the direct connected neighbours (“NHPassrate”) of the passed “CNodes”. Finally, the mean pass-rates were compared.

Table 20: Borders of centrality measure groups with 20% or less of all “CNodes”, where the centrality measure is bigger than:

Katz-centrality	**Degree-centrality**	**Betweenness-centrality**	**Eigenvector-centrality**	**Closeness-centrality**
> 0.14	> 3	> 0.10	> 0.13	> 0.24

Based on the summary in Table 21 it can be interpreted, that the mean “NHPassrate” is very closely centred around a mean pass-rate of 75.81% for all five observed centrality measures, with a similar close std. deviation.

Table 21: Mean “NHPassrate” in % for the analysis using the respective centrality measures.

	N	**Min**	**Max**	**Mean**	**Std. Dev.**	**Var.**
Katz-centrality > 0.14	13647	0	100	77.11	32.3658	1047.546
Degree-centrality > 3	8985	0	100	76.85	24.9339	621.699
Betweenness-centrality > 0.10	9046	0	100	75.94	26.6452	709.968
Eigenvector-centrality > 0.13	9154	0	100	74.34	29.1131	847.570
Closeness-centrality > 0.24	8238	0	100	74.82	28.9576	838.542
All CNodes	48711	0	100	76.87	35.0058	1225.407

While the observation is supporting a “the higher, the higher” trend, the comparison to the “all CNodes” comparison group, with a similar 76.87% mean, makes evident that this detected trend is universal for the passing of nodes in the neighbourhood of passed nodes, and therefore underlines no expected “better” trend for a high centrality measure, backing specifically the selected “CNodes”. So – surprisingly – the basic descriptive statistic indicates no room for a positive trend correlation.

8.3.2.3. Experiment II: The “Next Node” Performance

Following the preparation logic of the first experiment (select passed “CNodes”, replicate the set for each centrality measure and keep in each set only the rows which account for 20% or less of the highest centrality value bins), all passed nodes were collected and integrated as central nodes (“CNodes”) but in contrast the node were observed which were assessed next (“NextNodePassed”) by the adaptive assessment algorithm, if, and only if this next node still has a direct connection to the selected “CNode”. This framing is required as the assessment algorithm may jump within the structure from one path to another, if the earlier path were completely assessed.

For the resulting datasets, the mean of the performance was calculated again - this time for the next direct connected and assessed node. The summary is shown in Table 22. Again, the different centrality measures are on a comparable level for passed “CNodes”. Yet, in this experiment, the average mean across the centrality measures is 89.38% which is an 8.51% increase against the “all CNodes comparison group”, while especially the degree-centrality shows an increase with 10.46%.

Table 22: Mean “NextNodePassed” in % for the analysis using the respective centrality measures.

	N	Min.	Max.	Mean	Std. Dev.	Var.
Katz-centrality > 0.16	7446	0	100	88.34	32.0932	1029.973
Degree-centrality > 3	5852	0	100	90.99	28.6285	819.588
Betweenness-centrality > 0.10	6251	0	100	90.24	29.6776	880.758
Eigenvector-centrality > 0.14	5969	0	100	89.70	30.4027	924.322
Closeness-centrality > 0.24	5378	0	100	87.65	32.9002	1082.423
All CNodes	34363	0	100	82.37	38.1051	1452.000

Figure 21, takes a closer look at the degree-centrality against the “All CNodes” baseline. The graph plots a repetition of the experiment with different degree-centrality borders. The highest mean percentage of “NextNodePassed” is reached for a degree-centrality >= 4, while the percentage drops again when only >=5 relations are considered. With 92.08%, >= 4 is showing an enhancement of 11.79% against the overall mean.

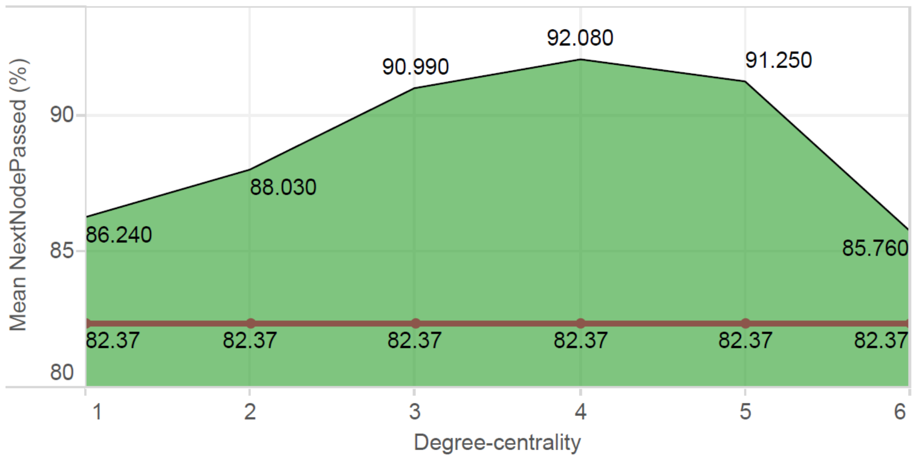


Figure 21: Mean "NextNodePassed" in percentage against an "AllCNodes" mean baseline. The x-axis is cumulative with "NextNodePassed" intervals of: [1), [2), [3), [4), [5), [6).

8.3.2.4. Pearson Correlations for Observed Variables

Table 23 shows the correlations between the observed variables. For "NextNodePassed", the positive correlations of the centrality measures indicate the results of the second experiment. For "NHPassed" the centrality measures show a weak negative correlation.

Table 23: Pearson correlations between the observed variables.

	NextNode-Passed	**NH-Pass-rate**	**Katz**	**De-gree**	**Be-tween-ness**	**Eigen-vector**	**Close-ness**
NextNode-Passed							
Pears. Corr.	1	0.820	0.070	0.142	0.097	0.104	0.118
Sig. (2-tailed)		0.000	0.000	0.000	0.000	0.000	0.000
N	45389	45389	45389	45389	45389	45389	45389
NH-Passrate							
Pears. Corr.	0.820	1	0.012	-0.001	-0.007	-0.042	-0.036
Sig. (2-tailed)	0.000		0.001	0.402	0.037	0.000	0.000
N	45389	64831	64831	64831	64831	64831	64831

Assuming that "Experiment II" shows a valid positive trend, then the negative correlation of the measures in "Experiment I" can be explained by the dominant prequisite of the assessment algorithm to sufficiently pass a rate of child nodes to be eligible to receive a question for the parent. This leads to a situation where a number of parent nodes may have potentially had been in the ability of the learner to pass, but weren't asked based on the assessment logic. While these cases are implicitly excluded by the criteria to take only passed "CNodes" into account for the experiments, the Pearson correlation includes these cases, which may tip the scale towards a negative correlation.

8.3.2.5. Pre-study Conclusion and Limitations

The question was tackled if, in the sense of connectivism, the connectivity of concepts can have an importance for assessment and learning and influence the observed performance of a learner. This study was conducted within the STUDIO learning and assessment environment and the analysed data was based on a real-world test with 247 bachelor level students, preparing for an examination in the domain of management information system.

Within two experiments it could be shown that a higher centrality measure for a given concept can play a role in predicting the passing of the next assessed, connected concept. In a broader neighbourhood, this trend couldn't be traced. Especially the later analysis has a strong overlap with the rationale of the bottom-up assessment logic to fail general nodes based on previous detailing nodes. Furthermore, the structure and rational of the used ontology will have an influence on the results and tests with other sub-domains of the ontology and an alternative ontology is recommended. Both, the influence of the testing algorithm and the specific ontology should be investigated in future studies.

8.3.3. A Concept Importance Measure for a Domain Ontology

The previous section has shown that within a STUDIO assessment test the degree of a node – as the basic measure for the connectivity of a concept – can be a predictor for the performance on the next tested concept. In this regard, the degree can be a predictor for the test flow within the assessment. Assuming that students indeed use the system to learn and not to practice the assessment, the finding that a passed high degree concept can forecast the performance of a connected concept, underlines that learning in a network of concepts *"is motivated by connectivity, connecting experiences but also external information, residing in external, potentially interconnected, sources"* (Siemens, 2005). Considering the statement of Siemens further, in line with Section 8.3.1, the learning material within STUDIO can be considered as a part of the experience of the learner if it is already “internalized” by learning or even if it is only connected to other experiences of the learner and known as a source of information. In both cases, the potential of information access and the connectivity would make the difference in the sense of connectivism.

The pre-study and the connectivity focused nature of connectivism highlight that considering the connectivity of the concepts (the degree of concepts) can be a suitable starting point for the development of a measure for the importance of concepts for learning in knowledge networks. Considering the pre-study, the correlation for the degree – comparing if a correctly answered concept with many relations forecasts the passing of the next assessed concept (the NextNode-performance) – is weak at a first glance. Yet, framing the consideration to a concept degree of >= 4, the passing of a concept with many relations becomes a ~92% predictor for the passing of the next concept and underlines the impact of connectivity on the assessment. The remaining discord in the correlation can be partially explained by considering that the degree may not account for all cases equally well, but rather there are connectivity ranges or groups where the concept connectivity is a particular good predictor for the performance on assessing the next concept (the main range is here around 4). Which can be seen, observing that a low concept degree yields results nearer to the performance average. Furthermore, later the performance within the experiment also drops for very well connected concepts with a degree higher than 4, which can outline that after a specific concept connectivity a change in the learning or conceptualization of the learned may occur (the students in the study learn in between the assessment phases, as shown in the concept Figure 10).

In contrast to the concept-degree, weaker correlations were observed for the remaining tested network measures. The measures are low performing in terms of their correlation to the performance measures and furthermore low performing in terms of their predictive power. So, taking into account that the degree is a good first stage indicator for an importance measure, what can be further integrated to strengthen a concept importance measure? Are there other sources of information available within the network to derive a measurement of concept importance for learning and assessment?

A network of concepts offers only a structural answer to these questions. In contrast the STUDIO domain ontology integrates a semantic model. Table 17 summarizes and focuses on three features which are be captured from the domain model, which can be exploited for a concept importance measure: *connectivity, (semantic) complexity* and *semantic hierarchy*. The connectivity is well represented by the concept degree – in line with the results of Section 8.3.2.4; the semantic complexity is represented by the different concept and relation types within the STUDIO ontology; and finally, the semantic-based hierarchy is given by the ordering nature of the relations, which model the decomposition of knowledge areas and the dependencies between concepts using the semantic definitions of concepts and relations. Accepting these three aspects as the source for the measure, the question of "how" to represent the aspects by numbers and how to combine the aspects to a measure.

The numeric degree of a concept (*connectivity*) is straightforward to translate into a measure, the translation of the semantic of concepts and their relations will need an interpretation (the *semantic complexity* and *semantic hierarchy*). To tackle the interpretation, two considerations will be explored about the semantic of concepts and relations, as gathered in Table 24: the *"dimensions of interpretation"* of the semantic (What are we looking for?) and the specific *"mapping of the interpretation"* to numeric values (How do we do the mapping?).

Table 24: Dual consideration on the translation of semantic concepts to numeric values.

Consideration	Description
Expected Modelling Dimensions	***What dimension or feature do the concepts and relations share?*** E.g. observing a common definition of a fox in a semantic framework, the fox may have different classes (like the STUDIO ontology has different types of knowledge classes) which carry the concept of a fox but are different in their features: colour, fur-texture, region, cultural meaning etc. For different cases of reasoning, different sets of features may be relevant and consequently differently important to derive a conclusion. For an abstract consideration of the fox, all the features are equally relevant to shape the concept of a fox, yet the goal of the interpretation of the fox for a specific task may differ in its dimensions of interpretation. Considering a cultural interpretation, it may be important how mischievous a fox can be, based on stories but also based on natural observations. So, for the dimension of mischievousness a fox may rate high being a good hunter, and for being known to be a shapeshifter in the world of stories – both expressed by the combination of concepts (e.g. hunter) and the relations linking to it (is_a_sneaky(hunter), enchants(humans)). Another dimension could be the worth of taming (de Saint-Exupéry, 2013), which would be interpreted well by considering classes like "wild animal" and relations as "is_afraid_of_hunters()" or "knows_about_roses()".
Numeric Interpretation and Mapping of Semantical Models	***How can an importance dimension be interpreted and then mapped to a numeric value?*** Considering a numeric interpretation of the semantic, for every chosen importance dimension (e.g. "mischievousness") a translation to a numeric value as a numeric interpretation must be modelled – either automatic through an algorithm, or expert-driven by directly framing and

	defining the values. To translate semantic concepts into a numeric interpretation a scale or number range for valid values has to be defined. The scale will be used as a target range to map concepts or relations which account for the identified dimensions on a numeric scale. The scale should be limited to a fixed range of numbers to keep mappings within one dimension comparable.

8.3.3.1. The Importance Dimensions of the Domain Ontology

To create a concept importance measure, the semantic of the domain ontology has to be interpreted – while taking into account the connectivity of the concepts. To do so the semantic will be interpreted. The final measure will integrate different factors, as the concept degree and furthermore aspects of the domain which can be interpreted as factors. These factors or "dimensions of interpretation" will be called further *"importance dimensions"*. To create the specific dimensions for the desired semantic interpretation – supporting the rating of the importance of concepts for learning – and to reason on the kind of importance dimensions, the domain ontology will be examined. To explore the abstract the idea of the importance dimensions, the following sections will use the example of a car to first abstract and collect observations, to then create a model of the desired importance dimensions.

Following the narrative of a car, a car represents a complex system composed of different components. Looking at the car in Figure 22, all parts are working together to make up the construct and experience of a car, yet not all components are equally needed. The windows are e.g. components which are part of the composition of a car, but they are not "needed" in an absolute sense, rather the windows "detail" the overall concept of a car. Furthermore, not all parts are equally needed. While, within the example of a car, the tires can be considered as highly needed, they also introduce a redundancy as one tire can be a spare tire.

Also, the locality of dimensions – so is an importance dimension interpreted in a general or local context – may differ. The seats are not observed as details but they are details in the context of the overall car. The cylinders of the engine on the other hand, are very small details in the context of the car as a whole, but in the context of the engine they are very prominent "details". Resulting, the different dimensions are interpreted in a locality context, which can be: 1) local, directly surrounding the current component, or 2) global, considering – as the context – e.g. the complete car.

Figure 22: A car as an example for a complex construct to interpret its parts in importance dimensions. [7]

The considerations regarding the importance dimensions, derived from the car example, can be applied similarly to the domain ontology. Revisiting the domain ontology structure, the concepts and relations can be equally semantically translated into the dimensions of:

1. "**need**" **–** "How much is a concept needed for another concept?", and
2. "**detail**" **–** "How much is a concept detailing another concept?"

The dimensions are selected in a way that a higher degree of "need" for a concept and a higher degree of "detail" of a concept is considered superior in the context of learning. The dimension of "need" directly connects to the assumptions of connectivism that selected, highly connected concepts are important enablers to learn surrounding concepts, which also correlates to the notion of "needed" concepts. Furthermore, the "detailing" aspect connects to the behaviouristic assumption that a detailed, fact-intensive basic set of concepts is needed to master a domain and access and master more complex general concepts.

[7] The image comes with no copyright restrictions. Science, Industry and Business Library: General Collection, The New York Public Library. (1918). 1918 - Oldsmobile - Model 45, 3 cylinders. Retrieved from http://digitalcollections.nypl.org/items/510d47db-bab6-a3d9-e040-e00a18064a99

As within the car example, the aspect of *locality* additionally plays a role in the application of the dimensions of “need” and “detail” onto the domain ontology, but it also introduces conflicts in terms of the related context. Are the dimensions valid in a local context, among single nodes/concepts; or in a global context, among whole networks of nodes/concepts? The answer can be derived jointly from the *concept of the domain ontology* and the *theory of connectivism*.

The domain ontology is tailored to the context of learning, but is – in these boundaries – a flexible construct for modelling knowledge, as visualized in Figure 11. STUDIO employs a dynamic process of tailoring, where concepts are tailored on-demand to specific assessment and learning scenarios. As gathered in Table 17, the ontology represents – based on the different types of relations between the modelled knowledge – also a knowledge hierarchy. E.g. two concepts can be connected by a “Has sub-knowledge area” relation – indicating that the linked concept is considered as a detail of the former concept – or concepts are connected through the “Has parts” relation – expressing a further decomposition. In this way, multiple concepts can be connected, and together build a longer chain of hierarchy which “details” the domain.

Looking closer on the hierarchical aspect of the domain ontology, the rational and limitations on how connections between concepts are made must be considered. The STUDIO domain ontology is an expert modelled and expert maintained ontology (see Section 12.3.1). Classes and realtions are defined in such a way in the STUDIO ontology, to ensure offer flexibility and enable the modelling of educational domain of different nature and complexity. Thus, the ontology – and any tailored sub-ontology – allows to include loops, or hierarchies where the chain of semantic relations may follow different orders and the semantic of the chain of concepts change over the sum of concepts: E.g. from the perspective of a medical practitioner the knowledge of biology may be regarded as sub-knowledge, while for a biologist the medical knowledge may be regarded as sub-knowledge. Both semantic viewpoints, have a rational, yet, without explicit constraining rules, this type of conflicts could only be resolved by direct design decisions by an ontology engineer (Neuhaus et al., 2011).

A semantic conflict between relations throughout the domain may stay undiscovered if has to be traced by the ontology engineer over many relations. In this case, algorithms or mechanisms for consistency could be used (Haase et al., 2005). In favour of the freedom of design and the speed of reasoning (consistency checking comes at a cost) the

STUDIO domain ontology implements no explicit (constraint-based) ontology consistency checking. An explicit approach is outside of the scope of this work (the software solution part of STUDIO implements solutions regarding the maintenance of the domain ontology). Yet, the general possibility of semantic inconsistencies is intended for the STUDIO domain ontology to implement the knowledge of experts more freely. One strategy to cope with it in the frame of STUDIO, is the use of Concept Groups to tailor the ontology to specific application cases which then have an improved consistency based on the additional framing to the tailoring goal.

Taking into account the collected considerations on semantic loops and the overall trend of the ontology to model a complex overall hierarchy of concepts (which is also complex to consider as a whole), the concept importance measure will focus on single concepts in the domain and not multiple concepts at one (as e.g. for the PageRank centrality (Page et al., 1999)).

Furthermore, from the perspective of **connectivism** and learning, connectivism doesn't motivate explicitly and exclusively to consider large networks, but to consider "*connections which offer the learner to learn more*" – as collected in Table 18. The theory of connectivism is neutral regarding the locality and supports a local context (but also isn't necessarily limited to it). Local, in this regard, focuses on concepts individually in the context of their immediate neighbours which are directly connected to the initial concept without an intermediate concept.

So, finally, to choose a flat starting point in line with the demands of connectivism and considering the complexity of the domain ontology, this work will implement: *a local concept to concept focus, considering single concepts and their direct relations to neighbours*. This way, the focus is on single concepts, to map the concepts of the domain ontology to numeric values for the importance dimensions of "detail" and "need" (nevertheless, a hybrid approach is introduced as an extension in Section 8.4.4).

But, what parts of the ontology should be considered to map the semantic of concepts and relations into values? Concepts within the STUDIO ontology are differentiated by different types – knowledge areas, basic concepts, examples, etc. – with a different semantic in the context of learning. Revisiting connectivism, the central idea is that the importance of concepts is based on how well they are connected and, more specifically, is based on how well a concept connects new sources of information. In this context, *not* *the existence of a concept itself but its ability to connect to other concepts is important* – so the relations to other concepts are in the focus and consequently their specific semantic. Concepts and relations are equally described by semantic but, comparing the semantic of the knowledge area focused core concepts and their relations, it becomes evident that the relations either mirror the semantic of the concepts or even detail them:

- **“Knowledge Area”** is detailed by three relation types ***“Has sub-knowledge area”***, ***“Requires knowledge of”*** and ***“Part of”***.
- **“Basic Concept”** relates to other knowledge elements with ***“Refers to”*** and is detailed by ***“Premise”***.
- **“Theorem”** is related to **“Basic Concepts”** and other **“Theorems”** by ***“Conclusion”*** and relates to knowledge elements by ***“Refers to”***.
- **“Examples”** are related to other knowledge elements by ***“Refers to”***.

Two general groups of concepts can be interpreted from the different types *Knowledge Areas* (Knowledge Area) and *Knowledge Elements* (Basic Concept, Theorem, and Example). The groups implement internally relations of equality and/or relations which further detail concepts. Specifically, with the “Part of” relation the group of Knowledge Elements details the group of Knowledge Areas. In general, the relations are either equal in discriminating the concepts and their semantic, or offer more detail to the relationships of the concepts, as the “Premise” and “Conclusion” relation.

Considering in more detail concepts which are part of the Knowledge Elements, Knowledge Elements are different in their types but share – beside the combination “Basic concept” / “Theorem” – the main relation “refers to”. “Refers to” on the other hand doesn’t offer a strong discriminating power and as such limits the assumed difference of the concepts in regards to a difference in the desired modelling dimensions “detail” and “need”. The groups are visualized in Figure 23. *In consequence, this work will focus on the semantic of the relations as the representative for the semantic of the concepts.*

For this work the importance dimensions of “need” and “detail” are derived from the individual concepts, based on the semantics of their relations. Every relation contributes with “need” and “detail” at the same time as every relation expresses the “need” and “detail” of a connected concept. The final measure will then account for all relations a single concept shares and enables to compare individual concepts. As such, the measure implicitly integrates the degree of each concept by considering all relations of a concept at once (so the degree of connectivity) – in line with the assumption of connectivism that the degree of a concept can indicate a better access in the context of learning and furthermore, in line with the findings of Section 8.3.2.4, that the concept degree can be a weak predictor for performance in a structured learning and assessment approach.

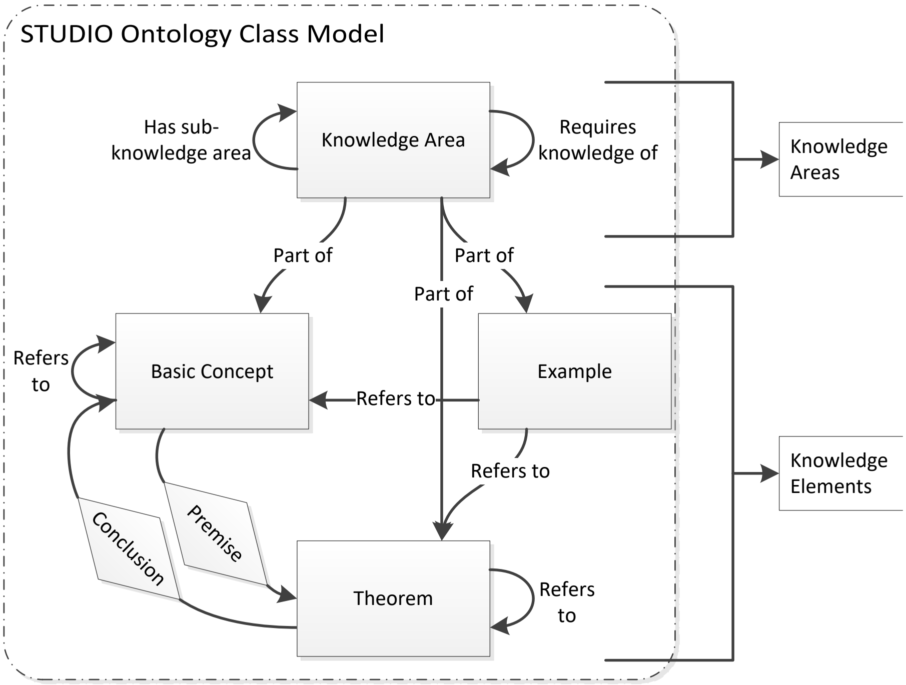


Figure 23: Grouping of concepts, visualized on the STUDIO domain model.

Equations (9) and (10) below are picturing the basic idea of the mapping of relations to their connected concepts. Each importance dimension is collected and summed separately over the n connected relations, which address the specific importance dimension mapped to the individual concept/node. The specific value of a relation is based on the semantic of the used relation and explored in detail in the following section 8.3.3.2. The number n is equal to the degree of a node (while single relations may contribute with a value of 0 to an importance dimension).

$$\mathrm{Concept}_{\mathrm{needed}} = \mathrm{Rel}_{\mathrm{need}_0} + \mathrm{Rel}_{\mathrm{need}_1} + \cdots + Rel_{need_n} \tag{9}$$

$$\mathrm{Concept}_{\mathrm{details}} = \mathrm{Rel}_{detail_0} + \mathrm{Rel}_{\mathrm{detail}_1} + \cdots + Rel_{detail_n} \tag{10}$$

8.3.3.2. Preparation of the Numeric Interpretation of the Semantic Relations of the Domain Ontology

Assigning or mapping numbers to semantic can happen in multiple ways. One well explored solution is the use of fuzzy sets. Fuzzy sets enable to account for uncertainties by allowing to describe environments in which “the goals and/or the constraints, but not necessarily the system under control, are fuzzy in nature” (Bellman and Zadeh, 1970). In fuzzy sets, entities can belong to multiple sets with a value set in an interval of [0,1] and it is possible to attach labels to specific values, which represent the fuzzy “semantic” of a group of entities, e.g. “tall”, “medium sized” and “short” for the set of sizes. Other alternatives are explored in the field of psychometrics, which tackles to measure psychological phenomenon. One example is the Item Response Theory (IRT) (Carlson and von Davier, 2013), as addressed in Section 3.4.

Also suitable are rating methods for goals, features and alternatives in the modelling of decision processes in the field of decision theory. The motivation is to model decision processes which can derive complex decisions in environments where parts or all factors are non-trivial to quantify and human perception and judgement is part of the decision process. Well explored examples are the Cost-Utility Analysis (CUA) (Nas, 2016) and the Analytic Hierarchic Process (AHP) (Saaty, 1987), as a more complex, extended solution. AHP defines a hierarchy for the decision process, decomposing goals into criteria and final alternatives. Within the process the goal is labelled with the value of 1 and all decomposing criteria in the hierarchy share a faction of the 1.0, adding up to 1 and representing the weight or priority of the criteria.

In contrast to the available methodologies to translate concept semantics to numeric values, this work will use an elementary approach, motivated by but not following the approach of AHP. AHP models a complex decision process, expressing the expert’s preferences in a decomposed structured decision process for rating decision alternatives, while the concept importance measure targets to measures individual concepts for a guided assessment and learning process. The aimed assignment of numbers will be based on an expert modelled, translating the semantic of the relations to numeric values. The assignment of numbers which will be proposed, is strongly connected to the goal of learning, which may introduce a degree of uncertainty. In such an environment, fuzzy logic could be an appropriate representation. Yet, it is not in the scope of this work to account for the uncertainty of different factors which (are known or unknown to) influence a measure for rating the importance of concepts for the goal of learning.

The study in Section 8.3.2.4 (Weber and Vas, 2016a) showed a weak correlation between the connectivity degree of a concept and the performance within a guided assessment test, based on the STUDIO domain ontology – which is addressed by this work. The study considered the performance of a student, which pictures the state of learning of a learner at one given time in the learning process, while learning is a continuous process. As such, correlating this “snap-shot” performance to the connectivity degree of a concept can only partially capture the learning of a learner (even if the learning process continues throughout the assessment by connecting the already learned to concepts and statements of the assessment questions). Nevertheless, the correlation to the connectivity degree is an indication that the assumption of connectivism that the connectivity of a concept has an influence on the learning of surrounding concepts (e.g. better access) is right. As such a measure to track the importance of concepts in the context of learning will benefit from integrating the connectivity degree as a source of information.

The domain ontology incorporates solely directed relations. To consider the connectivity degree of concepts, while equally considering the quality (semantic type) of the relation of an individual concept, this work will consider incoming and outgoing relations simultaneously. To do so, the concepts will consider the degree of existence and the degree of absence of an importance dimension, represented by the individual attached relations. To ease the modelling process and the human understanding of the translation of the importance dimensions of a relation to numbers, a relation will be rated for a specific importance dimension on a scale of [0,100]. E.g. a relation can be rated, for a single concept and the importance dimension of “need”, with the value 80 on a scale of 0 to 100, indicating that, based on the relation (and its direction), the concept is highly needed for the neighbour connected through the considered relation.

The mapping of the importance dimension additionally has to as well consider the absence of a specific importance dimension. The chosen solution is to split the interval into two ranges to consider either the existence or the absence of a importance dimension. The first interval with [0,50) expresses the degree of absence of the specific dimension and (50,100] the degree of presence of a given dimension. The mid-position of [50] is reserved to express an indeterminate state where the individual importance dimension has no effect on a given concept [8].

An example is given in Figure 24. The “Basic Concept” is connected through a “premise” relation to a “Theorem”, expressing that the “Basic Concept” “is a premise” for the “Theorem”. The attachment of the specific dimension to a directed relation and the following mapping of the connected concepts can be imagined as a slider attached to the given relation. The relation binds 100 value-“units” (integers) of the importance dimension – here “need” – and through the virtual slider the units get distributed to the connected concepts. In the given example “80” units are distributed to the “Basic Concept”, expressing that the “Basic Concept” is highly needed for the “Theorem” as a pre-condition – connected by the “premise” relation. At the same time “20” units are mapped to the “Theorem” to express that the concept needs another concept to a high degree (in contrast to the “is needed” expression of the inverse direction).

Figure 24 summarizes the intuition of the split interval and the slider distribution. Every relation has a “need” and a “detail” value attached, considering all importance dimensions at the same time.

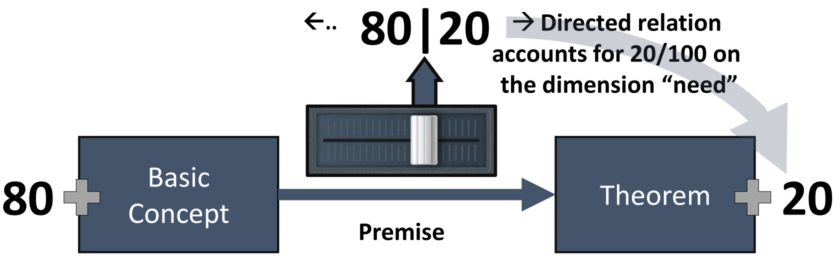


Figure 24: Splitting of the available value of the dimension onto the concepts connected with a given relation.

The translation of the semantic dimensions into numbers considers the direction of each relation. So, in the case of the relation “premise” the concept with the outgoing relation “premise” “is needed” by the concept to which it is pointing to, while the target concept “needs” the former concept. Depending on the defined translation of the importance dimension and the direction of the relation, the context of the translation changes between “needs” and “is needed” and equally “details” and “is detailed. Table 25 shows the translation of the importance dimension. The translation motivation (e.g.

[8] Following these interval definitions, the lower [0,50) interval includes one more “unit” – by including explicitly the 0 – than the slightly smaller interval (50,100]. This difference is accepted in favour of being able to express the complete absence of a dimension through 0 and the natural selection of “50” as the neutral middle value.

“needs” vs “is needed by”) changes based on the relation, the direction of the relation and the defined value of the translation for an importance dimension for a specific relation.

Table 25: Intuition for the numeric mapping of the importance dimensions.

Intuition:	“It is Needed”
Needs	< 50
Equally needed	50
Is needed by	> 50

Intuition:	“It Details”
Contains parts	< 50
Equally details	50
Is part of	> 50

To consider separately each relation a concept connects, every relation is regarded as a pair of mirrored relations: the regular relation, counting for the concept for which it is an outgoing relation and an implicit inverse relation, counting for the other connected concept, carrying the inverse values of the relation. In case of the relation “premise”, the concept from which the relation points to another concept, and which is regarded as the source of the premise (the outgoing relation), receives from the relation Rel_{need} = 80, while the other concept receives Rel_inv_{need} = 20 through the “virtual” inverse relation.

The latter concept receives a low value in terms of the “need” (20) as it is highly dependent on another concept. At the same time the low value still ensure that the relation is considered at all and considers that the concept has a relation which can be followed to access more information for learning. The inverse relation “inv. premise”[9] allocates, for the concept to which the relation points to as a “premise” (inbound relation), $Rel\_inv_{need} = Rel_{need_{max}} - Rel_{need} = 20$. This way the scenario of considering incoming and outgoing connections is projected to a scenario where every relation is mirrored by an inverse relation and only inbound relations are considered.

Figure 25 illustrates an example for the importance dimension of “detail”. The “Knowledge Area” “is detailed” by the outgoing relation “has part” and as such gathers a low “detailing” value by the relation, as it “contains parts” rather than detailing the local connection. Again, the low value ensures that the additional relation is counted, while translating at minimum a lower value. In contrast, the inverse relation “is part of” applies a high (inverse) value to the target concept, for which it is an outgoing relation. The idea of the mirroring of each relation with a virtual inverse relation is that for every concept

[9] In terms of the semantic of the domain ontology the (semantic) inverse relation of “premise” is “conclusion”. Considering “virtual” inverse relations here has only the rational to help and separate the calculation of the importance measure and is not an explicit semantic statement.

then only outgoing relations have to be considered, while the relations are a mix of real relations and of mirrored inverse relations. This way an algorithm can select all outgoing relations, extract the importance dimensions for each relation from a look-up table (addressed later in Table 26) and calculate the measure.

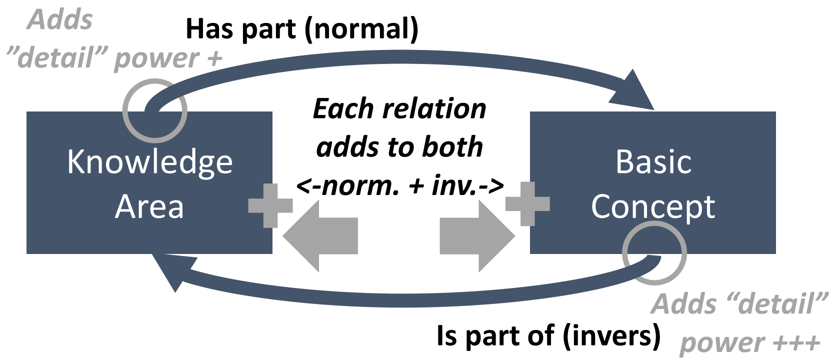


Figure 25: Every directed "normal/regular" relation is mirrored implicitly by an "inverse" relation to split how the semantic dimensions are translated to the connected concepts.

To count with all relations meaningful in the context of learning – which responds to the target of modelling the importance of single concepts for the goal of learning – those relations and concepts are considered which are relevant for the scope of learning. The STUDIO domain ontology implements two general sets of concepts and relations: 1) a curriculum related set, focusing on competences and actions and tasks connected to competences – which is not addressed and out of the scope of this work and available through (Vas, 2007) – and 2) a learning and assessment focused set, addressed in this work and elaborated in the previous sections and visualized in Figure 23. The later, learning focused set, will be the basement for modelling the concept importance measure.

8.3.3.3. The Numeric Interpretation of the Importance Dimensions

Multiple approaches (manual, semi-automatic, automatic) are applicable to define or derive numeric values for each concept, reflecting the semantic of the relations and thus concepts in the context of the importance dimensions "need" and "detail". A manual definition represents and an expert modelling approach where the number translations of the importance dimensions are manually defined based on the expert knowledge about the scope of learning, the domain to learn, and the structure and semantics of the domain ontology. Automatic approaches can use machine learning based solutions to e.g. observe concept and relation related patterns with (supervised) or without (unsupervised) an explicit numeric representative of the learning progress while using the STUDIO domain ontology for supporting the learning of a specific domain. Hybrid approaches and expert supervised approaches can frame semi-automatic processes. Considering the volume and scope of the presented research questions an expert modelled approach is chosen to directly reflect the experience of the users, maintainers, and architects of the STUDIO solution – selected for this work – and to furthermore frame the resources available for addressing the research questions.

The expert modelled mappings for the chosen importance dimensions of "need" and "detail" are shown in Table 26. The table shows on the left part the relations of the domain ontology which address the context of learning. The right part collects the mirrored / inverse relations. The mappings in the "detail" and "need" related dimensions also mirror for each inverse relation the regular relations and together sum to 100.

***Table 26: Collection and comparison of the expert modelled mapping values for the relations in the importance dimensions "need" and "detail"* [10].**

Relation	Relation "is needed" (0-50-100)	Relation "details node" (0-50-100)	InvRelation "details node" (0-50-100)	InvRelation "is needed" (0-50-100)	Inverse Relation
has part	40	20	80	60	is part of
has sub-knowledge-area	45	35	65	55	has parent domain
premise	80	90	10	20	inv. premise
conclusion	20	10	90	80	inv. conclusion
refers to	50	50	50	50	referred by
requires knowledge of	0	35	65	100	knowledge required by

A part of the relations tends to be underrepresented in the domain ontology as "conclusion" / "premise" and "refers to". They exist as a "blueprint" within the defined relations of the STUDIO domain ontology, yet the nature of the modelled content requires these relations less to express specific domains. The most frequent pair of relations are among the "Knowledge Areas" (bright coloured rows in Table 26) with "has sub-knowledge area" and "requires knowledge of".

[10] The inconsistent naming of inverse relations has historical reasons and is based on the software implementation of the STUDIO approach.

All translations of the importance dimensions to numbers and weights in this thesis will be expert modelled. They represent the natural observation and understanding of the relations in the scope of learning and are based on:

1. the experience collected in the **system exploration** (STUDIO, see Section 8.1),
2. the description of the **domain ontology** and
3. the interaction with the **system designers and learners** and
4. the **tracked usage-data** of the STUDIO system.

The basement for the selection is to design the mappings in a way that:

1. **concepts can be separated** by a single relation when combined to a single measure and
2. *especially* **the dimension of "need" can be differentiated well –** as the "need" dimension represents a strong statement about the requirements of learning, embedded into the domain ontology definition.

An extra benefit of mapping the semantics of the relations directly onto the concepts (instead of vice versa) and splitting the single importance dimensions across the relation is, that it is possible to consider concepts and relations even if the concept to which the relation is pointing to is unknown (the unknown concept could at best be narrowed based on the "allowed" relation in terms of the domain ontology). Furthermore, this "local" calculation of the indicator is beneficial for situations where connected concepts are unknown pieces of information. The local calculation also supports the learning, as learning is a step-wise process and may involve cases where it is still not (yet) known what will be mastered by an individual student. In such situations, a learning window can be applied and limit the set of concepts in focus, which would still enable a complete calculation of the concept importance measure.

Furthermore, the short-term "local" information what the student masters and should master in the learning process, may be of more use than a complete long-term forecast of the global single best connector to new information. To fully compare this "local" distance 1 neighbourhood concept importance against a possible >1 distance neighbourhood, further research outside of the scope and frame of this work is needed. A solution would need an extension to control and limit an observation window as e.g.

introducing a damping factor for the calculation, similar to the PageRank document importance rating (Page et al., 1999).

8.3.3.4. A Thought Experiment on Considering Incoming and Outgoing Relations

The allocation of the importance dimensions to concepts, based on their relations, considers incoming and outcoming relations. The explanation of the process is given earlier, yet it would be also an option to consider only incoming or outcoming relations. The rational of considering both is to differentiate concepts better, based on their relations.

The following figures Figure 26 and Figure 27, illustrate – based on a closed system of concepts – the difference of the differentiating power of considering either only inbound connections, or considering both connections. The circles are representing concepts and the arrows relations. All relations are using the same importance dimension with the same values (e.g. “need” of “requires knowledge of”). For both figures the outer green ring considers only inbound connections. The yellow ring considers both, in- and outbound connections.

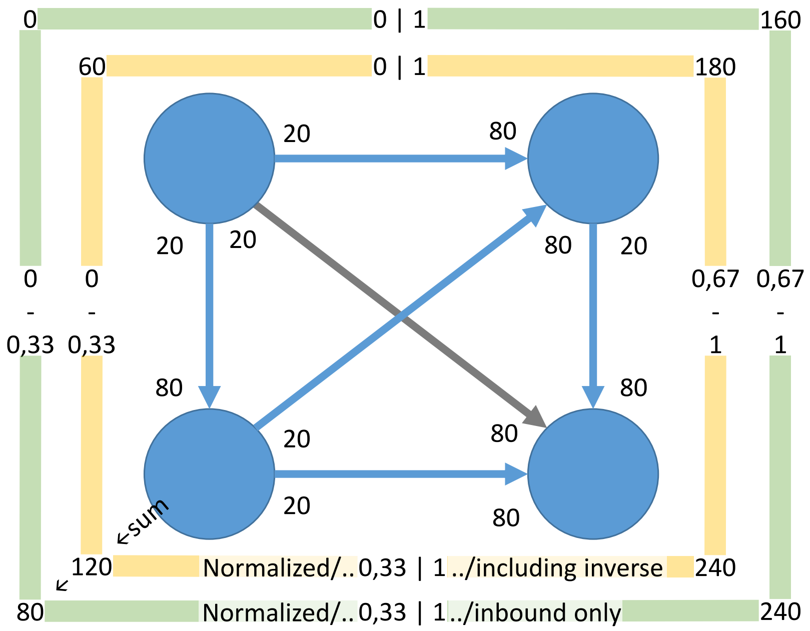


Figure 26: Comparing a normalized importance dimension mapping in a closed system with an equal number of relations per concept.

For both figures the corners of the coloured rings are showing the sum of the relations of the node in the corner. In both figures the sum of the mappings is normalized to the minimal and maximal sum in each ring to render the illustrated cases comparable. The normalized values are shown in the middle of the rings for each ring. E.g. in Figure 26, “0 | 1” means that for the green ring the “0” sum of the top-left concept is normalized to “0”, while the “160” sum of the top-right concept is normalized to “1”.

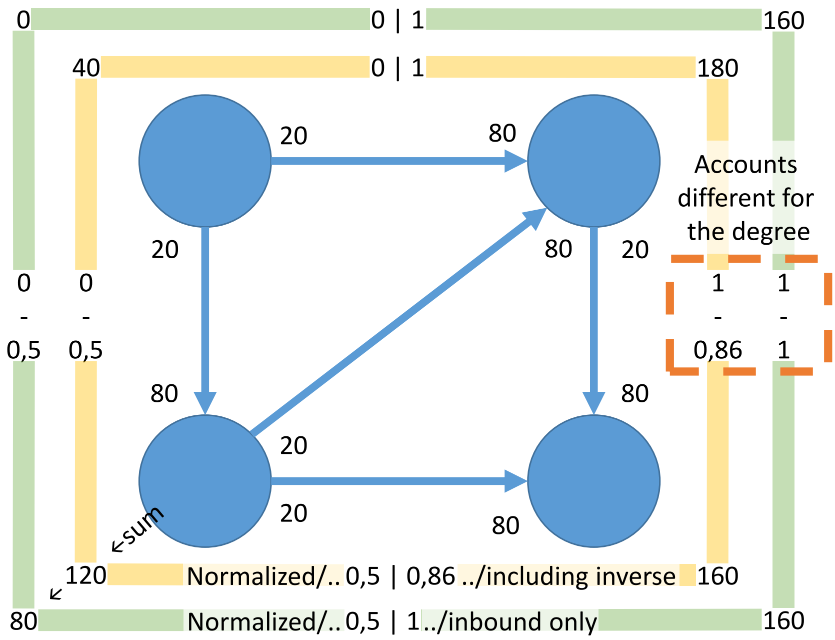


Figure 27: Comparing a normalized importance dimensions in a closed system with an different number of relations per concept.

Figure 26 models an equal number of relations per concept. Despite the different strategies to account for the relations (only incoming or both), all normalized sums of the mappings – illustrated in the middle of each coloured path between two concepts – are equal for both coloured summing approaches. In contrast, the system in Figure 27 models an unequal number of relations between different concepts.

The difference between the systems is the existence and absence of the grey coloured connection (from the top-left to the bottom-right concept) and the resulting change of the sums at the corner. The resulting sums of the relations are calculated in the same way. But the minimum and maximum values for the normalization of the sums changes with the different number of relations and the sums scale accordingly.

In contrast to the system in Figure 26, in Figure 27 the bottom-right concept cannot account for the change, introduced by the missing connection if only inbound connections are considered. Consequently, the top-right and the bottom-right concept share the same normalized value along the green edge and only the mapping which considers in- and outbound connections fully accounts for the modelled difference. This difference underlines the design decision to consider both: incoming and outgoing relations.

8.3.3.5. Fusing Importance Dimensions to an Integrated Concept Importance Measure

The previous sections isolated the *importance dimensions* to be map to numeric values and the actual *strategy for mapping concepts* – considering the *relations of the domain ontology*. Figure 28 visualizes a learning-path through the domain ontology. For each concept on the path a value has to be derived to express the importance of the individual concept for learning. To account for the "need" and "detail" of a concept, while considering multiple relations for a concept simultaneously, requires a strategy to fuse the dimensions of all relations to one measure. The goal is to reflect the dimensions and the degree equally.

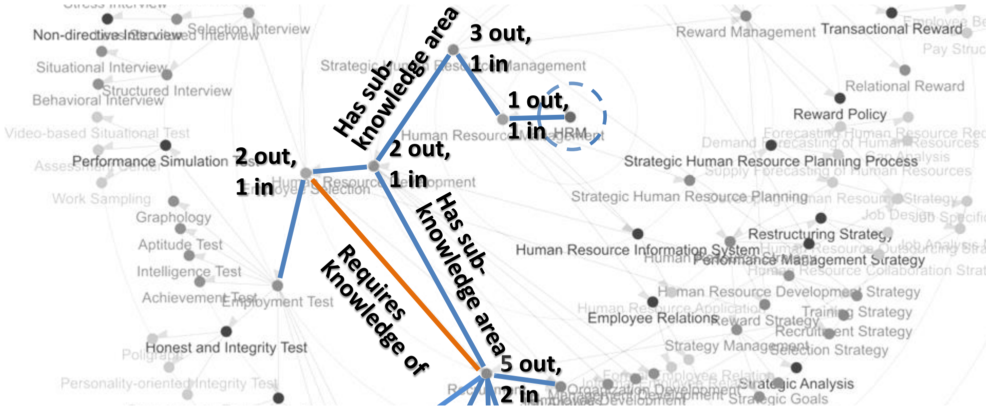


Figure 28: A learning-path through the domain, with concepts with a different degree and different relations to account for with the desired concept importance measure.

The aspect of "need" and "detail" of a concept can have a different weight based on the assumptions taken for the goal of learning. The goal is to derive a common measure which reflects that more needed and more detailed concepts result in a higher value of the importance dimensions and finally the combined measure, as pictured in the matrix in Figure 29. So, to combine the importance dimensions into one measure, representing the concept importance, requires to define how to weight them – according to their assumed relevance in the context of learning. To do so the measure considers three factors:

- **Dimension mapping:** importance dimensions to map for one individual relation.
- **Dimension weight:** specific weight of the importance dimension for the context of learning.
- **Sum of relations per concept**: the different relations of an individual concept.

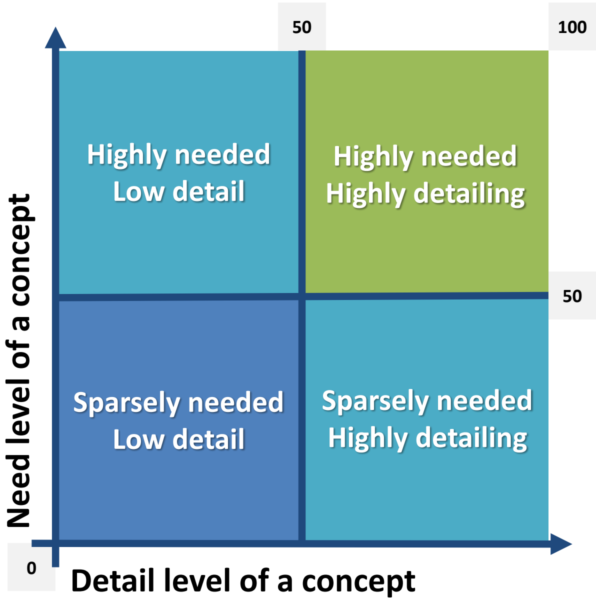


Figure 29: The concept importance matrix. More needed and more detailed concepts result in a higher value of the importance dimensions and finally the combined measure.

Connectivism follows the assumption that well-connected concepts are valuable connectors to access more knowledge and well-connected concepts enable more learning. This assumption can be extended: well-connected concepts more likely represent knowledge which is content-wise needed for surrounding concepts 1) from the network perspective – representing bottlenecks to other knowledge – and 2) from the perspective of understanding – representing central knowledge which is needed to master surrounding concepts, unifying the neighbourhood. The direct modelled need for other knowledge through an expert-modelled relation (e.g. the "requires knowledge of") can be interpreted as an alternative expression of this connectivity-based need, taking the viewpoint that an expert models an explicit requirement to express that a concept is needed for a range of other concepts, which again reflects a notion of connectivity.

From the perspective of behaviourism, fact knowledge and – more general – knowledge which "details" other knowledge, is "needed" to enable an effective learning process. In this regard, the identified importance dimension of "detail" can be interpreted as a remote substitute for the "need". While this work doesn't follow the idea of behaviourism, a fact or a detailing concept can be interpreted as more needed if it is connected to a range of other concepts.

Mapping the collected considerations onto the task of weighting the importance dimensions, the "need" is considered as the semantically stronger concept as it combines the connectivity-based "need" expressed by well-connected concepts and the expert's perspective of modelling essential concepts through different relations. The dimension "detail" is considered as the inferior indication in this context. Yet the "detail" has to be sufficiently accounted for to be able to "upvote" less-"needed", "detailed" concepts if they are well connected. Furthermore, the "detail" dimension cannot be under-represented within the weighting as it is expected that especially general concepts have many connections as they potentially unify multiple knowledge areas and less general concepts have to able to compete when being well-connected. Consequently, the weights will be balanced with the goal to prefer "needed" nodes but at the same time to enable "detailed" concepts with many connections "to catch up".

To enable a direct and intuitive modelling approach for weights, weights are defined on an interval of [0,100] for each dimension and are then set into relation to each other in the interval [0,1] by dividing each weight through the sum of weights. All normalised weights together then always sum up to 1 with the basic normalisation, where w_i is the dimension weight in focus and the denominator sums over all n available dimension weights w_j:

$$w_{i_{norm}} = w_i \Big/ \sum_{j=0}^{n} w_j \tag{11}$$

The goal and the intended profit of the specific weight selection and normalisation process is that further dimensions can be integrated on-demand without changing the expert's modelling intent (assuming the expert maintains a similar modelling intent throughout an extension, preserving a consistent weight modelling). Aligned to the gathered considerations on the importance dimensions, the domain ontology itself and related learning theories, the weights are selected as $w_{need} = 70$, $w_{detail} = 40$, resulting in a "detail" / "need" ratio of $\sim 2 \,/\, 3$, with:

$$w_{need_{norm}} = \frac{w_{need}}{w_{need} + w_{detail}} = \frac{70}{70 + 40} = 0.\overline{63} \tag{12}$$

$$w_{detail_{norm}} = \frac{w_{detail}}{w_{need} + w_{detail}} = \frac{40}{70 + 40} = 0.\overline{36} \tag{13}$$

To combine values and weights of the defined importance dimensions the weighted sum model is applied. The weighted sum model is a fundamental and well explored approach in the field of multi-criteria decision making (Triantaphyllou, 2000). The initial approach, captured in Equation (14), combines instances of weighted $(\mathrm{w_j})$, defined categories (c_{ij}) over which alternatives (A_i) are rated and finally combined to one single value per alternative (the higher the better). In case of the importance dimensions, the dimensions ($Rel_{\mathrm{dim}_{ij}}$) represent the categories for the weighted sum model, while the global weights of the categories are represented by the pre-defined weights (w_j) of the importance dimensions.

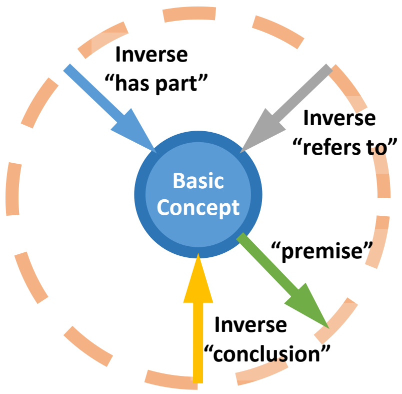
The weight/dimension pairs are summed over the m importance dimensions of each of the n relations of a given individual concept for which the concept importance measure ($Imp_{Concept}$) is calculated, as shown in Equation (15). Figuratively, each relation can be interpreted as an alternative route for learning, evaluated over the defined criteria of learning (“need”, “detail”) and the sum of these alternatives (the integrated connectivity) is the potential for a better learning, which is defined as the concept importance measure.

$$\mathrm{A_i} = \sum_{j=0}^{m} w_j * \ c_{ij}\,, \qquad for\ alternatives\ i = 1, 2, .., n \tag{14}$$

$$\mathrm{Imp_{Concept}} = \sum_{i=0}^{n} \sum_{j=0}^{m} w_j * \ Rel_{dim_{ij}}\,, \qquad for\ m\ dim.\ and\ n\ rel. \tag{15}$$

Figure 30 shows an example concept with relations to calculate the concept importance measure on. The calculation is done from an outbound perspective in terms of the relations. So, the relation “premise” announces that the “Basic concept” points to a “Theorem” as its “premise”. The “Basic Concept” highly “details” the “Theorem” and is highly “needed”, so from the outbound perspective of the “Basic concept”, rating the “Basic concept” a high “need” and high “detail” is accounted (the low accounting happens at the “Theorem”). All inbound connections are considered based on their inverse relation “converting” them to an outbound relation, based on Table 26. Based on the composition of relations, the measure is calculated as:

$$\begin{aligned}
\mathrm{Imp_{Concept}} = {} & \mathrm{w_{detail}} * premise_{detail} + w_{need} * premise_{need} \\
& + \mathrm{w_{detail}} * isPartOf_{detail} + w_{need} * isPartOf_{need} \\
& + \mathrm{w_{detail}} * invConclusion_{detail} + w_{need} * invConclusion_{need} \\
& + \mathrm{w_{detail}} * referredBy_{detail} + w_{need} * referredBy_{need} \\
& = 0.\overline{36} * 90 + 0.\overline{63} * 80 \\
& + 0.\overline{36} * 80 + 0.\overline{63} * 60 \\
& + 0.\overline{36} * 90 + 0.\overline{63} * 80 \\
& + 0.\overline{36} * 50 + 0.\overline{63} * 50 = 133.\overline{63}
\end{aligned} \tag{16}$$



Relation	Relation “is needed” (0-50-100)	Relation “details node” (0-50-100)
premise	80	90

InvRelation “details node” (0-50-100)	InvRelation “is needed” (0-50-100)	Inverse Relation
80	60	is part of
90	80	inv. conclusions
50	50	referred by

Figure 30: Concept importance example case for calculating the measure.

Revisiting the importance dimension mappings and the mappings of the relations, the defined concept importance measure can distinguish nodes with one or more relations. An example is given in Table 27, calculating the concept importance measure for a node with one outbound relation. The relation pair “premise” / “conclusion” yields the same result if calculated for a single relation and for one of the relations the inverse relation is considered as “inv. premise” / “conclusion”. In the frame of this work this is an intended modelling outcome to account for the mirrored nature of the relation.

Table 27: Concepts are already distinguishable based on a concept importance measure for nodes with single relations.

Relation	need	detail	w_{need} * need	w_{detail} * detail	sum
has part	40	20	25.45	7.27	32.73
has sub-knowledge-area	45	35	28.64	12.73	41.36
premise	80	90	50.91	32.73	83.64
conclusion	20	10	12.73	3.64	16.36
refers to	50	50	31.82	18.18	50.00
requires knowledge of	0	35	0.00	12.73	12.73

Considering the logic of the concept importance measure a wide range of modifications in the presence of more information is possible. As gathered earlier, additional dimensions are possible, importance dimensions can re-weighted on demand and different sets of dimension mappings for relations can be stored to be applied on demand e.g. to support the adaptive behaviour of an implementing system. A dimension extension is proposed as part of the measure integration starting in Section 8.4.4 to account for the learning progress and better account for the modelling intent of concept groups.

8.4. System Integration of the Concept Importance Measure

To integrate the measure into the STUDIO system, the system hast to be extended in three main pillars:

1. **Concept importance measure module:** the system will be extended by a new module to derive and manage the concept importance measure to enable a system-wide seamless use of the measure. The module will integrate a parametrisation interface to adapt the concept importance measure for different scenarios of assessment and evaluation.

2. **Path based assessment:** To utilize the concept importance measure a new path-based assessment algorithm will be created. The path-based assessment enables to create paths through the tailored domain ontology used for the assessment. From the created paths, concepts can be selected on demand, following different specific assessment strategies – like the selection based on the importance measure to select the next important concept for assessment and learning. To utilize the new concept importance measure, the implemented path extension will be used to integrate the new measure and to customize the assessment and later evaluation algorithm.

3. **Path-based evaluation:** A concept importance algorithm follows paths but will select the next concept to test within the same path based on the highest calculated value of the measure for a specific concept. Furthermore, concepts could be connected through different paths to the respective root concept and depending on the complexity of the structure the next “important” element may be accessible through multiple paths. The algorithm will make use of assessment paths and implements the introduced evaluation strategy.

To address the intended three pillars first a path-based assessment and evaluation strategy will be introduced universally. Based on the path-based assessment different assessment and evaluation strategies are possible, in particular a concept importance based assessment and evaluation.

The proposed path-based solution for assessment and evaluation will be followed by a pre-evaluation using a less complex bottom-up assessment and evaluation. Following the trial finally the concept-importance measure will be integrated and replace the bottom-up logic for a concept-importance based assessment and evaluation.

8.4.1. A Path-based Exploratory Knowledge Assessment

The initial algorithm for assessment and evaluation integrated into the STUDIO system, explores the knowledge structure in a top-down manner. In this regard the system “drills-down” from the general to the detailed, more factual knowledge – in line with the idea of constructivism and following a process related logic. To account for the behaviour the initial algorithm is called the “drill-down” algorithm. To create a different solution which explores the structure freely – as needed for a concept importance measure implementation – the system has to be extended with a new open algorithmic logic. The extension will act as a potential fundament for different smart knowledge exploration and evaluation solutions. A new knowledge exploration framework will be developed, which enables the implementation of various new assessment and evaluation approaches in STUDIO. In this regard the extension will be, beside the final implementation, also a theoretical extension to the concept of the STUDIO assessment and learning system.

A challenge is to ensure the to cope with the complexity of the semantic domain model and the assessment paths and how to translate the handling into the existing system. Furthermore, the new approach has to solve specific algorithmic problems as e.g. loop-prevention and solve conceptual problems as e.g. defining the criteria for evaluating and terminating the assessment. A summary of the concepts needed here with an event-based perspective is published in (Weber, 2014) and the approach for a path-based assessment is reported in (Weber, 2016a).

One of the major benefits of the domain ontology supported self-assessment is the potential to use the knowledge structure to see single answers to questions in the context of the “bigger” knowledge structure of the ontology. If the system stops in a too early stage of the assessment this may discourage students to retry the test after gaining more insights, as it slows down the exploration of the knowledge structure. Different cases are reasonable in which it may have sense and reason to continue to ask questions for related knowledge areas, even after a sequence of wrong answers. E.g. a student may not have the overall knowledge of an educational area but may already know perquisite knowledge areas, which are sub-areas to the current test, or the test questions for more general knowledge-elements may be biased and slow down the „drill-down”. In both cases an exploration of knowledge-elements deeper in the knowledge structure may shed light on different further improvement potentials of the student.

A solution to explore deeper the knowledge of the student comes in the form of assessment paths. For the initial „drill-down” assessment and evaluation algorithm – presented in Section 8.1.6 – the criterion for evaluation is based on the second assumption for testing, presented in Table 12 of the same section, expressing that continues failing on general concepts forecast failure on detailed concepts. To promote a different exploration of the knowledge structure within the assessment, and to trial the path-based assessment the drill-down thesis will be replaced by a bottom-up thesis. The bottom-up assessment and evaluation will act as a reference implementation for the path-based assessment and evaluation and is conceptualized in a way that it can be reconfigured and reframed for other path-based strategies to traverse the domain.

Using this changed assumption about the structure exploration and their evaluation, finally the new concept is introduced: assessment paths. An assessment path is a set of knowledge-elements (concepts) which are connected by relations while the path is connected with one knowledge-element to the start-element. With this new approach the tree shaped knowledge structure, which is extracted for each assessment, could be interpreted differently, as a set of any possible paths from one knowledge-element to the start-element.

An example for a path is shown in Figure 28. These paths are routes through the knowledge structure. How far a student succeeds on a path, marks the personal knowledge of a student, while the set of succeeded paths highlights knowledge-areas in which the student excels. Beside the existing feedback on the depth the learner reached within the structure, the new assessment also yields the potential to group the assessed knowledge, based on aggregating single passed concepts to sets of finished and unfinished paths.

In the chosen bottom-up (reference) implementation of the path-based assessment, in contrast to the drill-down testing, the assumption is that learners will know details about the represented domain, even if they cannot answer questions for high level, general concepts. A failing on earlier concepts may be based on a yet missing comprehension of correlations and consequences – which is reflected on the mirrored dependency assumption, phrased in Table 28 – while detail knowledge is already known by the learner.

Table 28: Necessary assumption for traversing the knowledge structure in a bottom-up, path-based setup for assessment.

Assumption	Description
Bottom-up knowledge dependency	If a test-taker fails on more detailed concepts the system will assume that he or she will also fail on more general concepts. If a test-taker fails on more general concepts he or she could potentially still succeed on more detailed concepts. Resulting, each knowledge-area or element is relevant for the main test goal. As higher level concepts are comprised of an aggregation of detailed concepts, sub-level concepts have to be tested as complete as possible to explain failing on higher levels.

Further questions for high level knowledge elements may be considerable harder to phrase and create as they have to represent a trade-off between size, concept dependencies and the numbers of concepts needed to make a statement about the core of a concept and its implications. As such the probability for flawed or biased questions on higher levels is higher than for detailed concepts.

Taking a process related view, learners may have sufficient knowledge to fulfil tasks of the target job roles, attached to the processes but may lack a higher-level understanding of the reasoning behind processes. Further they may already have an understanding of target processes, paired with the power to apply it in daily life situations, fitting to the classification of the blooms taxonomy (Krathwohl, 2002), but yet lack the analytical proficiency to transform them to the specific question.

As such it has rational, especially in early learning cycles, to start to assess more detailed knowledge first to create an understanding of the current skill level and the compliance to processes, rather than stopping the assessment on high level concepts which are hard to decompose to derive learning feedback. A solution to explore the knowledge of learners broader and more detailed in a bottom-up approach – while defining an evaluation frame – comes in the form of the creation of assessment paths.

Assessment paths are a generalization of the concept of connected knowledge elements and describe paths through the knowledge structure which connect one knowledge element to the respective start-element. A path can thereby include an unlimited amount of intermediate knowledge elements which are needed to connect to the start-element. To prevent loops in the directed graph, the final algorithm makes use of a strategy to black-list visited nodes, combined with a backtracking algorithm to continue to create and explore alternative paths.

To enable the new path concept, the STUDIO assessment assumptions about the structure, based on the drill-down algorithm, have to be modified and extended, resulting in the assumptions phrased in Table 29. The bottom-up algorithm will start from bottom knowledge-elements. As such a path from a passed element to the start-node may include failed elements. To cope with it the assessment and evaluation are done in separate ways for the bottom-up assessment, as reflected on Table 29. Following the revised assumptions, passed elements will be only accepted if they are connected to a path of other passed elements, connecting without interruption to the start-element.

Table 29: Revised assumptions for a path based assessment (with a bottom-up specialization).

Assumption	Description
Extended Ordering for Bottom-up Paths	All knowledge areas are connected either with "part_of" or equivalent relations or with "requires_knowledge_of" relations. Resulting, for every tailored concept group, paths, starting with a start-element, will develop on average from general concepts to detailed concepts. To sufficient explore the knowledge structure for each set of knowledge-elements, reachable through a path of connected relations, the test will first select knowledge-elements which are connected through the highest number of intermediate knowledge-elements to the start concept.
General Path / Knowledge Evaluation Assumption	If a test-taker fails on more general concepts he or she can still succeed on more detailed concepts. Multiple correctly answered paths may connect to the same correctly answered concept. Resulting, each knowledge-area or element is relevant for the main test goal, if there is a path of knowledge-elements to the start-element, which includes only knowledge-elements which are marked as passed.

8.4.2. The Path-based Assessment and Evaluation Strategy

With the new concept of assessment paths, the tree shaped knowledge structure extracted for the assessment, describes a set of possible paths from each knowledge-element to the start-element. These paths represent routes through the knowledge structure and traces how far a learner masters connected concepts. They may show further how complete are the capabilities to meet certain learning outcomes, while the set of succeeded paths together mark knowledge-areas in which the learner excels.

The trade-off for the use of paths comes in the shape of an additional algorithmic costs. The loading and creation of paths has to be partially pre-fetched and – based on the complexity of the structure – a multitude of different paths could be complex and expensive in extraction. To lessen this resource strain, the new algorithm makes preemptive use of the knowledge about the evaluation of paths. Only paths between a node and the start-element which include no failed knowledge-element will be evaluated. Resulting, each failed knowledge-element leads to an automatic “block” and removal of the path below the failed element (including the failed element) for the current path. As such, every alternative path which only connects through a “blocked” concept to the start-element will be omitted, as elements can be only evaluated if there is a path to the start-concept open. So, while the algorithm explores the knowledge structure path by path, it successively decreases the set of unexplored paths by evaluating them and at the same time by eliminating them based on blocked concepts.

Figure 31 shows the creation cycle of paths. The input sub-ontology is based on the concept group-based extraction process, as described in the previous Section 12.3.1. The resulting sub-ontology delivers the structure which is used to run the assessment.

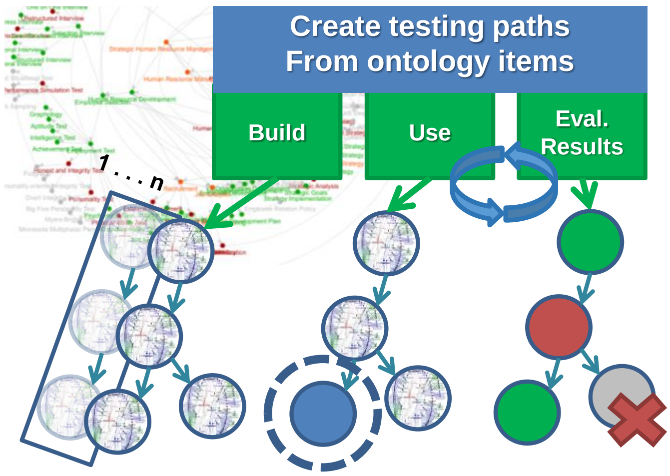


Figure 31: Path based assessment process, succeeding the sub-ontology extraction process.

The path creation over the course of the assessment will continuously trigger three steps:

1. **“Build”** extracts fitting knowledge-elements from the knowledge structure, starting from the start-element and combining them into paths.
2. **“Use”**, triggers the central assessment which then assesses the path, concept by concept, based on the connected questions – in case of a bottom-up assessment from the bottom to the top element.
3. The step of **“Store results”** is a concurrent process, storing the success or failure of knowledge elements and cuts the path of at elements which are marked as failed.

So, if any element fails, it will be marked as failed and block the later part of the current path to the start-element. The block is twofold: 1) no element below the blocked concept will be tested for this specific path (which is relevant for a non-bottom-up assessment) and 2) if later a new path is created, the creation process will be stopped for the path when encountering a blocked concept and the last added concept will be the end of the current path. With this cut-off procedure, the algorithm minimizes the set of future sub-paths to assess. As designed, the system accepts every path of knowledge-elements, with any length, which reaches the start-element through offered relations and passed knowledge elements. Resulting, a failed knowledge-element splits a path into a “top” part which could still reach the start-element and a failed bottom part which won’t be considered.

Figure 32 visualizes the assessment and evaluation process in more detail. In the first 1) stage a bottom concept of the path is selected for assessment – which for the concept importance based assessment could be selected by a different criterion. In step 2) the earlier concept is passed and the next concept is selected. In contrast to the drill-down algorithm for evaluation no multi-concept pass-threshold is required to be passed. This time the learner fails to answer and the concept is evaluated as “failed”. As the concept also connects another concept, which could be added later for a new path, it is considered as “blocked”. So, the small structure used in this example includes two paths – one path which includes the left and one which includes the right bottom concept. The other concepts are shared in both paths.

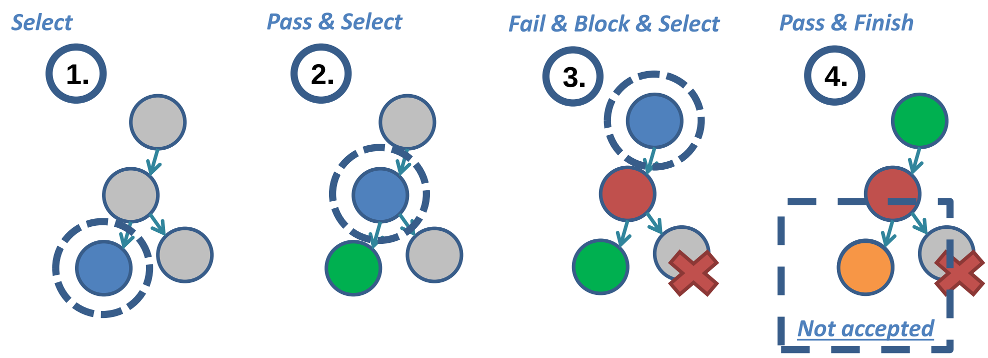


Figure 32: Detailed path creation and blocking process.

Assuming the example of four concepts the “blocked” concept eliminates the second potential path, as shown in 3), and reduces the search space. In step 4 the top concept is passed and marked as passed. Now follows an additional stage of evaluation. Only concepts which are connected through a path of passed concepts to the top element are "accepted" for the final evaluation. So, even though the learner answered 2/4 of the concepts correct, only 1/4 are accepted, so the final result is 25%.

Figure 33 shows the result summary visualization of the bottom-up self-assessment test. Red/dark dots signal knowledge-elements which the learner failed, while green/light dots identify knowledge-elements which are passed. The image visualizes efficiently the potential to reason on cleared and not yet cleared areas of the domain. While some concepts are known, higher level knowledge-elements could not be passed and mark sections for further learning. Orange concepts mark concepts which were successfully answered but couldn't be accepted as no path with passed concepts could be found to the middle "top" concept. The brackets in Figure 33 already display normalized concept importance measure values. The implementation is addressed in later sections.

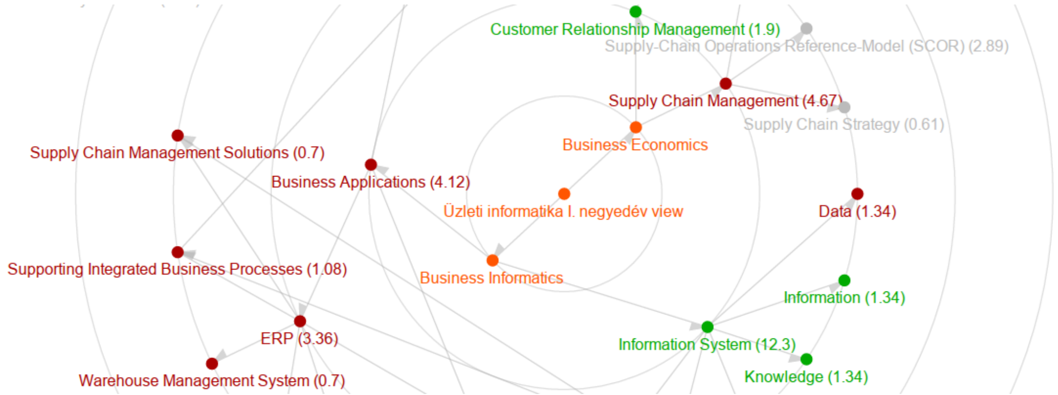


Figure 33: Result visualization as educational feedback for the learner, based on the path-based evaluation.

The path exploration and evaluation-based blocking is one way to control for the complexity of the domain and the assessment alike. This way also potential loops are minimized as every branching of the structure triggers the creation of a new (overlapping) path and each path is created "just-in-time" of its use and accounts for a) blocked concepts, which cut off paths and b) already asked concepts are not asked again. Still a part of the search space for paths still allows the creation of loops where paths "point" back to themselves.

The solution for this bottleneck is that it is not allowed to add the same concept to the same path twice. So, for the first iteration the path will include all the concepts till the point where the loop enters the existing path again and after the assessment the concepts in the path will assessed (or even blocked) and the next path will be created including a different sub-branch of the domain. As the creation of the paths (in contrast to the assessment) follows only directed relations in their intended direction, the building of paths is unique.

The core path-creation algorithm is implemented in two straightforward loops. When a concept was failed, or has no other concepts to go to – either because there are none or because all others are “blocked” – then the concept is “closed” and won’t be touched any more. The closed state can thus be considered as “infective” and over time all concepts will be “infected” and the assessment or evaluation is over. The loops work like this:

1. **Removal loop:** Take the path and travers it top-down. If the concept in focus has no connections to non-closed concepts, then close the concept. If the concept is closed, then remove all further concepts from the path. Else look at the next concept in the path.

2. **Add loop:** If the path is empty then add the start concept. Select the last concept in the path. If the concept has a connection to a non-closed concept, then add the concept and consider it next to add more concepts.

If after both loops there are concepts in the path, then start the assessment of the next concept. If the path is empty, then all the accessible concepts (some may be blocked) are asked and the assessment ends – which is the finishing criteria. The loops are running after each assessment / evaluation of a concept. So, after each iteration the path is different and adjusting to the set criteria. One major profit is that not only the exploration and the loop-prevention is solved this way, also for a complete evaluation or re-evaluation of an existing test, the structure only has to be pre-initialized with the information passed / failed and the system evaluates the given test through running the assessment again while using the known answers to evaluate / close the concepts. This way evaluations for different answer sets can be simulated on demand.

8.4.3. Experiment on the Application of a Path-based Strategy for Assessment

This experiment will cover the extended, path-based assessment and evaluation framework – based on a bottom-up example implementation – and is conducted on the Business Information Systems information systems bachelor course, which ensures a comparability to other experiments in this work. The learner used the system to prepare for the mid-term exam in March and the final-term exam in April 2015.

The study will have a twofold focus – to evaluate the new path-based assessment and evaluation framework, and to get a comparative insight into the influence of different testing strategies on the assessment of the learners. Students learned using the drill-down assessment and using the path-based “bottom-up” assessment approach, which applies a behaviourism driven implementation (bottom-up - facts first as an average tendency). The concept of assessment paths is used to connect concepts through paths of connected concepts to the root concept of the knowledge structure. The experiment builds on the lessons-learned in the system selection and use variables which were explored within the architecture extension of STUDIO in Section 8.2.

A challenge is that the notion of “direction” in learning is connected to specific learning theories. As such, the derived feedback will incorporate different factors influencing the results of the assessment. Further a connection to the later behaviour in the learning module will depend on a sufficient availability of qualitative and quantitative feedback. The top-down assessment is based on an existing solution within the STUDIO system and the bottom-up assessment with assessment paths is published within (Weber, 2016b) and the first level results are summarized in the publication (Weber and Vas, 2016b).

8.4.3.1. Analysing Drill-down and Path-based Bottom-up Testing Within a Course on Management Information Systems

The field study explores and evaluates the path-based bottom-up assessment in contrast to the drill down assessment and evaluation approach. The study took place in a blended learning environment, where students had access to the extended STUDIO system and to traditional learning material and the regular bachelor’s seminar in the field of Management Information Systems. Based on the STUDIO performance the students had access learning objects, provided through STUDIO. The study had two stages: students used the system with a drill-down implementation throughout 14 days to get prepared for the mid-term exam and with a bottom-up approach throughout 13 days, a month later, to get prepared for their final exam. 287 students took part in the top-down test (61,897 tested knowledge elements) and 213 in the bottom-up test (25,919 tested knowledge elements). In the first stage the students had the motivation to prepare for the mid-term exam (the concept group were tailored to the mid-term topic range). In the second stage an additional incentive for students were provided in the form of extra points for the final grade. Two third of the students from the first stage also joined the second stage of testing.

Figure 34 describes how many times each knowledge element had been tested using each of the two approaches (the larger and darker the circle, the more times the element had been tested), while Figure 35 accounts for the number of times a knowledge element was passed across all test runs. The graph visualisation of the visited and passed knowledge elements in Figure 34 and Figure 35 traces the different exploration of the two algorithms. For the top down, elements are visited more frequently when they are near to the start elements in the centre, the focus on the right part tributes partially to the clockwise selection of initial nodes.

In case of the bottom-up testing, bottom elements are visited more frequently and more equally, which partially goes back to a stronger random selection component (if a path can branch in different directions, the direction is chosen at random). The sets of points within the graph are scaled within the respective testing algorithms to compare the distribution, so Figure 34 (a/b, scale [60,3300]) and Figure 35 (a/b, scale [10,900]). The overall pass/fail distribution among knowledge elements is 73.18%/26.82% for top-down and 69.87%/30.13% for bottom-up testing and in this regards comparable.

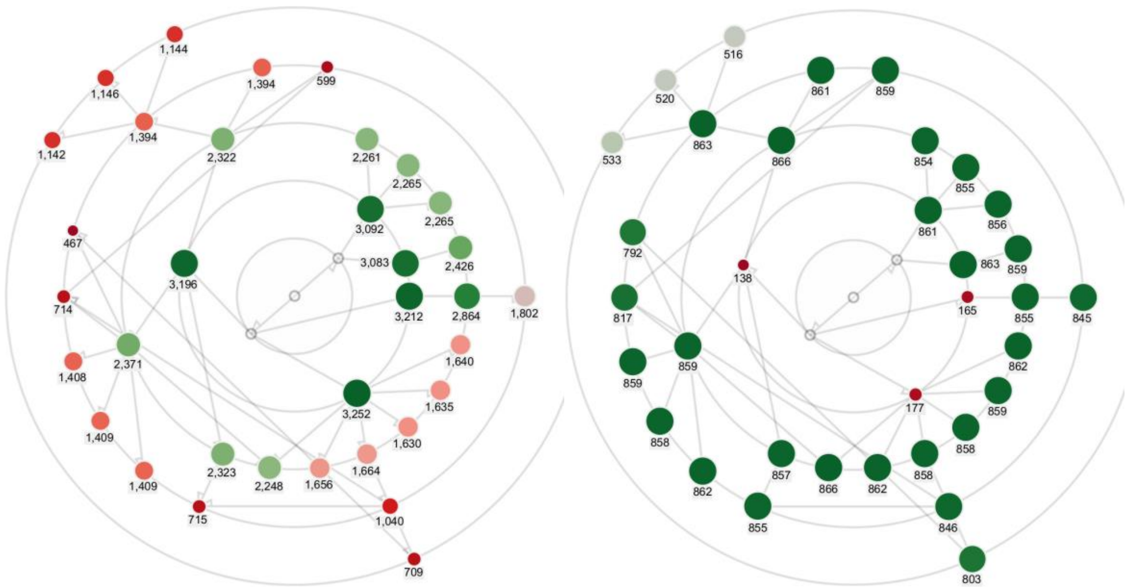


Figure 34: Aggregation of how frequent a knowledge area was visited for top-down (a/left) and bottom-up (b/right) testing visualised on the course's knowledge structure (see Figure 1). Each graph is scaled based on its own internal distribution.

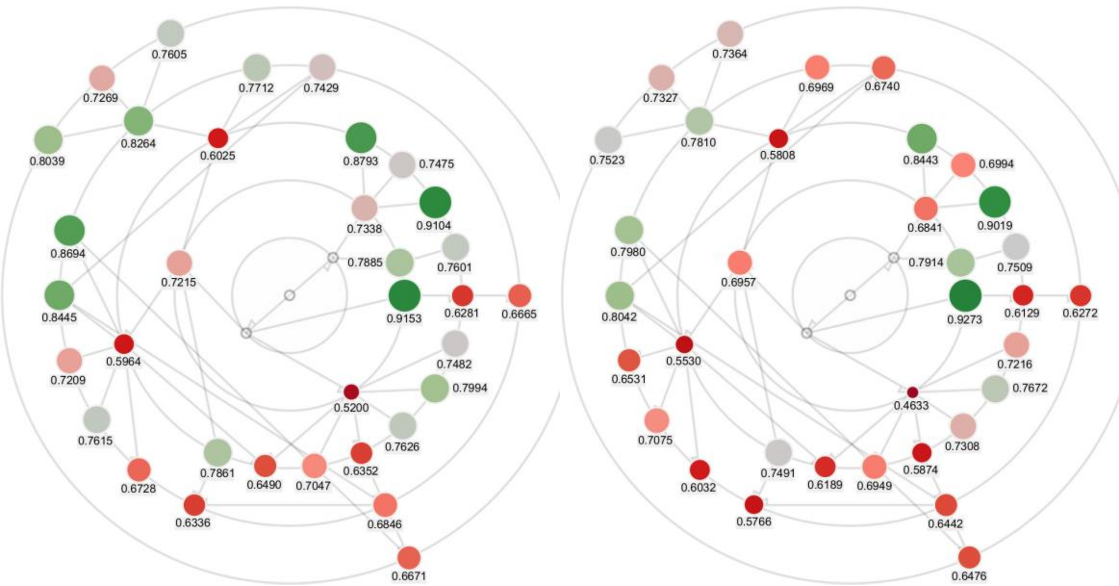


Figure 35: Aggregation of how frequent a knowledge area was answered correctly for top-down (a - left) and bottom-up (b - right) testing. Each graph is scaled based on its own internal distribution of passed elements.

For Figure 36 and Figure 37, only the first and the last test for each student were considered to trace an overall trend across tests. The x-axis shows individual students, sorted by the performance of their last test, where the right part picture higher performing students. What is visible here is that the drill-down test in Figure 36 has a low and flat trend for the rate of passed nodes across first tests.

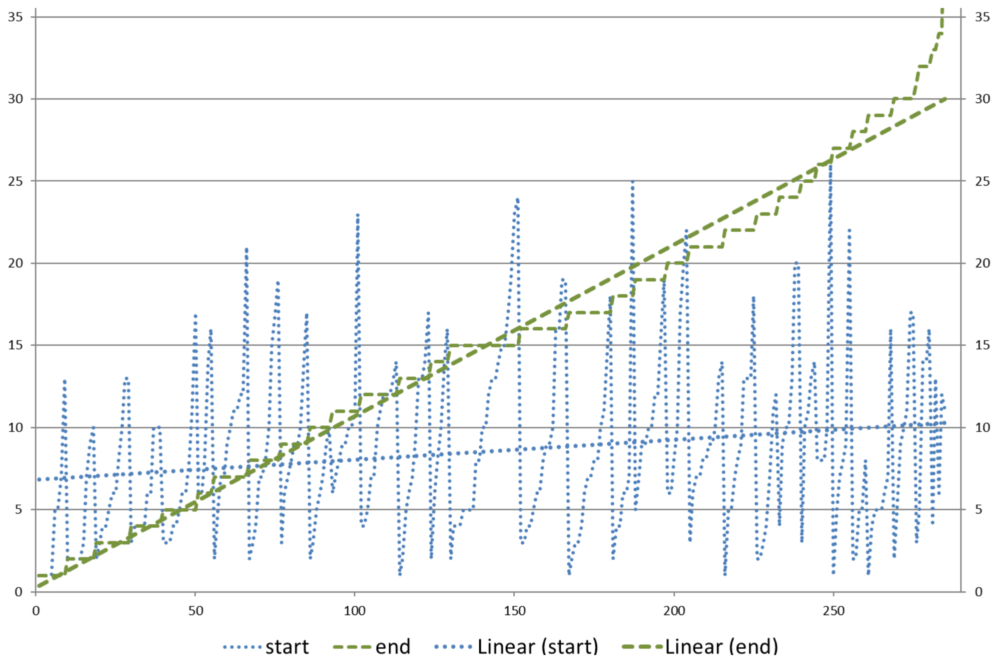


Figure 36: Amount of knowledge elements passed in the first and last test of a user within the drill-down approach, sorted by the amount of knowledge elements passed within the last test.

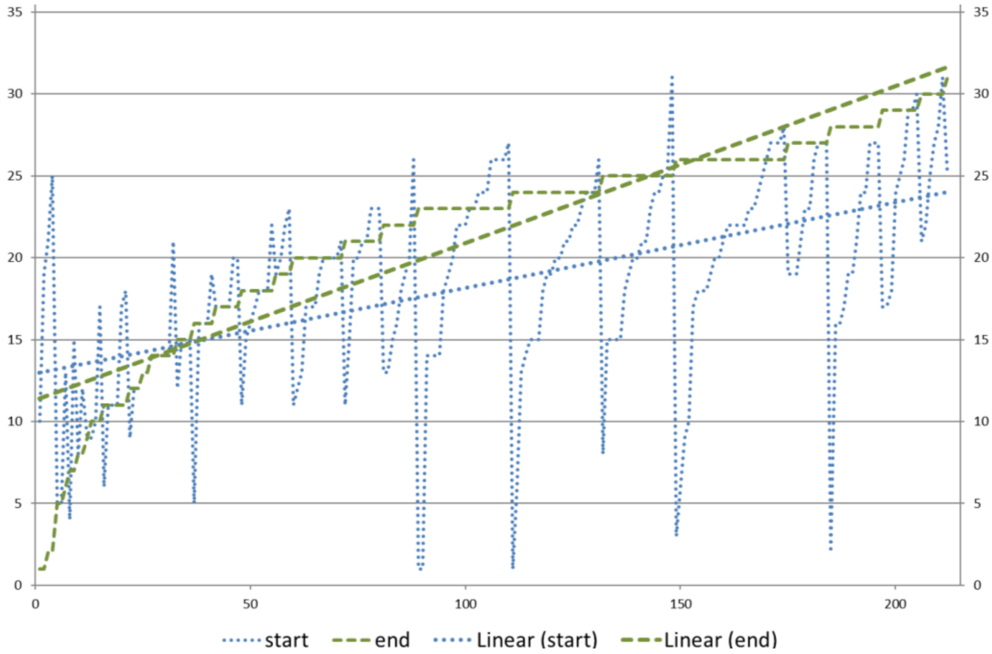


Figure 37: Amount of knowledge elements passed in the first and last test of a user within the bottom-up (b/right) approach, sorted by the amount of knowledge elements passed within the last test.

In a direct comparison to the bottom-up testing, the drill-down tests show in average a higher performance boost till the final test, yet the bottom-up test starts with a higher average pass-rate and rises lighter and more stable across all users, with a similar std-deviation of 6.16 for the start and 6.34 for the final number of passed nodes, against 5.21 and 8.70 for the drill-down approach. The average of the std-deviations of all tests for each user is 4.97 for the drill-down and 2.79 for bottom-up testing. So, within the drill-down testing, the performance of passed nodes changes in average stronger than within the bottom-up testing. This observation is especially of interest as the higher number of observations within the top-down testing (more observations) should smooth the variation of the test results.

8.4.3.2. Drill-down and Path-based Bottom-up Testing Comparison Considerations

This study addresses and compares the drill-down and bottom-up implementation of the STUDIO educational test and investigates how different the outcomes of the specific approaches are. For the drill-down/bottom-up testing, the overall pass/fail level for high performing learners – taking into consideration every answer regarding every knowledge element – is on a comparable level. Yet the average improvement in the test takers' performance is different, starting on different performance levels and showing different gradients towards the high performing groups of learners. The drill-down approach encourages higher results from high performing testers on the costs of more low performing testers, while the bottom-up approach tends to stronger equalise the performance.

Technology enhanced solutions for learning and testing are expected to have an initial phase of familiarisation of the tester, resulting in lower initial tests, where the learners "learn" how to use the system, which will account for a part of the overall low starting point for the drill-down assessment. As the bottom-up test were a repeated test scenario, this initialization "bias" will play a lesser role. It is likely that the influence of the different stopping-criteria for both algorithms influences the pass rates within test runs and create special "early"-finished test runs, which may account for the different std-deviation across the drill-down assessment.

The strongest factor accounting for a part of the difference will be the repetition of the test phase with STUDIO by students which took part in the first drill-down stage and later in the bottom-up assessment stage. This will account partially for the higher average start/end performance in the bottom-up group of students, yet it doesn't explain the strong "bend" between low performing and high performing student groups across the bottom-up tests. Future analysis on more extensive knowledge structures may here reveal further insights and help to better distinguish the core performance of testers from other factors.

A likely factor for the difference between the student groups across the two assessment and learning approaches will be the different exploration of the path-based bottom-up test, which starts with more concepts in early iterations of the test and explores more freely the tailored domain. While a deeper exploration of the differences will depend on repeated studies with parallel groups of students [11], this study shows independent of the existing influencing factors that a path-based assessment "performs" comparable to the drill-down assessment and can deliver different results and different learning gains.

8.4.4. A Concept Importance Based Knowledge Assessment Algorithm

The integration, implementation and utilization of the concept importance measure will be based on the path-based assessment and evaluation defined in the previous Section 8.4.1, which were proposed and implemented as a flexible starting point for a range of

[11] This study doesn't feature an explicit A/B testing scenario (which would have been recommendable) because the path-based implementation was only available to support the final exam preparation.

algorithms. This part will address the integration of the concept importance based assessment, the evaluation, and introduce an additional extension to further grain the granularity of the selection within the assessment.

The path-based assessment introduces paths to provide a trade-off between the exploration – in terms of the assessment and learning – and the complexity of the domain and the underlying domain model, which may impact the computation of complex approaches for assessment and evaluation. Every path represents one possible “walk” through the domain, starting with the start-concept and ending in another concept. Along the walk, the path “connects” concepts which are relevant for the current part of the assessment. The path is that way representing a closed unit which separates the assessment and evaluation into smaller steps. Every path is considered separately across the assessment, while within every path, all concepts are considered at once for the active algorithm. The path can be considered as a window for the algorithm, its respective needed calculations, and the specific selection of the concepts on the path, while the path length is a flexible window size. The overall assessment process progresses from one path or window to another.

For the integration of the concept importance measure – to select concepts based on the measure – the path generation of the path-based assessment will be used. For a reference, see Section 8.4.2. For every created path, the bottom-up reference implementation selected concepts for the assessment, starting from the concept which were furthest away from the start-concept – and thus considered as the “bottom” concept. In contrast, the integration of the concept importance based assessment will select the concepts from an active path, based on the concept importance measure.

Concepts which rate higher in terms of the concept importance value are considered as higher important for the learning of concepts in a given domain. Following the logic, the concept which has the highest importance value within an active path, has to be selected and assessed first, followed by the next less high value, etc. So, within an extracted path the concepts are ordered for the assessment based on the concept importance measure. An example is given in Figure 38. Above each concept in Figure 38, the concept importance based value is given as the base for sorting (the values are examples) and below the concepts the resulting assessment order is shown.

Concepts in a path which a learner failed can change the order of the assessment by blocking remaining parts of the path. This is an expected feature of the path-based framework, yet it differs for the concept importance based assessment in contrast to the bottom-up reference implementation. The final selection of a concept to assess is based on the highest value of the measure per (un-assessed) concept in a given path. Thus, the failing of a concept will change how the domain is explored. Furthermore, it will also change the selection itself as the algorithm selects concepts from each path by "jumping" based on the measure. These jumps will result in cases where the order of testing in a given path is already derived but the path closes and cuts off concepts which were expected to be asked next. The list of next concepts is then automatic rearranged as a part of the concepts isn't available (in this path) anymore, as shown in Figure 39, thus limiting the search space for assessment and consequently learning and further addressing the desired adaptivity.

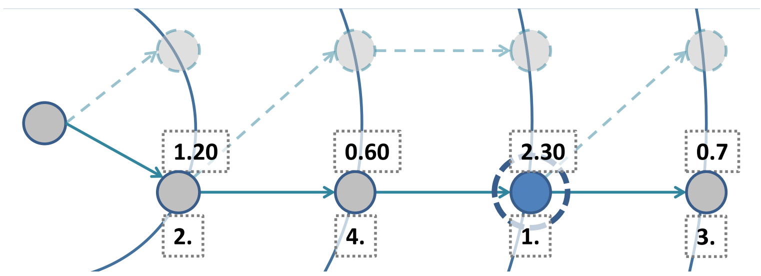


Figure 38: Paths are ordered and assessed based on the individual concept importance measure.

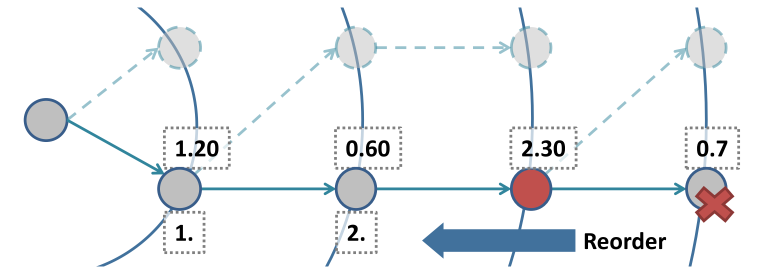


Figure 39: Failed concepts can trigger a re-ordering of the remaining concepts in the path for the further assessment when parts become unreachable.

The evaluation of the concept importance based algorithm will re-use the initial evaluation of the path-based assessment, which is also used for the bottom-up reference implementation. The process is described in Section 8.4.2. Every "passed" concept is considered, while only concepts which connect through a continuous path of passed concepts to the start-concept are "accepted" and used to calculate the final performance ("accepted_per_all"), as visualized in Figure 33.

8.4.5. Extending the Concept Importance Measure with a Distance-Based Importance Dimension

In the proposed implementation of the concept importance measure, concepts will have the same concept importance measure if they have the same relations. Especially concepts which are at the end of paths may connect through the same relations, as the tailored domains tend to end in sets of multiple, equal connected concepts, connected to the same parent concept. An example is shown in Figure 40. All concepts with the same measure will be considered equal and cannot be differentiated by the current concept of the concept importance based assessment. In this case concepts will be selected at random from the list of equals, yet a further differentiation can be derived from the structure.

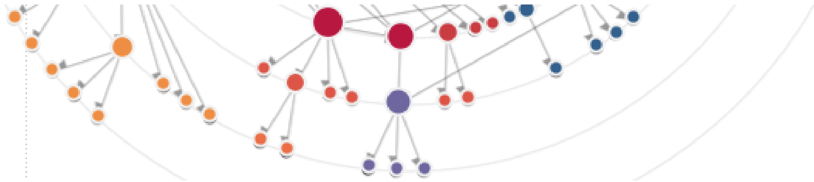


Figure 40: "Bottom" elements of a concept group, where – as a tendency – multiple concepts connect with the same relation to one single parent concept.

The concept of the assessment algorithm is to differentiate and select concepts – in the case of the concept importance, based on the individual value of the measure. So far, the measure exploits the semantic and the connectivity of the domain ontology. Yet another aspect of the structure can be exploited – the extracted paths.

The tailoring of the domain ontology through concept groups expresses a modelling intention: To frame and order the domain to a set of concepts which represent well the specific domain for the intention of learning. The start-concept is considered as the entry-point for the concept group. All concepts in a concept group can be considered as "detailing" the start-concept – independent of their individual type of concept or relations. The further away a concept is the more it details the central concept. This distance can be utilized to further distinguish concepts. To consider this "detailing" trend, an additional factor will be added to the concept importance measure.

The overall "detailing" trend of the concept group – based on the concept distance to the start-concept – can be expressed using the assessment-paths, created for the latter assessment. The idea is that, based on the distance of a concept to the top-concept in a path, concepts with the same concept importance measure can be distinguished, if they are in different locations of the domain. Equation (15), summarizes the concept importance measure. To implement the extension, the equation will be extended by an additional weigh and relation pair for the "detail-distance", with: $w_j * Rel_{dist\_detail_{ij}}$.

The core Equation (15) doesn't change, as the sum over m dimensions and n relations is defined for any number of weight and relation pairs. Also, the set of weights is extended by adding to the pre-set of "70" for the "need" and "40" for the "detail" dimension, "5" for the "detail distance". The low additional weight of "5" reflects that the extension should help to differentiate concepts, rather than to substantially change the resulting values. All importance dimension weights are re-normalized as shown in Equation (12) and (13).

By considering the path-length, the "detail distance" resembles a counter. As such it is defined in an open interval of [0, n], where n is the number of concepts between a concept in a path and the start-concept. The remaining importance dimension mappings are defined in an interval of [0, 100]. To combine the ranges of the intervals, the "detail-distance" interval is normalized using a features scaling or min-max scaling. The normalized value is then multiplied by the difference between the minimal and maximal expected mapping value with $(Dim_{max} - Dim_{min}) = 100$, to fit to the other importance dimensions (shown in Equation (17))

$$Concept_{dist\_detail_{norm}} = \frac{\mathrm{conc}_{dist} - conc_{\mathrm{min}\_dist}}{\mathrm{conc}_{\mathrm{max\_dist}} - conc_{\mathrm{min}\_dist}} * 100 \tag{17}$$

Another alternative to further differentiate concepts, is to integrate the betweennes measure as an extension to the need dimension. The betweennes centrality measure (Section 6.1.4) rates concepts high which are part of many shortest paths between two concepts. As such the measure rates a concept as high if it is a gateway between clusters of concepts and highlights concepts. They can be considered as "needed" to pass from one cluster into another. In this regard, the "need", modelled by the betweennes, shows similarities to the domain ontology based "need" dimension.

The concept importance measure transforms the domain into a topology of learning, where high measures can be considered as modelled mountains of learning, as visualized in the Human Resource Management domain in Figure 41. Climbing the mountain early for learning enables a better understanding and reach through more available connections to follow "down the mountain", while reaching the peak through assessment requires an initial understanding of the represented topics to climb the mountain. The measure models the mountains and the implementing, path-based assessment is modelling the strategy for climbing the mountains and finally guiding and exploring the domain of learning.

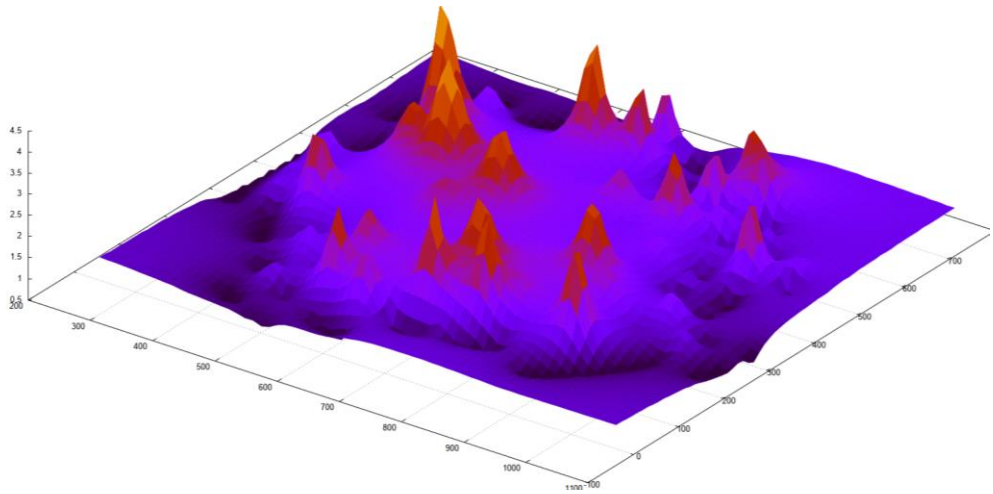


Figure 41: The measure converts the domain into a topology for learning.

8.4.6. Implications and Evaluation of the Concept Importance Measure Integration

To derive data about concepts and their implemented solutions within a field, two type of experiments can be distinguished in the frame of this work: controlled experiments and field experiments. Controlled experiments are planned in a controlled environment where the circumstances and most of the factors of the experiment can be controlled. Field experiment takes place in a real-world environment, where some but not all factors and circumstances can be controlled. The field studies in this work – analysing the interaction of the learner with STUDIO – can be considered as field experiments. The core environment of STUDIO is well-controlled, but there is no control when, where and how the learners interact with the system. All tracked data is passively observed and collected without intervention during the field studies. “Interventions”, in terms of changes which are expected to additionally influence the dependent variable, are applied only from one experiment to another, based on the derived feedback. This way the insights, which can be gathered throughout the experiments on the new concept importance measure and its implementation, are limited – considering the also limited control of factors and the passive data collection and analysis of the experiment.

Furthermore, the studies, conducted and evaluated for this work, consider that the performance of students on an online assessment test can reflect the learning progress –

especially by recording multiple test results over time within the learning and assessment cycle of STUDIO. Yet, the concept importance measure – based on the idea of connectivism – targets to derive an indication about the learning in an environment of interconnected information and not necessary target to derive a statement about the performance.

Finally, the assessment performance as a measure can – as the learning itself – be only fully set into context by considering multiple additional factors as: the composition of the domain, previous education, working experience, the environment of learning, motivation, and personal traits as cognitive styles or finally cultural differences. In this regard, tracking the influence of the concept importance measure on a dependent variable – as e.g. a performance measure – may be overlaid by these factors and be further influenced by the degree of openness of the field experiments. Additionally, the curriculum of the domain to learn may be inconsistently defined and aligned. E.g. a software engineering course may define a learning outcome on planning an effective software testing, while the course itself addresses the topic with a specific, yet interchangeable framework, which introduces its own learning barriers.

As such, the collected feedback and the conducted analyses – based on the concept importance implementation in STUDIO – can cover and uncover only parts of the overall picture. More and other learning related variables may have to be taken into account as dependent variables for future analyses. Throughout the course of this work five trial experiments were conducted, addressing different domains and considerations, as gathered in Table 30:

Table 30: Considered trial studies within STUDIO, conducted between 2015 and 2016.

Domain / Tailoring	Time	Algorithms / Application
Management Information Systems / mid-term	2015, March	Drill-down System exploration, utilized for Section 8.1
Management Information Systems / mid- and final-term	2015, April / May	Drill-down (final-term) and Bottom-up/path-based (mid-term and final-term) Exploring the event-tracking (Section 8.2) and exploring/evaluating the path-based algorithm (Section 8.4.1) (Section 8.4.3 (Weber and Vas, 2016b))(Section 8.3.2 (Weber and Vas, 2016a))
Human Resource Information Systems	2015, November	Drill-down and Bottom-up/path-based Exploring/Evaluating the path-based assessment

		(Section 8.4.1)
Management Information Systems / final-term, revised	2016, April / May	Drill-down and Concept Importance based Exploring/evaluating the concept importance based algorithm (Section 8.4.4) (Gkoumas et al., 2016) (Includes a cognitive style test)
Human Resource Information Systems	2016, October – 2017, January	Drill-down Continuous tracking (Includes a cognitive style test)

All trial studies could underline the operation of the examined components of STUDIO (existing and implemented), as: the drill-down assessment and evaluation, the event-tracking implementation, the path-based assessment and evaluation and finally the concept importance based assessment and evaluation. Furthermore, the learning of the individual students was analysed and reflected, based on their performance and behaviour while using the system. All considered assessment and evaluation algorithms enabled the students to learn and improve over the course of using the STUDIO system. In the process if doing so, the algorithms performed differently but lead to results on a comparable level regarding the assessment performance of the learner.

A comparative analysis of the learner behaviour and interaction using STUDIO, while utilizing both, the concept importance and the drill-down algorithm for assessment and evaluation is presented in (Gkoumas et al., 2016). The analysis proposes a self-organizing map (SOM) approach to detect hidden patterns in the learner’s interaction, while comparing the differences between the first and the best round of the assessment. The study could isolate in both test cases three groups of learners – a low performing, a good performing and an excellently performing group independently of the applied algorithm. The excellently performing group distinguishes itself by a higher number of interactions of the learners with the learning material. The members of the excellently performing and the good performing groups take more tests and check their performance on the single concepts of the domain more frequently.

The study also featured a Myers-Briggs Type Indicator (MBTI) cognitive styles test (Myers, 1998), which learners took before starting to use the STUDIO system. All students were split into two major groups, where each group took the STUDIO tests with a different assessment algorithm (drill-down, concept importance based). The student groups which used a different algorithm were balanced in terms of their MBTI types. The study couldn’t isolate a significant correlation between the selected algorithm selection and the performance and behaviour of specific MBTI types. Yet, the performance was different between MBTI types and may show a more significant difference for a different curriculum or with a differently controlled environment.

Figure 42 visualizes a comparison of the average final performance of the learner – the “accepted per all”[12] performance measure –against the average concept importance measure, averaged by all concepts which were asked from the same individual user. The graphs are based on the STUDIO based assessment of the Management Information Systems (MIS) curriculum, which were addressed in the same published study. A visualization of the tailored concept group is shown in Figure 44. The left graph in Figure 42 is based on the use of the concept importance based assessment, while the right graph is based on the drill-down algorithm. For the drill-down algorithm the concept importance measure is mapped against the assessed concepts “retrospectively”. The measure was not used within this assessment group.

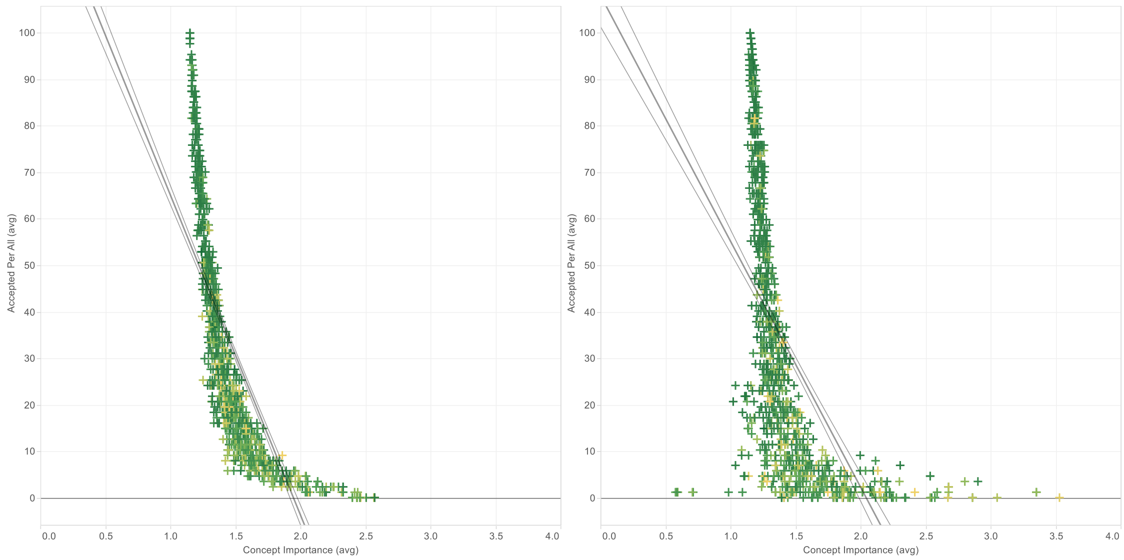


Figure 42: MIS concept group with the acceptance per all against the concept importance measure per use. The left shows the concept importance and the right the drill-down based testing.

The correlation between the passing of concepts with a high concept importance measure and the final performance is not significant. Nevertheless, for low performance tests the concept importance average is substantially higher. These difference is not tracing a specific assessment and learning behaviour but reflecting the structure of the

[12] The “accepted per all” performance measure is the overall final result of each assessment run of STUDIO and is calculated over all passed concepts which are also “accepted” by the active evaluation strategy, against the number of all concepts in the concept group. The description of the criteria is given in the sub-chapter, covering the different algorithms.

concept group graph. The size of the concepts in the concept group graph in Figure 44 depict the concept importance measure (the larger the circle is, the higher the importance of the concept is). Looking closer it becomes evident that based on the different connectivity of the concepts, the “middle” ring around the central concept collects the most higher measures, as the concepts in that area have a higher connectivity degree.

As such these concepts are asked first for the drill-down algorithm – which starts from the start-concept – while for the concept importance based algorithm concepts with a high measure are always selected first. The latter is also the reason for the more compact trend in the left concept importance visualization in Figure 42, which also approves visually the expected behaviour of the algorithm.

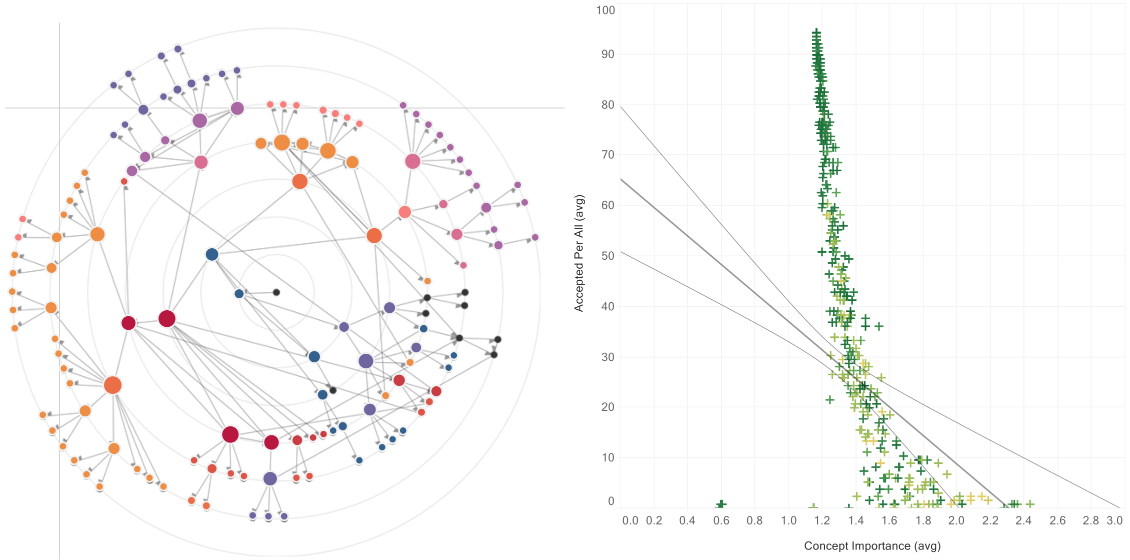


Figure 43: HRIS concept group with the concept importance measure on the graph and per user against the acceptance rate.

When the leaner explore more of the domain, more and more lower measured concepts are asked and lower the average of the concept importance measure in the comparison graph on Figure 42 by exploring more and more less connected and therefore lower measured concepts. The same trend can be observed for Human Resource Information Systems (HRIS) concept group in Figure 43. Here also the drill-down algorithm is shown which is mapped against the concept importance measure. The three performance/importance comparing graphs in Figure 42 and Figure 43, show a characteristic shape which can be reproduced in different domains and even across different algorithms, which indicates that the measure can differentiate concepts well. But, the average concept importance measure for differently performed tests is strongly influenced by lower values of the concept importance measure, which are more common in broader areas of the concept groups. As such the performance seems to be an insufficient proxy to track a correlation between the concept importance measure and learning.

An alternate approach is visualized in Figure 44 and Figure 45. Here the MIS concept group is matched against the results of a traditional multiple choice (MC) exam and then combined with the concept importance measure. To do so, the set of questions were mapped against fitting concepts of the concept group and combined to one merged data set. As for Figure 43, the size of concepts depicts the concept importance. The colours in contrast tracks the modularity class of the concepts. The modularity is a network measure and the related modularity algorithm uses the modularity to cluster a network into smaller networks with an optimal modularity value (Newman, 2006). Modularity itself is a measure which tracks how similar a network is to a random network of the same size.

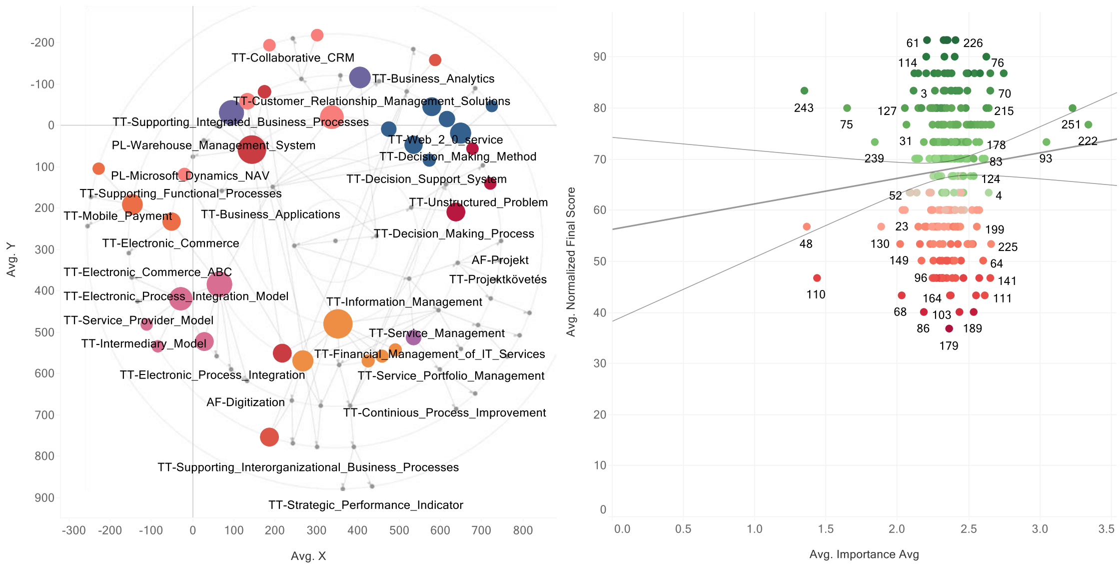


Figure 44: MIS concept group with the concept importance measure on the graph and the acceptance rate against the results of the traditional MC (final) exam.

The modularity algorithm splits the concept group into sub-groups for a more differentiated statement about the tailored domain. Each group includes a sub-set of concepts. Figure 45 visualizes on the left part the MC mid-term exam and on the right the final exam and as such also pictures a time horizon from left to right. Both graphs are clustered by the modularity classes addressed by the assessed concepts. The green area shows the average performance on questions of these concepts in a cluster, the orange part averages the concept importance, while the blue parts gather the frequency of assessed questions and their grouping concepts.

What makes the graphs remarkable is the change of the performance within the modularity classes, in connection with the calculated concept importance within the modularity classes. In the left block, the classes with the highest importance are 5 and 8, while both also mark the lowest performance. For the final-term in contrast, the same groups with an even stronger importance value now mark the highest performance classes. In both terms the exams ask different questions for the same concepts. Considering the connectivistic assumption, that a better connectivity of concepts to learn, also enable a better learning, the visualized results can be interpreted in a way that the high importance in the left graph leads to a behaviour where it takes considerably longer to master the concepts, while, once mastered, they are a high performing "well-memorized" starting point for more learning.

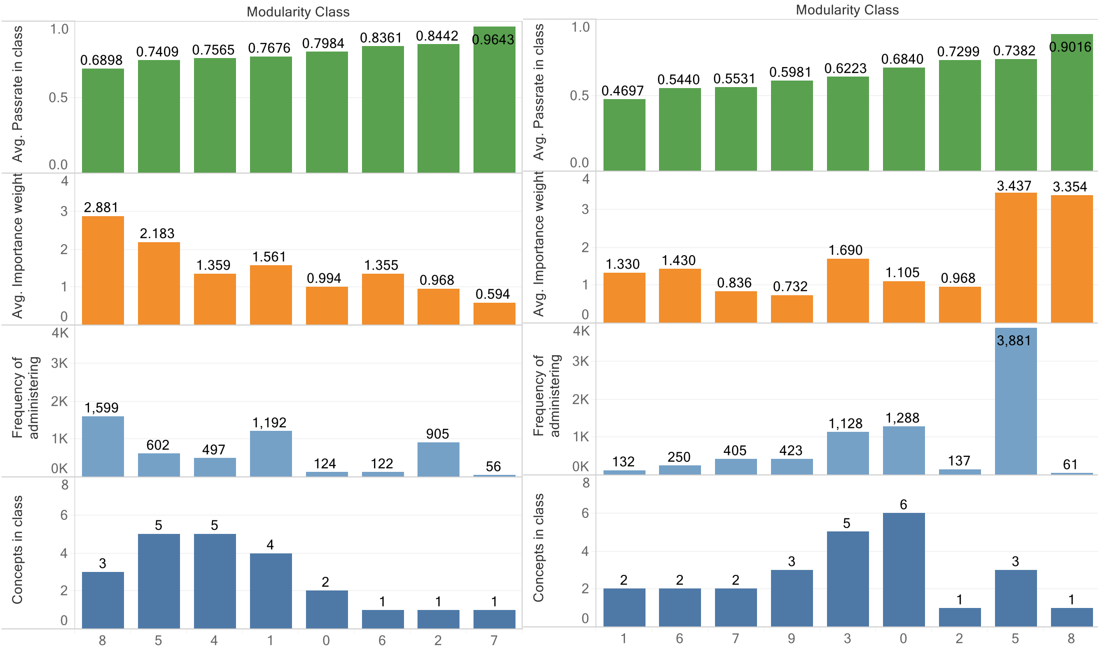


Figure 45: MIS traditional MC exam, clustered into modularity classes based on the graph, showing the acceptance rate and the concept importance measure per class.

A further support is shown in Figure 46, visualizing the flow of the performance over the test rounds a learner takes while using STUDIO. The green curve shows the concept importance, while the blue curve tracks the drill-down performance. Additionally, the curves are filtered on users with at least 5 test rounds. Furthermore, the top 10% performer are excluded as there is a likelihood that they are performing that well that their behaviour is independent of the provided environment. What can be seen is that for the concept importance based algorithm in green, learners start with a short low performing valley in contrast to the drill-down version. This would back that highly important concepts (which are assessed early) are harder to learn but are then a good starting point to master more concepts– which would also reflect the jump in the performance towards the 5th test round.

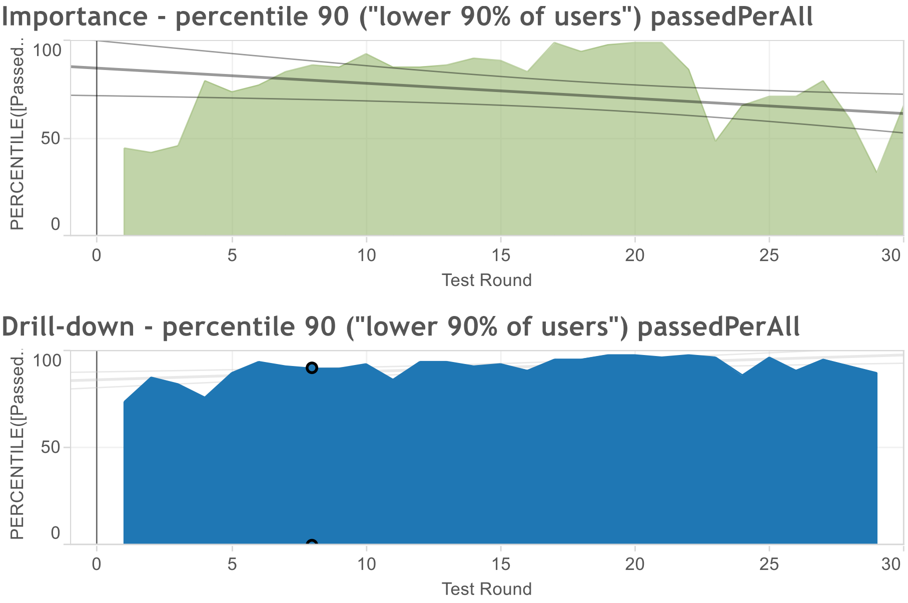


Figure 46: Development of the passing of concepts over the drill-down and concept importance based algorithm in MIS.

Even though there is no trivial correlation between the concept importance measure for learning and specific performance measures, defined in STUDIO, the concept importance can still reflect the importance of concepts for learning. Beside the indications collected here, other types of experiments and other experiment configurations and environments may improve the insights into the impact of the developed concept importance measure. Independent of the presented results and the gathered assumptions, many more factors may influence the learning process and as such are worth to be explored further. The created and presented measure can differentiate concepts within the domain, based on the domain structure and its semantic. Furthermore, the proposed algorithm successfully uses the concept importance to explore the domain, based on a path-based approach to assessment and evaluation. Future studies and extensions of the measure will uncover more of its potential to derive a statement about the influence of the connectivity and the semantic of a domain on the learning of a learner – in line with the learning theory of connectivism.

9. Conclusion, Impact, and Future Research

This thesis researches a new approach to implement a smart technology enhanced knowledge assessment, exploiting the structure and semantics of a learning domain model. The concept is based on the vision, that knowledge is connected and can be captured as a semantically enhanced knowledge network in the form of a domain ontology and that **the connectivity and semantics of a knowledge network can be utilized to derive the importance of a concept** and to recommend concepts to learn, based on the importance rating of concepts. This concept importance measure can support a continuous education of flexible workers in a smart economy by using the resulting rating to select which concepts to assess and learn first to explore a domain for learning.

This work draws on considerations of different fields in the context of education to offer a solid background for the new approach of concept importance: Taking a viewpoint of **organizational learning**, learning has to be flexible in terms of situation and time, offering an adaptive learning support which uses an adaptive assessment to narrow the individual training needs. An **adaptive assessment test** can enable to tailor the training need but furthermore it also helps to overcome the potentially flawed self-assessment of the learner and implement short assessment and learning cycles to support a better learning. To adapt the assessment to the learner, an idea of the user in terms of a **user model** is needed to tailor the training needs and to envision the learning progress of the learner on the domain ontology – which is used to drive the assessment. Finally, to explore the network of knowledge in line with the **learning theory** of connectivism, **network measures** are a fitting narrative and basement to translate the connectivity and the semantics of the specific "domain to learn" to a measure of **concept importance** of single concepts – for the assessment of the individual learning progress of the learner.

To design, model and implement a new concept importance measure this work builds on the STUDIO system for technology enhanced assessment and learning, based on the use of a domain ontology. In the frame of this research, different extensions to the STUDIO system are developed to enable and integrate the concept importance measure for a smart assessment and learning. A **data collection framework** extends the system to track the flow of the assessment and those features and variables that enable the application of the concept importance and guide the assessment. A **path-based assessment and evaluation algorithm** is developed and then integrated as a basement for an open knowledge exploration within the domain ontology, which is exploited to conduct the assessment. Finally**, the modelled concept importance is implemented within the path-based exploration** extension. To support and evaluate findings, each stage of **the research is backed by experiments** using the STUDIO system in blended learning environments, providing feedback on how to customize the application of the novel concept importance measure in semantically enhanced domain networks used for assessment and learning.

The research results can be considered as successful, since:

- The applied exploratory research methodology mix of modelling and evaluation, and the resulting stages are suitable to explore and answer the research question
- The concept importance measure is defined sound and functional, building on the laid foundation in the literature analysis.
- The importance dimensions are well-explored and cover the structure and semantic of the domain ontology, and provide a functional and flexible fundament for additional extensions of the modelled dimensions.
- Experiments show that the implemented path-based assessment and evaluation to explore the domain, is operating on pair with the existing selected online assessment and learning solution. The domain is explored well and the implementation hosts the concept importance measure for a working, concept importance based adaptive online assessment.
- The concept importance measure and the final integration, address the features of the connectivistic learning theory. In a field-study the connectivity were underlined as an indicator for assessment and learning.

The major contribution and result of this work is the definition of a new domain ontology aware measure to rate the importance of concepts for learning, and addresses the connectivistic learning idea – while its implementation is well-suited to support a flexible learning for rapidly changing requirements of the labour market.

The new concept importance measure targets to help to solve the question of “What to learn first?” – to master a given domain and to explore the knowledge in a field of learning. It introduces an extendable concept to translate important dimensions of the domain ontology and to integrate new dimensions flexible to consider more information about the domain but also about the user. Future extensions can connect here and enable an online and on-demand modification of the concept importance to improve the adaptivity of the system in the context of learning to the learner, the learner's performance, and external factors. In this regard, the approach of deriving the concept importance can be connected to new sources of information for a better and further tailored learning experience.

10. Published Research Results

The following publications are a selection, selected based on their relevance for the conducted research:

2016

Weber, Christian, and Neusch, Gábor, and Vas, Réka. 2016. 'STUDIO: A Domain Ontology Based Solution for Knowledge Discovery in Learning and Assessment'. International Conference on Information Systems Education and Research – of the AIS Special Interest Group for Education (AIS-SIGEd) 2016, Dublin, Ireland.

Weber, Christian, and Vas, Réka. 2016. 'Applying connectivism? Does the connectivity of concepts make a difference for learning and assessment?' In Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics. Budapest, Hungary: IEEE.

Weber, Christian, and Vas, Réka. 2016. 'How to Learn More from Knowledge Networks Through Social Network Analysis Measures'. In Proceedings of the 1st Pro-Nursing Symposium 2016. Bonn, Germany.

Weber, Christian, and Vas, Réka. 2016. 'TOP-DOWN OR BOTTOM UP: A COMPARATIVE STUDY ON ASSESSMENT STRATEGIES IN THE STUDIO ADAPTIVE LEARNING ENVIRONMENT'. In Proceedings of the European Distance and E-learning Network 2016 Annual Conference. Budapest, Hungary: EDEN.

Weber, Christian. 2016. 'Context-Aware Self-Assessment Path Generation for Personalised Education'. Journal of the Scientific and Educational on Forum on Business Information Systems 10 (10).

Weber, Christian. 2016. 'STUDIO: A Solution on Adaptive Testing'. In Corporate Knowledge Discovery and Organizational Learning, edited by András Gabor and Andrea Kö. Knowledge Management and Organizational Learning. Springer International Publishing.

2015

Weber, Christian, and Truong Huong, May, and Vas, Réka. 2015. 'Context-Aware Self-Assessment in Higher Education'. In *EDULEARN15 Proceedings*, 5910–20. Barcelona, Spain: IATED.

Weber, Christian, and Vas, Réka. 2015. 'Studio: Ontology-Based Educational Self-Assessment'. In *Workshops Proceedings of EDM 2015 8th International Conference on Educational Data Mining, EDM 2015, Madrid, Spain, June 26-29, 2015.*, 1446:33–40. CEUR Workshop Proceedings (CEUR-WS.org). Madrid, Spain. http://ceur-ws.org/Vol-1446/GEDM_2015_Submission_5.pdf.

2014

Weber, Christian. 2014. 'Enabling a Context Aware Knowledge-Intense Computerized Adaptive Test through Complex Event Processing'. *Journal of the Scientific and Educational on Forum on Business Information Systems* 9 (9): 66–74.

Weber, Christian, and Vas, Réka. 2014. 'Extending Computerized Adaptive Testing to Multiple Objectives: Envisioned on a Case from the Health Care'. In *Electronic Government and the Information Systems Perspective*, edited by Andrea Kö and Enrico Francesconi, 8650:148–62. Lecture Notes in Computer Science. Springer International Publishing. http://dx.doi.org/10.1007/978-3-319-10178-1_12.

11. References

Adomavicius, G., Tuzhilin, A., 2005. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE transactions on knowledge and data engineering 17, 734–749. doi:10.1109/TKDE.2005.99

Ahonen, H., Ahonen, H., Heikkinen, B., Heikkinen, B., Kilpeläinen, P., Heinonen, O., Heinonen, O., Jaakkola, J., Jaakkola, J., Mannila, H., Mannila, H., 1996. Intelligent Assembly of Structured Documents, Series of Publications C. Department of Computer Science, University of Helsinki, Helsinki.

Ajwani, D., Cosgaya-Lozano, A., Zeh, N., 2012. A Topological Sorting Algorithm for Large Graphs. J. Exp. Algorithmics 17, 3.2:3.1–3.2:3.21. doi:10.1145/2133803.2330083

Akbulut, Y., Cardak, C.S., 2012. Adaptive educational hypermedia accommodating learning styles: A content analysis of publications from 2000 to 2011. Computers & Education 58, 835–842. doi:10.1016/j.compedu.2011.10.008

Amaral, J.N., Elio, R., Hoover, J., Nikolaidis, I., Salavatipour, M., Stewart, L., Wong, K., 2011. About Computing Science Research Methodology.

Anderson, J.R., 1988. The expert module, in: Polson, M., Richardson, J. (Eds.), Handbook of Intelligent Training Systems. Erlbaum, Hillsdale, NJ, pp. 21–53.

Arru, M., 2014. Application of Process Ontology to improve the funding allocation process at the European Institute of Innovation and Technology., in: 3rd International Conference on Electronic Government and the Information Systems Perspective (EGOVIS). Munich, Germany.

Ausburn, L.J., Ausburn, F.B., 1978. Cognitive styles: Some information and implications for instructional design. ECTJ 26, 337–354. doi:10.1007/BF02766370

Barabási, A.-L., Pósfai, M., 2016. Network science. Cambridge University Press.

Beaumont, I.H., 1994. User modelling in the interactive anatomy tutoring system ANATOM-TUTOR. User Model User-Adap Inter 4, 21–45. doi:10.1007/BF01142356

Bednar, A.K., Cunningham, D., Duffy, T.M., 1991. Theory into practice: How do we link, in: Duffy, T.M. (Ed.), Constructivism and the Technology of Instruction: A Conversation. Routledge, pp. 17–34.

Bellman, R.E., Zadeh, L.A., 1970. Decision-Making in a Fuzzy Environment. Management Science 17, B-141. doi:10.1287/mnsc.17.4.B141

Benyon, D., Murray, D., 1988. Experience with adaptive interfaces. The Computer Journal 31, 465–473. doi:10.1093/comjnl/31.5.465

Bloom, B.S., 1984. The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. Educational Leadership 41, 4–17.

Blundell, R., Dearden, L., Meghir, C., Sianesi, B., 1999. Human Capital Investment: The Returns from Education and Training to the Individual, the Firm and the Economy. Fiscal Studies 20, 1–23. doi:10.1111/j.1475-5890.1999.tb00001.x

Bontcheva, K., Wilks, Y., 2005. Tailoring Automatically Generated Hypertext. User Model User-Adap Inter 15, 135–168. doi:10.1007/s11257-004-5637-6

Bouzeghoub, A., Carpentier, C., Defude, B., Duitama, F., 2003. A model of reusable educational components for the generation of adaptive courses, in: Proc. First International Workshop on Semantic Web for Web-Based Learning in

Conjunction with CAISE.
Brusilovsky, P., 2003. Developing Adaptive Educational Hypermedia Systems: From Design Models to Authoring Tools, in: Murray, T., Blessing, S.B., Ainsworth, S. (Eds.), Authoring Tools for Advanced Technology Learning Environments. Springer Netherlands, pp. 377–409. doi:10.1007/978-94-017-0819-7_13
Brusilovsky, P., 1996. Methods and techniques of adaptive hypermedia. User Model User-Adap Inter 6, 87–129. doi:10.1007/BF00143964
Brusilovsky, P., 1994. Student model centered architecture for intelligent learning environments, in: Proc. of Fourth International Conference on User Modeling,(Hyannis, MA, 15-19 August 1994), MITRE. pp. 31–36.
Brusilovsky, P., Millán, E., 2007. User Models for Adaptive Hypermedia and Adaptive Educational Systems, in: Brusilovsky, P., Kobsa, A., Nejdl, W. (Eds.), The Adaptive Web, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 3–53. doi:10.1007/978-3-540-72079-9_1
Brusilovsky, P., Peylo, C., 2003. Adaptive and Intelligent Web-based Educational Systems. International Journal of Artificial Intelligence in Education (IJAIED) 13, 159–172.
Brusilovsky, P., Sosnovsky, S., Yudelson, M., 2005. Ontology-based framework for user model interoperability in distributed learning environments, in: World Conference on E-Learning, E-Learn. pp. 2851–2855.
Brusilovsky, P., Vassileva, J., 2003. Course sequencing techniques for large-scale web-based education. International Journal of Continuing Engineering Education and Life Long Learning 13, 75–94. doi:10.1504/IJCEELL.2003.002154
Brusilovsky, P.L., 1992. A framework for intelligent knowledge sequencing and task sequencing, in: Frasson, C., Gauthier, G., McCalla, G.I. (Eds.), Intelligent Tutoring Systems, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 499–506. doi:10.1007/3-540-55606-0_59
Bunderson, C.V., Inouye, D.K., Olsen, J.B., 1988. The Four Generations of Computerized Educational Measurement. ETS Research Report Series 1988, i-148. doi:10.1002/j.2330-8516.1988.tb00291.x
Burgos, D., Tattersall, C., Koper, R., 2007. How to represent adaptation in e-learning with IMS learning design. Interactive Learning Environments 15, 161–170. doi:10.1080/10494820701343736
Carlson, J.E., von Davier, M., 2013. Item Response Theory. Universiteit Twente, Enschede.
Castello, V., Mahajan, L., Flores, E., Gabor, M., Neusch, G., Szabó, I.B., Caballero, J.G., Vettraino, L., Luna, J.M., Blackburn, C., Ramos, F.J., 2014. THE SKILL MATCH CHALLENGE. EVIDENCES FROM THE SMART PROJECT. ICERI2014 Proceedings 1182–1189.
CEDEFOP, 2011. Glossary - Quality in education and training, Cedefop information series. CEDEFOP, Luxembourg.
Čekerevac Zoran, A.S., Petar, Č., 2013. Knowledge assessment and application of computer adaptive testing. MEST Journal 1, 16–30. doi:10.12709/mest.01.01.02.02
Chandy, M.K., Etzion, O., Ammon, R. von, 2011. Executive Summary and Manifesto – Event Processing, in: Chandy, K.M., Etzion, O., Ammon, R. von (Eds.), Event Processing, Dagstuhl Seminar Proceedings. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany, Dagstuhl, Germany.
Chen, W., Mizoguchi, R., 2004. Learner model ontology and learner model agent, in: Kommers, P. (Ed.), Cognitive Support for Learning - Imagining the Unknown.

IOS Press, pp. 189–200.
Chrysafiadi, K., Virvou, M., 2013. Student modeling approaches: A literature review for the last decade. Expert Systems with Applications 40, 4715–4729. doi:10.1016/j.eswa.2013.02.007
Cicconi, S., 1999. Hypertextuality. Mediapolis. Berlino & New York: De Gruyter 21–43.
Clancey, W.J., 1987. Knowledge-based Tutoring: The GUIDON Program. MIT Press, Cambridge, MA, USA.
Coffield, F., Moseley, D., Hall, E., Ecclestone, K., 2004a. Learning styles and pedagogy in post-16 learning: A systematic and critical review. London: Learning and Skills Research Centre.
Coffield, F., Moseley, D., Hall, E., Ecclestone, K., 2004b. Should We Be Using Learning Styles? What Research Has to Say to Practice. Learning and Skills Research Centre, London.
Commission, E., 2001. Making a European Area of Lifelong Learning a Reality: Communication from the Commission. Office for Official Publications of the European Communities.
Conlan, O., O'Keeffe, I., Tallon, S., 2006. Combining Adaptive Hypermedia Techniques and Ontology Reasoning to Produce Dynamic Personalized News Services, in: Wade, V.P., Ashman, H., Smyth, B. (Eds.), Adaptive Hypermedia and Adaptive Web-Based Systems, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 81–90. doi:10.1007/11768012_10
Corbett, A.T., Koedinger, K.R., Anderson, J.R., 1997. Intelligent Tutoring Systems, in: Landauer, T.K., Prabhu, P.V. (Eds.), Handbook of Human-Computer Interaction (Second Edition). North-Holland, Amsterdam, pp. 849–874.
Dagger, D., Wade, V., Conlan, O., 2002. Towards a standards-based approach to e-Learning personalization using reusable learning objects, in: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education. pp. 210–217.
Dahab, M.Y., Hassan, H.A., Rafea, A., 2008. TextOntoEx: Automatic ontology construction from natural English text. Expert Systems with Applications 34, 1474–1480. doi:10.1016/j.eswa.2007.01.043
de Saint-Exupéry, A., 2013. The little prince. Houghton Mifflin Harcourt.
Del Corso, D., Ovcin, E., Morrone, G., Gianesini, D., Salojarvi, S., Kvist, T., 2001. 3DE: an environment for the development of learner-centered custom educational packages, in: Frontiers in Education Conference, 2001. 31st Annual. Presented at the Frontiers in Education Conference, 2001. 31st Annual, p. F2C–21–6 vol.2. doi:10.1109/FIE.2001.963703
Downes, S., 2008. An Introduction to Connective Knowledge, in: Hug, T. (Ed.), Media, Knowledge & Education - Exploring New Spaces, Relations and Dynamics in Digital Media Ecologies. Innsbruck University Press.
Drachsler, H., Hummel, H.G.K., Koper, R., 2009. Identifying the Goal, User model and Conditions of Recommender Systems for Formal and Informal Learning. Journal of Digital Information 10.
Driscoll, M.P., 2005. Psychology of Learning for Instruction. Pearson Allyn and Bacon.
Dunning, D., Heath, C., Suls, J.M., 2004. Flawed Self-Assessment: Implications for Health, Education, and the Workplace. Psychological Science in the Public Interest 5, 69–106. doi:10.1111/j.1529-1006.2004.00018.x
Easterby-Smith, M., Crossan, M., Nicolini, D., 2000. Organizational Learning: Debates Past, Present And Future. Journal of Management Studies 37, 783–796.

doi:10.1111/1467-6486.00203
Encarnação, L.M., 1997. Multi-level User Support Through Adaptive Hypermedia: A Highly Application-independent Help Component, in: Proceedings of the 2Nd International Conference on Intelligent User Interfaces, IUI ’97. ACM, New York, NY, USA, pp. 187–194. doi:10.1145/238218.238322
Erkan, G., Radev, D.R., 2004. LexRank: Graph-based lexical centrality as salience in text summarization. Journal of Artificial Intelligence Research 22, 457–479. doi:doi:10.1613/jair.1523
Ertmer, P.A., Newby, T.J., 2013. Behaviorism, Cognitivism, Constructivism: Comparing Critical Features From an Instructional Design Perspective. Perf. Improvement Qrtly 26, 43–71. doi:10.1002/piq.21143
Etzion, O., Niblett, P., 2010. Event processing in action. Manning Publications Co.
Falmagne, J.-C., Koppen, M., Villano, M., Doignon, J.-P., Johannesen, L., 1990. Introduction to knowledge spaces: How to build, test, and search them. Psychological Review 97(2), 201–224. doi:10.1037/0033-295X.97.2.201
Fan, X., 1998. Item Response Theory and Classical Test Theory: An Empirical Comparison of their Item/Person Statistics. Educational and Psychological Measurement 58, 357–381. doi:10.1177/0013164498058003001
Fiol, C.M., Lyles, M.A., 1985. Organizational Learning. The Academy of Management Review 10, 803–813. doi:10.2307/258048
Fischer, G., 2001. User Modeling in Human–Computer Interaction. User Modeling and User-Adapted Interaction 11, 65–86. doi:10.1023/A:1011145532042
Fletcher, J.D., 2003. Evidence for learning from technology-assisted instruction, in: O’Neil, H.F., Perez, R.S. (Eds.), Technology Applications in Education: A Learning View. Lawrence Erlbaum Associates, Mahwah, pp. 323–339.
Francisco-Revilla, L., Shipman III, F.M., 2000. Adaptive medical information delivery combining user, task and situation models, in: Proceedings of the 5th International Conference on Intelligent User Interfaces. ACM, pp. 94–97.
Frasincar, F., IJntema, W., Goossen, F., Hogenboom, F., 2011. A semantic approach for news recommendation, in: Business Intelligence Applications and the Web: Models, Systems and Technologies. IGI Global.
Gabor, A., Ko, A. (Eds.), 2016. Corporate Knowledge Discovery and Organizational Learning, Knowledge Management and Organizational Learning. Springer International Publishing.
Garcia, P., Amandi, A., Schiaffino, S., Campo, M., 2007. Evaluating Bayesian Networks’ Precision for Detecting Students’ Learning Styles. Computers & Education 49, 794–808. doi:10.1016/j.compedu.2005.11.017
Garratt, B., 1987. The Learning Organization: and the need for directors who think. Fontana Paperbacks.
Gauch, S., Speretta, M., Chandramouli, A., Micarelli, A., 2007. User profiles for personalized information access, in: The Adaptive Web. Springer, pp. 54–89.
Gavrilova, T.A., Leshcheva, I.A., 2015. Ontology design and individual cognitive peculiarities: A pilot study. Expert Systems with Applications 42, 3883–3892. doi:10.1016/j.eswa.2015.01.008
Gershon, R.C., 1992. Test Anxiety and Item Order: New Concerns for Item Response Theory, in: Wilson, M. (Ed.), Objective Measurement: Theory into Practice. Ablex, Norwood NJ.
Gkoumas, D., Gausz, B., Vas, R., 2016. An analysis of learning behaviour and patterns in a technology-enhanced learning environment, in: Proceedings of the Special Interest Group for Education (SIGED) 2016. Presented at the Special Interest

Group for Education (SIGED), Dublin, Ireland.
Graf, S., Kinshuk, 2010. Using Cognitive Traits for Improving the Detection of Learning Styles, in: 2010 Workshop on Database and Expert Systems Applications (DEXA). Presented at the 2010 Workshop on Database and Expert Systems Applications (DEXA), pp. 74–78. doi:10.1109/DEXA.2010.35
Grauwe, P.D., 2010. Top-Down versus Bottom-Up Macroeconomics. CESifo Economic Studies 56, 465–497. doi:10.1093/cesifo/ifq014
Gündüz, Ş., Özsu, M.T., 2003. A Web Page Prediction Model Based on Click-stream Tree Representation of User Behavior, in: Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '03. ACM, New York, NY, USA, pp. 535–540. doi:10.1145/956750.956815
Haase, P., van Harmelen, F., Huang, Z., Stuckenschmidt, H., Sure, Y., 2005. A framework for handling inconsistency in changing ontologies, in: Proceedings of the 4th International Conference on The Semantic Web, ISWC'05. Springer-Verlag, Galway, Ireland, pp. 353–367. doi:10.1007/11574620_27
Hebeler, J., Fisher, M., Blace, R., Perez-Lopez, A., Dean, M., 2009. Semantic Web Programming. Wiley.
Heckmann, D., Schwartz, T., Brandherm, B., Schmitz, M., Wilamowitz-Moellendorff, M. von, 2005. Gumo – The General User Model Ontology, in: Ardissono, L., Brna, P., Mitrovic, A. (Eds.), User Modeling 2005, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 428–432. doi:10.1007/11527886_58
Henning, G., 1987. A guide to language testing: development, evaluation, research, Methodology Series. Newberry House Publishers.
Hnida, M., Idrissi, M.K., Bennani, S., 2014. Towards an adaptive e-learning system based on individualized paths in a competency-based approach. RECENT ADVANCES in EDUCATIONAL TECHNOLOGIES and EDUCATION 73.
Hoffart, J., Altun, Y., Weikum, G., 2014. Discovering emerging entities with ambiguous names, in: Proceedings of the 23rd International Conference on World Wide Web. ACM, pp. 385–396.
Holland, P.W., Wainer, H., 1993. Differential item functioning. Routledge.
Huxley, J., others, 1940. The new systematics. The new systematics.
Interview with Thomas Duffy [WWW Document], 2000. URL http://www2.gsu.edu/~wwwitr/interviews/duffy.htm (accessed 12.29.15).
Jameson, A., 2003. The Human-computer Interaction Handbook, in: Jacko, J.A., Sears, A. (Eds.), . L. Erlbaum Associates Inc., Hillsdale, NJ, USA, pp. 305–330.
Jeong, H., Mason, S.P., Barabási, A.-L., Oltvai, Z.N., 2001. Lethality and centrality in protein networks. Nature 411, 41–42. doi:10.1038/35075138
Joksimović, S., Manataki, A., Gašević, D., Dawson, S., Kovanović, V., de Kereki, I.F., 2016. Translating Network Position into Performance: Importance of Centrality in Different Network Configurations, in: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, LAK '16. ACM, New York, NY, USA, pp. 314–323. doi:10.1145/2883851.2883928
Jonassen, D., 2009. Reconciling a Human Cognitive Architecture, in: Tobias, S., Duffy, T.M. (Eds.), Constructivist Instruction: Success Or Failure? Routledge, pp. 13–33.
Jonassen, D.H., 1991. Objectivism versus constructivism: Do we need a new philosophical paradigm? Educational Technology Research and Development 39, 5–14. doi:10.1007/BF02296434
Jowett, B., 1901. The dialogues of Plato, translated into English with analyses and

introductions, Vol. 3. Charles Scribner's Sons, New York, NY.
Kaplan, C., Fenwick, J., Chen, J., 1993. Adaptive hypertext navigation based on user goals and context. User Model User-Adap Inter 3, 193–220. doi:10.1007/BF01257889
Kay, J., 2000. Stereotypes, Student Models and Scrutability, in: Gauthier, G., Frasson, C., VanLehn, K. (Eds.), Intelligent Tutoring Systems, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 19–30. doi:10.1007/3-540-45108-0_5
Khobreh, M., Fathi, M., Ansari, F., Dornhöfer, M., 2013. An ontology-based Recommender System to Support Nursing Education and Training, in: LWA Workshop Proceedings. Presented at the LWA 2013. Lernen, Wissen & Adaptivität, Lehrstuhl für Medieninformatik, Universität Bamberg, Bamberg, Germany.
Koper, R., Tattersall, C., 2004. New directions for lifelong learning using network technologies. British Journal of Educational Technology 35, 689–700. doi:10.1111/j.1467-8535.2004.00427.x
Krathwohl, D.R., 2002. A Revision of Bloom's Taxonomy: An Overview. Theory Into Practice 41, 212–218. doi:10.1207/s15430421tip4104_2
Kuckulenz, A., 2007. Continuing Training, in: Studies on Continuing Vocational Training in Germany, ZEW Economic Studies. Physica-Verlag HD, pp. 9–52.
Lawson, S., 1991. One-Parameter Latent Trait Measurement: Do the Results justify the Effort? Advances in educational research : substantive findings, methodological developments.
Leondes, C.T., 2010. Intelligent Knowledge-Based Systems: Business and Technology in the New Millennium. Springer Science & Business Media.
Lester, J.C., Stone, B.A., Stelling, G.D., 1999. Lifelike pedagogical agents for mixed-initiative problem solving in constructivist learning environments, in: Computational Models of Mixed-Initiative Interaction. Springer, pp. 185–228.
Linacre, J.M., 2000. Computer-adaptive testing: A methodology whose time has come, in: Chae, J.M., S. Kang, U.–Jeon, E.–.Linacre (Ed.), Development of Computerized Middle School Achievement Tests, MESA Memorandum. Komesa Press, Seoul.
Linden, G., Smith, B., York, J., 2003. Amazon. com recommendations: Item-to-item collaborative filtering. IEEE Internet computing 7, 76–80.
Liu, Y., Ginther, D., 1999. Cognitive Styles and Distance Education. Online Journal of Distance Learning Administration 2.
Lonn, S., Teasley, S.D., 2009. Saving time or innovating practice: Investigating perceptions and uses of Learning Management Systems. Computers & Education 53, 686–694. doi:10.1016/j.compedu.2009.04.008
Lord, F.M., 2012. Applications of Item Response Theory To Practical Testing Problems. Taylor & Francis.
Lu, Z., Zha, H., Yang, X., Lin, W., Zheng, Z., 2013. A New Algorithm for Inferring User Search Goals with Feedback Sessions. IEEE Transactions on Knowledge and Data Engineering 25, 502–513. doi:10.1109/TKDE.2011.248
Luckham, D., 2002. The Power of Events. Addison-Wesley.
Makransky, G., Glas, C.A.W., 2014. An Automatic Online Calibration Design in Adaptive Testing. Journal of Applied Testing Technology 11, 1–20.
Makransky, G., Glas, C.A.W., 2013. Modeling differential item functioning with group-specific item parameters: A computerized adaptive testing application. Measurement 46, 3228–3237. doi:10.1016/j.measurement.2013.06.020

Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H., Koper, R., 2011. Recommender Systems in Technology Enhanced Learning, in: Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (Eds.), Recommender Systems Handbook. Springer US, pp. 387–415.

Maria Zajac, 2009. Using learning styles to personalize online learning. Campus-Wide Info Systems 26, 256–265. doi:10.1108/10650740910967410

Martins, A.C., Faria, L., Carvalho, C.V. de, Carrapatoso, E., 2008. User Modeling in Adaptive Hypermedia Educational Systems. Journal of Educational Technology & Society 11, 194–207.

Micarelli, A., Gasparetti, F., Sciarrone, F., Gauch, S., 2007. Personalized Search on the World Wide Web, in: Brusilovsky, P., Kobsa, A., Nejdl, W. (Eds.), The Adaptive Web, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 195–230. doi:10.1007/978-3-540-72079-9_6

Missikoff, M., Navigli, R., Velardi, P., 2002. The Usable Ontology: An Environment for Building and Assessing a Domain Ontology, in: Horrocks, I., Hendler, J. (Eds.), The Semantic Web — ISWC 2002, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 39–53.

Morita, M., Shinoda, Y., 1994. Information Filtering Based on User Behavior Analysis and Best Match Text Retrieval, in: Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR ’94. Springer-Verlag New York, Inc., New York, NY, USA, pp. 272–281.

Murray, T., 1999. Authoring Intelligent Tutoring Systems: An analysis of the state of the art. International Journal of Artificial Intelligence in Education (IJAIED) 10, 98–129.

Myers, I.B., 1998. MBTI manual: a guide to the development and use of the Myers-Briggs Type Indicator. Consulting Psychologists Press, Palo Alto, Calif.

Nas, T.F., 2016. Cost-Benefit Analysis: Theory and Application. Lexington Books.

Neuhaus, F., Florescu, E., Galton, A., Gruninger, M., Guarino, N., Obrst, L., Sanchez, A., Vizedom, A., Yim, P., Smith, B., 2011. Creating the ontologists of the future. Applied Ontology 6, 91–98. doi:10.3233/AO-2011-0083

Neusch, G., Gábor, A., 2014. PROKEX – INTEGRATED PLATFORM FOR PROCESS-BASED KNOWLEDGE EXTRACTION. ICERI2014 Proceedings 3972–3977.

Newman, M., 2010. Networks: An Introduction. OUP Oxford.

Newman, M.E.J., 2006. Modularity and community structure in networks. PNAS 103, 8577–8582. doi:10.1073/pnas.0601602103

Nodenot, T., Marquesuzaá, C., Laforcade, P., Sallaberry, C., 2004. Model Based Engineering of Learning Situations for Adaptive Web Based Educational Systems, in: Proceedings of the 13th International World Wide Web Conference on Alternate Track Papers &Amp; Posters, WWW Alt. ’04. ACM, New York, NY, USA, pp. 94–103. doi:10.1145/1013367.1013385

Obermaisser, R. (Ed.), 2011. Time-Triggered Communication. CRC Press, Taylour & Francis Group.

OECD, 2013. Recognition of Non-formal and Informal Learning [WWW Document]. URL http://www.oecd.org/education/skills-beyond-school/recognitionofnon-formalandinformallearning-home.htm (accessed 6.14.15).

OECD, 2005. E-learning in Tertiary Education: Where do we stand? 292. doi:http://dx.doi.org/10.1787/9789264009219-en

Ohlsson, S., 1986. Some principles of intelligent tutoring. Instr Sci 14, 293–326.

doi:10.1007/BF00051825
Page, L., Brin, S., Motwani, R., Winograd, T., 1999. The PageRank Citation Ranking: Bringing Order to the Web. (Technical Report No. 1999–66). Stanford InfoLab.
Panagiotopoulos, I., Kalou, A., Pierrakeas, C., Kameas, A., 2012. An Ontology-Based Model for Student Representation in Intelligent Tutoring Systems for Distance Learning, in: Iliadis, L., Maglogiannis, I., Papadopoulos, H. (Eds.), Artificial Intelligence Applications and Innovations, IFIP Advances in Information and Communication Technology. Springer Berlin Heidelberg, pp. 296–305. doi:10.1007/978-3-642-33409-2_31
Pazzani, M.J., Billsus, D., 2007. Content-Based Recommendation Systems, in: Brusilovsky, P., Kobsa, A., Nejdl, W. (Eds.), The Adaptive Web, Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 325–341. doi:10.1007/978-3-540-72079-9_10
Perkins, D.N., 1992. Technology meets constructivism: Do they make a marriage. Constructivism and the technology of instruction: A conversation 45–55.
Piaget, J., 1954. The construction of reality in the child. Routledge.
Plotly, 2016. 7 Plotly Graphs in 3D: Stocks, Cats, and Lakes [WWW Document]. Plotly Blog. URL http://blog.plot.ly/post/101360048217/7-plotly-graphs-in-3d-stocks-cats-and-lakes (accessed 11.26.16).
Premlatha, K.R., Geetha, T.V., 2015. Learning content design and learner adaptation for adaptive e-learning environment: a survey. Artif Intell Rev 44, 443–465. doi:10.1007/s10462-015-9432-z
Prieto, M., Garcia, F., 2006. METHADIS: Methodology for the Design of Adaptive Hypermedia Systems for Learning based on Learning and Cognitive Styles, in: Sixth International Conference on Advanced Learning Technologies, 2006. Presented at the Sixth International Conference on Advanced Learning Technologies, 2006, pp. 1137–1138. doi:10.1109/ICALT.2006.1652661
Psyché, V., Bourdeau, J., Nkambou, R., Mizoguchi, R., 2005. Making learning design standards work with an ontology of educational theories, in: 12th Artificial Intelligence in Education (AIED2005). IOS Press, pp. 539–546.
Rasch, G., 1960. Probabilistic Models for Some Intelligence and Attainment Tests. MESA Press.
Reckase, M., 1974. An interactive computer program for tailored testing based on the one-parameter logistic model. Behavior Research Methods & Instrumentation 6, 208–212. doi:10.3758/BF03200330
Reckase, M.D., 2009. Multidimensional Item Response Theory. Springer New York, New York, NY.
Ricci, F., Rokach, L., Shapira, B., 2011. Introduction to recommender systems handbook, in: Recommender Systems Handbook. Springer, pp. 1–35.
Rich, E., 1979. Building and Exploiting User Models, in: Proceedings of the 6th International Joint Conference on Artificial Intelligence - Volume 2, IJCAI'79. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp. 720–722.
Riding, R., Cheema, I., 1991. Cognitive Styles—an overview and integration. Educational Psychology 11, 193–215. doi:10.1080/0144341910110301
Riding, R., Rayner, S., 2013. Cognitive Styles and Learning Strategies: Understanding Style Differences in Learning and Behavior. Routledge.
Robotham, D., 1999. The application of learning style theory in higher education teaching. GDN Discussion Papers.
Roediger, H.L., III, 2008. Relativity of Remembering: Why the Laws of Memory Vanished. Annual Review of Psychology 59, 225–254.

doi:10.1146/annurev.psych.57.102904.190139
Roskam, E.E., Jansen, P.G.W., 1984. A New Derivation of the Rasch Model, in: Degreef, E., Buggenhaut, J.V. (Eds.), Trends in Mathematical Psychology, Advances in Psychology. North-Holland, pp. 293–307.
Rossi, F., van Beek, P., Walsh, T., 2006. Handbook of Constraint Programming, Foundations of Artificial Intelligence. Elsevier Science.
Rumetshofer, H., Wöß, W., 2003. XML-based Adaptation Framework for Psychological-driven E-learning Systems. Journal of Educational Technology & Society 6, 18–29.
Saaty, R.W., 1987. The analytic hierarchy process—what it is and how it is used. Mathematical Modelling 9, 161–176. doi:10.1016/0270-0255(87)90473-8
Sacerdoti, E., 1977. A Structure for Plans and Behavior. Elsevier North-Holland, Amsterdam and New York.
Schafer, J.B., Konstan, J., Riedl, J., 1999. Recommender systems in e-commerce, in: Proceedings of the 1st ACM Conference on Electronic Commerce. ACM, pp. 158–166.
Schmeck, R.R., 1983. Learning styles of college students., in: Schmeck, R.R., Dillon, R. (Eds.), Individual Differences in Cognition. Academic Press, New York, NY, USA.
Schult, D.A., Swart, P., 2008. Exploring network structure, dynamics, and function using NetworkX, in: Proceedings of the 7th Python in Science Conferences (SciPy 2008). pp. 11–16.
Schunk, D.H., 1991. Learning theories: An educational perspective. Macmillan Publishing Co, Inc, New York, NY, England.
Scott, J.P., Carrington, P.J., 2011. The SAGE Handbook of Social Network Analysis. Sage Publications Ltd.
Seidel, R.J., Perencevich, K.C., Kett, A.L., 2005. Suggestions for the Integration of Technology, in: From Principles of Learning to Strategies for Instruction. Springer US, pp. 185–209.
Self, J.A., 1990. Bypassing the intractable problem of student modelling, in: Gauthier, G., Frasson, C. (Eds.), Intelligent Tutoring Systems: At the Crossroads of Artificial Intelligence and Education. Intellect Ltd, pp. 107–123.
Sheehan, K., Lewis, C., 1992. Computerized Mastery Testing With Nonequivalent Testlets. Applied Psychological Measurement 16, 65–76. doi:10.1177/014662169201600108
Shrivastava, P., 1983. A typology of organizational learning systems. Journal of management studies 20, 7–28.
Siemens, G., 2005. Connectivism: A learning theory for the digital age. ITDL.
Simon, T., Binet, A., 1904. Méthodes nouvelles pour le diagnostic du niveau intellectuel des anormaux. L'année psychologique 11, 191–244. doi:10.3406/psy.1904.3675
Singhal, A., 2012. Introducing the Knowledge Graph: things, not strings. Official Google Blog.
Skinner, B.F., 1974. About behaviorism. Vintage.
Sleeman, D.H., Smith, M.J., 1981. Modelling student's problem solving. Artificial Intelligence 16, 171–187. doi:10.1016/0004-3702(81)90009-6
Sosnovsky, S., Dicheva, D., 2010. Ontological technologies for user modelling. International Journal of Metadata, Semantics and Ontologies 5, 32–71. doi:10.1504/IJMSO.2010.032649
Sosnovsky, S., Gavrilova, T., 2006. Development of Educational Ontology for C-programming. International Journal ITA 13.

Stage, C., 2003. Classical Test Theory or Item Response Theory: The Swedish Experience (No. 90). CENTRO DE ESTUDIOS PÚBLICOS, Santiago de Chile, Chile.
Stansfield, J.L., Carr, B.P., Goldstein, I.P., 1976. Wumpus Advisor I. A First Implementation of a Program That Tutors Logical and Probabilistic Reasoning Skills. AI Memo 381. (No. 381), Lab Memo. Massachusetts Institute of Technology, Cambridge, Massachusetts.
Tabaa, Y., Medouri, A., 2013. LASyM: A Learning Analytics System for MOOCs. International Journal of Advanced Computer Science and Applications (IJACSA) 4.
Thissen, R.., Mislevy, R.J., 1990. Testing Algorithms, in: Wainer, H., Dorans, N.J., Flaugher, R., Green, B.F., Mislevy, R.J. (Eds.), Computerized Adaptive Testing: A Primer. Lawrence Erlbaum Associates, pp. 101–134.
Thorndike, E.L., Linacre, J.M., 2000. The Rasch Model derived from E. L. Thorndike's 1904 Criteria. Rasch Measurement Transactions 14, 763.
Triantafillou, E., Georgiadou, E., Economides, A.A., 2008. The design and evaluation of a computerized adaptive test on mobile devices. Computers & Education 50, 1319–1330. doi:http://dx.doi.org/10.1016/j.compedu.2006.12.005
Triantafillou, E., Pomportsis, A., Demetriadis, S., 2003. The design and the formative evaluation of an adaptive educational system based on cognitive styles. Computers & Education 41, 87–103.
Triantafillou, E., Pomportsis, A., Demetriadis, S., Georgiadou, E., 2004. The value of adaptivity based on cognitive style: an empirical study. British Journal of Educational Technology 35, 95–106. doi:10.1111/j.1467-8535.2004.00371.x
Triantaphyllou, E., 2000. Multi-criteria Decision Making Methods, Applied Optimization. Springer US.
Truong, H.M., 2015. Integrating learning styles and adaptive e-learning system: Current developments, problems and opportunities. Computers in Human Behavior.
Tseng, J.C.R., Chu, H.-C., Hwang, G.-J., Tsai, C.-C., 2008. Development of an adaptive learning system with two sources of personalization information. Computers & Education 51, 776–786. doi:10.1016/j.compedu.2007.08.002
Tsiriga, V., Virvou, M., 2004. A Framework for the Initialization of Student Models in Web-based Intelligent Tutoring Systems. User Modeling and User-Adapted Interaction 14, 289–316. doi:10.1023/B:USER.0000043396.14788.cc
Tsiriga, V., Virvou, M., 2002. Initializing the student model using stereotypes and machine learning, in: 2002 IEEE International Conference on Systems, Man and Cybernetics. Presented at the 2002 IEEE International Conference on Systems, Man and Cybernetics, pp. 404–409. doi:10.1109/ICSMC.2002.1173446
UNESCO, n.d. The Learning Continuum [WWW Document]. URL http://www.unesco.org/new/en/education/themes/strengthening-education-systems/quality-framework/technical-notes/learning-continuum/ (accessed 6.15.15).
Van Plon Verhagen, B., 2006. Connectivism: a new learning theory. Bijdrage van Pløn Verhagen (University of Twente) 29, 08–12.
Van Rossum, G., Drake Jr, F.L., 1995. Python tutorial. Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands.
Vandenberg, R.J., Lance, C.E. (2000), 2000. A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. Organizational Research Methods 3, 4–70.

Vandewaetere, M., Desmet, P., Clarebout, G., 2011. The contribution of learner characteristics in the development of computer-based adaptive learning environments. Computers in Human Behavior, Current Research Topics in Cognitive Load Theory, Third International Cognitive Load Theory Conference 27, 118–130. doi:10.1016/j.chb.2010.07.038

Vas, R., 2016. STUDIO – Ontology-Centric Knowledge-Based System, in: Gabor, A., Ko, A. (Eds.), Corporate Knowledge Discovery and Organizational Learning, Knowledge Management and Organizational Learning. Springer International Publishing.

Vas, R., 2007. Educational Ontology and Knowledge Testing. Electronic Journal of Knowledge Management 5, 123–130.

Vaughan, K., Cameron, M., Aotearoa, A., others, 2009. Assessment of learning in the workplace: A background paper.

Wang, H.-C., Li, T.-Y., Chang, C.-Y., 2006. A web-based tutoring system with styles-matching strategy for spatial geometric transformation. Interact. Comput. 18, 331–355. doi:10.1016/j.intcom.2005.11.002

Weber, C., 2016a. Context-Aware Self-Assessment Path Generation for Personalised Education. SEFBIS 10, 53–66.

Weber, C., 2016b. STUDIO: A Solution on Adaptive Testing, in: Gabor, A., Ko, A. (Eds.), Corporate Knowledge Discovery and Organizational Learning, Knowledge Management and Organizational Learning. Springer International Publishing.

Weber, C., 2014. Enabling a Context Aware Knowledge-Intense Computerized Adaptive Test through Complex Event Processing. SEFBIS 9, 66–74.

Weber, C., Gabor, N., Vas, R., 2016. STUDIO: A Domain Ontology Based Solution for Knowledge Discovery in Learning and Assessment, in: Proceedings of the Special Interest Group for Education (SIGED) 2016. Presented at the Special Interest Group for Education (SIGED), Dublin, Ireland.

Weber, C., Truong, H.M., Vas, R., 2015. Context-Aware Self-Assessment in Higher education, in: EDULEARN15 Proceedings. Presented at the 7th International Conference on Education and New Learning Technologies, IATED, Barcelona, Spain, pp. 5910–5920.

Weber, C., Vas, R., 2016a. Applying connectivism? Does the connectivity of concepts make a difference for learning and assessment? (In publishing), in: Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics. Presented at the IEEE International Conference on Systems, Man, and Cybernetics, IEEE, Budapest, Hungary.

Weber, C., Vas, R., 2016b. TOP-DOWN OR BOTTOM UP: A COMPARATIVE STUDY ON ASSESSMENT STRATEGIES IN THE STUDIO ADAPTIVE LEARNING ENVIRONMENT, in: Proceedings of the European Distance and E-Learning Network 2016 Annual Conference (in Press.). Presented at the European Distance and E-learning Network 2016 Annual Conference, EDEN, Budapest, Hungary, pp. 41–49.

Weber, C., Vas, R., 2015. Studio: Ontology-Based Educational Self-Assessment, in: Workshops Proceedings of EDM 2015 8th International Conference on Educational Data Mining, EDM 2015, Madrid, Spain, June 26-29, 2015., CEUR Workshop Proceedings (CEUR-WS.org). Madrid, Spain, pp. 33–40.

Weber, C., Vas, R., 2014. Extending Computerized Adaptive Testing to Multiple Objectives: Envisioned on a Case from the Health Care, in: Kő, A., Francesconi, E. (Eds.), Electronic Government and the Information Systems Perspective,

Lecture Notes in Computer Science. Springer International Publishing, pp. 148–162.
Weber, G., Specht, M., 1997. User modeling and adaptive navigation support in WWW-based tutoring systems, in: User Modeling. Springer, pp. 289–300.
Welch, E.R., Frick, T.W., 1993. Computerized adaptive testing in instructional settings. Educational Technology Research and Development 41, 47–62. doi:10.1007/BF02297357
Wenger, E., 1987. Artificial intelligence and tutoring systems: computational and cognitive approaches to the communication of knowledge. Morgan Kaufmann.
Werquin, P., 2007. Terms, concepts and models for analyzing the value of recognition programmes, in: Report to RNFIL: Third Meeting of National Representatives and International Organisations, Vienna, October. pp. 2–3.
Willingham, D.T., Hughes, E.M., Dobolyi, D.G., 2015. The Scientific Status of Learning Styles Theories. Teaching of Psychology 42, 266–271. doi:10.1177/0098628315589505
Wolf, T., 2015. Assessing the criticality of IT projects in a portfolio context using centrality measures, in: Smart Enterprise Engineering. 12. Internationale Tagung Wirtschaftsinformatik, WI 2015. Tagungsband. pp. 706–721.
Wright, B.D., 1988. Practical adaptive testing., in: Rasch Measurement Transactions. Rasch Measurement SIG.
Wu, S.-H., Hsu, W.-L., 2002. SOAT: a semi-automatic domain ontology acquisition tool from Chinese corpus, in: Proceedings of the 19th International Conference on Computational Linguistics-Volume 2. Association for Computational Linguistics, pp. 1–5.

12. Appendix

12.1. Extending the Path-Based Assessment to a Framework for Path-based Strategies

In the example of a bottom-up path-based test, while assessing detailed knowledge-elements, failed higher level knowledge-elements could be decomposed into areas of missing knowledge. This additional information would for the drill-down approach become accessible for later, quite complete passed tests and may increase the number of learning iterations for mastering the domain knowledge. Opposing to these additional learning profits, a single bottom-up test would take in the average, and even more for early tests, considerable longer than a top-down test and psychological suggests a less strict process knowledge compliance, which may be outside of the assessment and learning policies implemented in an organisation. Considering these two exemplary algorithms as a starting point, other algorithms are possible, based on other learning theories and philosophies like cognitivism. As addressed as part of Section 0, the construction of a learning and teaching approach has to happen in the presence of a clear instructional design. So, a conscious design, aligned to the goals of teaching and learning is needed, e.g. by explicitly addressing the desired learning outcomes and identifying the right arrangements to meet the requirements of the context of learning.

The choice of an assessment algorithm for a specific assessment goal – in presence of different strategies – has to include an initial analysis on the requirements of the assessment, including considerations on a fitting instructional design. E.g. in cases of a large-scale selection of well-prepared learners for the assignment on specific job profiles the strict drill-down testing scheme may be more suitable as it aligns stronger to the logic of organisational processes. Vice versa, to pre-filtered groups of candidates or to isolate training needs, the bottom-up assessment may provide a wider profile of the capabilities of each individual and enable a more profound decision.

To further implement the freedom to allow different configurable assessment and evaluation solutions the path-based assessment and evaluation is aided by an additional extension. Paths can be traversed based on different selection strategies to select the next concept to assess in line with the context of the assessment. The evaluation of single concepts in terms of passing can be based on single or multiple concepts. Furthermore, the final consideration for the overall evaluation result can be based on different acceptance criteria. Yet, considering the freedom of adaptive systems, further adaptation within the assessment process would be possible. What is missing is a way and strategy to configure the assessment and evaluation process. The proposed extension, is a first step to a more configurable and controllable adaptive assessment experience. The solution incorporates two pillars: the introduction of assessment stages and the integration of a configuration frame.

Assessment stages are targeted to enable the creation and management of different assessment phases, which implement a modified concept selection behaviour for sections of the assessment and enforce a specific exploration of the concept group. Stages are intended as a sub-modification and extension to the overall assessment and shouldn't break or challenge the intention of selection of the pre-selected assessment strategy (as for modelling tasks in general – the degree of modification is finally in the hands of the designer). In this regard, it is important to scale how much of the overall concepts are selected in which stage of the assessment with what modification. To especially support the path-based assessment, four stages are defined and implemented as a starting point, described in Table 31. The stages are open for extension and represent exemplary the stages implemented for this work to test different assessment behaviour:

Table 31: Proposed stages for a more flexible assessment.

Stage	Description	Parameters
Stage 1: Initialization	Pre-selection and pre-ordering of the overall set of concepts accessible for the test. The initialization can derive information about the domain to improve the later assessment.	(pot. maximum of considered concepts, longest paths, etc.)
Stage 2: Randomized Assessment	The assessment engine will select a given number of concepts at random to be assessed before the selected assessment strategy. The random phase can help to variate the exploration of the concept group by providing a different starting seed (passed/non-passed concepts).	Number of random selected concepts to assess.
Stage 3: Assessing Required Concepts	The assessment engine will select all concepts which are referenced by a “required knowledge of” relation and interprets and assess them as a strict requirement to pass. Again, the specialized assessment phase	Number of required concepts to assess.

	can help to variate the exploration of the concept group by providing a starting seed.	
Stage 4: Returning to the Regular Assessment Strategy	The assessment engine will continue to follow the regular assessment strategy, e.g. the concept importance based assessment.	Parameters are selected based on the selected overall assessment strategy.

A configuration frame for the implementation and instantiation of assessment and evaluation strategies can be given in different ways and focused on different aspects of the system and strategies. The current extension considers three main aspects – the blueprint of the assessment and evaluation class, an evaluation object which encapsulates settings for the stages of assessment and evaluation and a customizable listing of relations which are considered during the assessment and evaluation.

The blueprint of the assessment class goes back to the initial implementation of the STUDIO system and defines a set of functions which have to be supported to implement an assessment and evaluation strategy as e.g. “starting the test”, “suspending” and “resuming”. The defined extension additionally requires to define an instance of the list of considered relations and the evaluation object.

The list of considered relations, gathers the relations which are relevant for the specific assessment and evaluation strategy. The strategy may be flexible in the range of allowed relations and can be parameterized and adapted through reframing the considered relations. E.g. cases could be defined where non-listed relations are omitted for the calculation of the concept importance measure of each concept or concepts could be filtered above the initial tailoring to sets which don’t include concepts which are connected to omitted relations.

The selection of the evaluation object sets the specific evaluation algorithm which is used for the assessment. The object furthermore encapsulates a range of parameters for the assessment stages and the evaluation itself. Potential stage-parameters are collected in Table 31. Evaluation specific parameter can regulate the threshold for the evaluation of single concepts or groups of concepts or – in the case of concept groups – the weighting of concepts for evaluation within the groups – depending on the selected evaluation (e.g. path-based) and the selected assessment strategy (e.g. concept importance based assessment).

Together, the extension is providing a starting point for a more thorough and detailed parameterization of the assessment and evaluation of learners and finally for an extended adaptation to different detected indications about the learner. While parameters could even be changed on-line and on-demand, a careful tracking and evaluation is needed to discover and establish the equally extended goals of adaptation and the boundaries of the resulting required parameterization.

12.2. A Heuristic Approach for Finding the Maximum Path Length in a Concept Group

The extension of the concept importance measure by the “detail-distance” needs the maximum path length to enable the normalization of the “concept-distance”. In a complex structure this maximal path length is costly to retrieve. The problem is known as the “longest path problem” and is NP-hard and not solvable in polynomial time for a random graph. Even though the structures, extracted from the STUDIO domain ontology have a feasible complexity and size (concept groups tend to include not much more than 100 concepts, while the domain ontology defines ~5000 concepts at the time of this writing), the provided implementation has to address the problem.

The solution for finding the longest path in the frame of STUDIO is heuristic in nature. In contrast to the frame of heuristic solutions for the longest path problem e.g. in the field of scheduling (Ajwani et al., 2012), the STUDIO extracted structures are not guaranteed to be acyclic (loops are allowed) or strictly directed (explicit modelling of inverse relations is allowed).

The implemented solution is triggered in line with the initialization of the path-based assessment and evaluation. In the initialization phase the software solution collects all concepts for a given concept group. The system then initializes all relevant values, which are used to track – concept-centred – the assessment and enables the continuous evaluation. To do so, the concepts are initialized recursively, starting with the start-concept, following the directly connected concept for the concept in focus and finally re-calling the initialization for the selected concepts. This way the initialization spreads in parallel branches of the concept group – from the start-concept in the centre to the most distant concepts, as indicated in Figure 47.

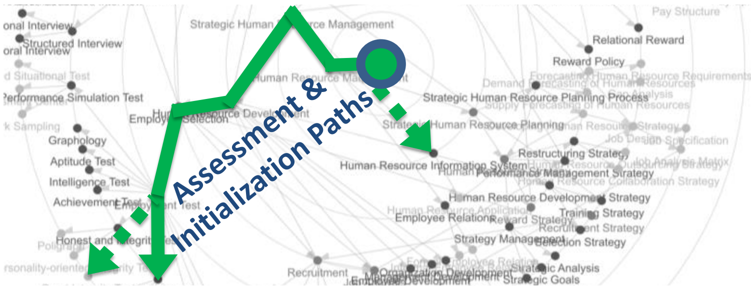


Figure 47: The initialization of the path-based assessment spreads in branches through the concept group – equal to the latter assessment.

To frame the recursion, the initialization keeps a record of visited concepts. Every concept which is in the list of visited concepts, can't be considered again within the initialization. Resulting, the rule 1) prevents loops and 2) frames which paths are considered in the initialization. Each time a recursive initialization is triggered for a concept which connects to no more concepts, a maximum path length is found and stored and the recursion starts the backtracking. Over the course of the initialization a set of longest path lengths is collected and the largest value within the set is selected as the longest path length.

Different paths may lead to the same concepts with different path lengths. Furthermore, within the initialization not all paths are discovered. As such, the "detail-distance" may also be detected differently during the assessment, depending on the specific considered path. The algorithmic compromise for this difference is to calculate the concept importance for a concept in a path every time the path is created. The values for the calculated concept importance is then tracked per concept across all generated paths. To finally derive the order of concepts for the assessment of one individual path, the average of the tracked concept importance measures for the concept is used. Consequently, the concept importance improves over the course of the assessment.

This approach can handle even complex structures, yet it may miss out the absolute longest path. The assessment creates paths different (see Section 8.4.1) and may encounter a path which is longer than the already considered paths. The new maximum path length is stored and leads to a different scaled run-time consideration of the "distance-detail" dimension. Yet, the rational of the "detail-distance" is to provide small value differences for concepts with an equal set of relations and an equal concept importance measure. The value of the differentiation of concepts in this case outweighs the potential error in the dimension "detail-distance" itself.

Figure 48 visualizes the “Managing Business” concept group. The numbers entailing the concept name show the individually calculated concept importance values of the concepts[13]. “Managing Business” is the start-concept. “International sales agreement” and “EU Internal Market” are both concepts at the end of a path, even though in different locations of the concept group. Both concepts on the right are connected through the same relation type and have the same number of relations. The minor difference in the calculated and visualized concept importance measure, goes back to the “detail-distance”, which successfully differentiates the concepts based on the distance from the start-concept.

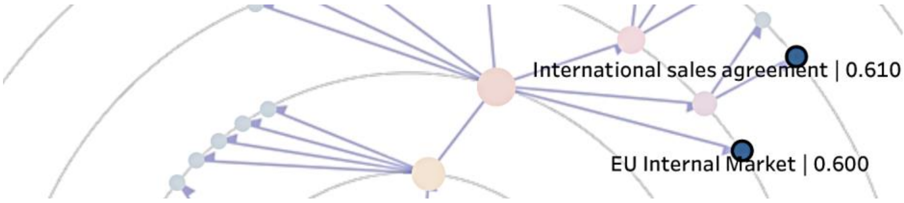


Figure 48: Concept group visualization, showing the concept importance based measure.

12.3. An Algorithmic Extension for the STUDIO Assessment to Resolve Learning Traits Biases

In this section, the requirements, concepts, and technologies of enhancing the STUDIO learning system are presented, which offers a systematic solution for personalised self-assessment. The STUDIO system uses a model of different educational areas, represented as a domain ontology. By testing students based on the domain structure, the system detects the real knowledge of each student in a complex and developing setting. To adapt to students, a test has to have the ability to react depending on the student’s performance, and trigger questions by taking into consideration the always changing and improving profile of the student. Based on this assumption a new frame of rules will be introduced that enables the student-profiling and determines how to branch the assessment for the individual student. This study is based on (Weber et al., 2015).

[13] The final implementation of the concept importance integrates an additional normalization of the [0, 100] interval of the dimensions to an [0, 1] interval, which results in smaller final results than shown here.

12.3.1. Creating and Maintaining Tests

Within STUDIO, the creation and maintenance of the domain ontology is a continuous task of ontology engineering. As the maintainer of the ontology the ontology engineer, – also called the ontologist, – is responsible for creating, using and evaluating the ontology (Neuhaus et al., 2011). Within the range of tasks, maintaining the structure and content has a strong focus, in favour of a more reduced set of formalized constraints. As such, within STUDIO, the task of creation and maintenance is a guided process. The system defines a specialized administration workflow which splits into a basic set of three consecutive task areas. Each task is supported by different, defined user roles with scaled access rights. The details are described in Table 32.

Table 32: The tasks, roles and concept levels involved in the ontology maintenance.

Task	Role	Concept level	Description
Ontology engineering	Ontologist	Ontology (class instances only)	Creating and linking of instances of knowledge-elements into the overall domain ontology.
Test definition	Ontologist/ Tutor	Concept Groups	For each new assessment test, relevant knowledge areas are selected and grouped into specialized containers called Concept Groups (CG). The concept groups could be organized further into trees of concept groups depending on the target of assessment. The final tree is equivalent to a sub-ontology. Concept groups are organized internally, based on the overall domain ontology and include all relations between knowledge elements, which are also defined within the domain ontology.
Question and learning material creation	Tutor	Knowledge Elements	Questions and learning materials are directly linked to single knowledge areas and could be created and extended on demand. Based on the designed concept group tree (test frame), existing questions and materials are imported, from the repositories. More questions and learning materials are defined now by the tutor, completing the assessment frame to meet the need of the target education. Every extension of questions and learning material is then also available for future assessment tests.

Independent of new content extensions to the domain ontology, the central structure of classes and relations, is fixed, acting as a blueprint for all instances. A view on the administration interface of the system is given in Figure 49. The left area pictures circular visualization of an excerpt of the ontology, while the right area pictures the question listing and editing interface. Each tab gives access to additional views for editing, including the learning material management and interfaces to modify node relations and node descriptions.

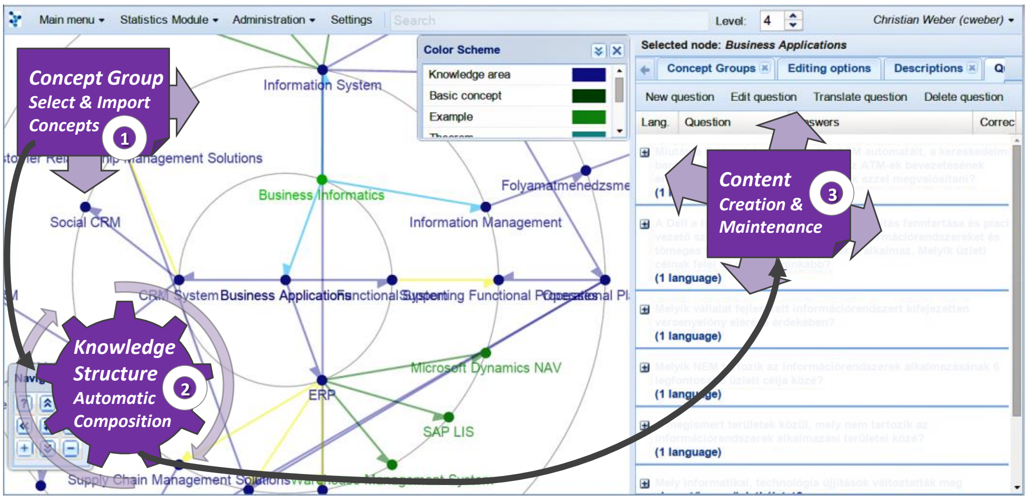


Figure 49: The STUDIO administration and maintenance interface.

12.3.2. Adaptive Self-Assessment in the Context of STUDIO

Adaptive self-assessment in the sense of STUDIO, addresses two dimensions of adaptivity:

- **Adapting to the student’s performance:** The knowledge structure, represented by the domain ontology, is a conceptualization of specific knowledge areas of education. Through adapting the exploration of the knowledge structure during testing, based on the responses of the student and the knowledge structure of the given domain (modelled in the ontology), the system implicitly draws a map of the student’s knowledge over time. This knowledge map then is used to adapt the learning feedback to the student’s performance
- **Adapting to the student’s profile:** Each student comes with personal traits and different learning backgrounds. Accessing and triggering questions and materials in connection to the known student’s learning style, means triggering questions

with a known lower bias, which enables the system to adapt better to the true core of the student's knowledge.

Additionally, adapting in the sense of self-assessment could be described as a compromise between the "time to learn" and the "time to test". The previous sections isolated that one main enabler for a better learning success is the continuous cycle of learning and assessment. This comes with two pitfalls which are opposing the potentials for an improved overall learning time:

- **Time to test:** Assuming a non-limited repository of questions, an extensive test lowers the concentration of the student and the readiness for continuing, while a too short test may collect not enough information to conclude on framing recommendations for learning.
- **Time to learn:** An unframed recommendation for learning material will result in long learning sequences including already memorized concepts, while few recommendations lead to an unbalanced relation between learning time and assessment time.

12.3.3. Context-aware Assessment paths

The domain ontology is a shared conceptualization of the represented knowledge, while the test-model, defined by the tutor, is the blueprint for the self-assessment test. The result is a knowledge-structure as a sub-ontology of the domain. This structure is the input for the adaptive test algorithm, which will then continuously trigger self-assessment questions till the knowledge structure is explored. Within the assessment, the created sub-ontology is used as a directed graph, where the highest knowledge area is a start element. Coming from the start-element, the testing algorithm then only uses detailing relation-types, essentially creating a knowledge tree, starting with the most general concept and branching to the most detailed concepts.

In line with the observation of a potential bias of questions, Vandenberg and Lance (Vandenberg and Lance, 2000) underlined, that a missing assessment of differences across assumptions may lead to an inaccurate reasoning. In the cultural context, this phenomenon is identified and formalized by Differential Item Functioning (DIF). In case a person with the same underlying ability has a different chance to answer a question correctly because of differences in the cultural background, this question is showing DIF.

Makransky and Glas (Makransky and Glas, 2013) modelled here a solution for Computerized Adaptive Testing, virtualising a question with DIF and attaching a context label to account it differently according to the cultural context, while still assessing the same question. For the following concept, the idea will be adapted in the opposite direction.

Following the rules of the path creation, two framing assumptions are introduced to ease the introduction of the algorithmic extension:

- **No “required_of” relations:** In contrast to previous publications it is assumed that there are no knowledge-elements with a “required_of”-relation, which may modify the number of needed child-elements to pass a parent-knowledge-element to fulfil the requirement criteria.
- **No hierarchical weighting:** The position of a knowledge-element within the knowledge structure has no impact on the weight of the element within the evaluation, neglecting the different implications of knowledge-elements, based on the position within the hierarchy of concepts.

Following the idea of Makransky and Glas and adapting it to the current frame, knowing a learning style for a question means that for a matching learning style the question is nearer to the core of the true knowledge of a student. As such the question has less learning style-based bias, while, when having an opposing learning style the question has a higher bias. Further, translated to the extent of known information about the question, this means that an answer to a question with a known learning style has a higher value for discriminating between persons than an answer to a question with a non-known learning style. In the latter case, the system would have to ask more questions to get a similar impact of the feedback, countering the potential but unknown bias.

In this regard, an adaptive testing algorithm, incorporating the knowledge about the learning style of the student and the assessment questions has access to two types of modifiers: 1) changing the number of questions which are asked for each knowledge-element to counter a known bias and 2) changing the weight of the questions within the evaluation, based on the discriminating power of the question - all in dependency of the detected learning style combination of student and content. The resulting modifiers are shown below in Table 33.

Table 33: Fit between content- and student-context and resulting assessment modifiers.

Learning style of content and student is: / Modifier:	The same	Unknown	Different
Number of questions for assessment:	Lower/--	Unchanged/~	Higher/++
Weight of the question for evaluation:	Higher/++	Unchanged/~	Lower/--

Following, a fusion of the modifiers is used. To combine the influence for the algorithm and the evaluation alike, a label in dependency of the fit between the content- and student-learning style fit is introduced and a weight is attached to express the impact on the evaluation. This weight also acts as a modifier on the count of a question. The number of questions will be compared against a fixed maximum number of questions asked and evaluate per knowledge-element. In this regards a supporting learning style introduces a higher weight for the evaluation, since it has a higher discrimination potential and will also count as more than one question against the maximum number of assessed questions per node. The assessment function will then use this assessment algorithm steps to select and evaluate questions for each knowledge-element, while a picture of the later process is shown in Figure 50:

1. **Labelling:** Label all questions for the current pool with a count modifier, where the labels are given in dependency of the influence of the learning style on the evaluation power of the question concerning the true knowledge of the student. A supporting learning style gains the label "positive", counting as 1.5, questions with an unknown learning style gain the label "neutral" with 1.0 and learning styles straining the connection gain the label "negative", counting as 0.5.

2. **Trigger question:** Select a random question from the pool, attached to the current knowledge-element. The order for selecting questions is: take all questions with an attached "positive" label, then take all questions with a "neutral" label and then take all questions with a "negative" label.

3. **Count:** Assess the result of the question. If the question were answered right, add the label counter to the variable Q_{pass}, else-wise add the counter to Q_{fail}.

4. **Evaluate:** If Q_{pass} is bigger than the evaluation border with: $Q_{pass} > Q_{max} * Q_{eval}$, where Q_{max} is the maximum number of assessed questions per knowledge-element and Q_{eval} is the percentage border which has to be overcome to pass, the knowledge-element is marked as passed. Further if $Q_{fail} \geq Q_{max} * Q_{eval}$, the

knowledge-element is marked as failed. In both evaluation cases the algorithm will stop assessing the knowledge-element and the overall assessment will select the next knowledge-element to assess. If none of the evaluation criteria is met, the algorithm will go back to step 2 and trigger the next question.

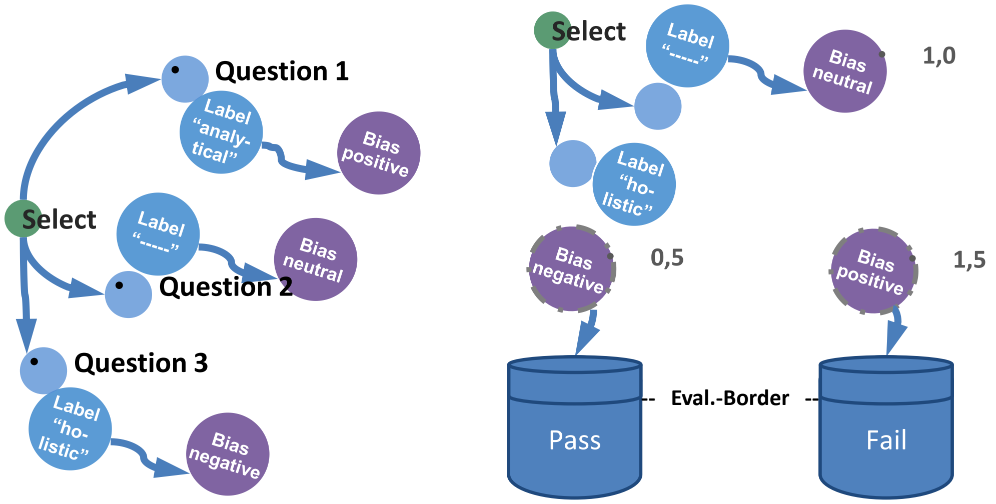


Figure 50: Using the detected question bias to modify the evaluation.

The consequence is: The higher the differentiation power of the questions is, the fewer questions are asked and the higher is the impact of the single questions. Therefore, the algorithm will stop assessing questions for a knowledge-element, as soon as the evaluation criteria is met. This way, the algorithm further strengthens the goal of improving the dependency between the "time to test" and the "time to learn".

The test itself is creating data over time about the responses and in case of an online detection of the learning style of the learner, the extended algorithm enables an additional implication: If a context is not known at the start of the assessment and gets detected within the assessment, the evaluation of passed and failed knowledge-elements will be re-evaluated at the time of the assessment. As the creation of assessment paths is a dynamic process, using the evaluation criteria and known responses of the student, the re-evaluation of the evaluation criteria in the light of the new "knowledge" about the student's learning style may re-open paths or close already taken paths in dependency of the changed observation. In this regards the picture of the student, created as a map based on the student's response, becomes more precise and the discrimination power of the assessment increases.

Within this frame, no measures to ensure an equal distribution of labelled questions is introduced. As such a not equally distributed repository of questions in terms of their labels, may introduce an additional bias. Also, the approach shown here were framed with assumptions to ease the introduction of the extended testing algorithm. Future publications will introduce here an extended set of variables and highlight further potentials for adaptation as highlighted in (Weber, 2014).

12.3.4. Feedback and Test results

Feedback plays an essential role in learning as it not only offers guidance for improvement but also motivate learners. It is widely agreed among educators and researchers that feedback should fit to individuals rather than using one feedback for all. Nevertheless, providing timely individualised responses is not always easy for teacher especially with high number of students for several tests. This is where the information of learning styles, in addition to their performances on adaptive tests, support to provide personalised feedback to fit to the ways students' process and understand information which offers benefits for both teachers and students themselves.

There are several ways where learning styles can contribute in a system like STUDIO. For example, the format can be adapted to include visual or verbal content to fit learners' input channel; languages can be changed to adapt to their ways of processing information such as more action verb for active learners etc. There are many research studies in the area of automatic feedback system which have shown very positive results. As far as being aware, while there are few detailed guidelines available on different matching strategies, it is required to evaluate the impact of future adaptive strategies on students' performances as well as other factors such as motivation before mass adaptation.

12.3.5. Learning Material Adaptation in STUDIO

The central concept of the STUDIO educational self-assessment is the continuous adaptation to the student's performance in cycles of testing and learning. To complete this cycle of continuously identifying and fulfilling knowledge gaps, personalised guidance on how the knowledge gap can be effectively fulfilled is essential. In STUDIO, one possible value of monitoring learning styles is that it can be used to personalise learning materials and learning activities for students. For visual learners, for instance, diagrams and presentations etc., can be provided, while for verbal learners, texts and lecture notes can be suggested. In the current STUDIO-extension, as it is still under development, there is not yet enough alternative learning material, supporting every style, to carry out the complete adaptation process for each knowledge area.

In addition, it may not always be easy for teachers to prepare several alternatives for the same content. Thus, while developing our learning material repository, STUDIO, firstly, provides a diverse set of learning materials that will cater different styles for every individual learner. Secondly, to overcome the limitation of a reduced set of overall learning material, at this stage the materials are structured into sections which correspond to different learning styles with their heading providing “clues” and “hints” emphasizing the preferences of different styles for easy navigation such as “examples”, “definitions”, “diagram” etc.

12.3.6. Conclusion and future directions

The STUDIO learning system offers a systematic solution for personalised self-assessment, using a domain ontology to capture the different areas of education. By making use of the domain structure for testing, a continuous profiling of students is possible and enables a better reflection of the learning progress for learners. Different people may have different learning styles and perceive information differently. This paper has shown how this bias in learning could be detected through an implementation of the proven Felder Silverman’s framework for learning style detection. Integrated into a new extension of the STUDIO learning system, the learning style detection directly impacts the learning process, improving the adaptation of testing and selection of learning materials as educational feedback to the individual student.

There are a number of work-in-progress fields, as well as exciting future development plans for the enhanced STUDIO system. Variable engineering and classification algorithm fine-tuning continues in order to improve the performance of the detection model. The next stage of adaptive assessment will make use of additional variables and exploit further the knowledge structure of the domain ontology for a better knowledge gap analysis. In addition, in future studies, different feedback strategies will be tested and evaluated to identify optimised strategies for each group of students.

12.4. Visualization Case Study: Alternate Domain Visualization Potentials

Section 8.3 defined the concept importance measure as an indicator of the importance of a concept in a domain ontology for the task of learning. Section 8.4 introduced an implementation of the measure in an assessment and evaluation algorithm to assess first concepts with a high concept importance measure to then learn first these concepts in the reflection phase – supported by STUDIO. The recommendation to learning happened here based on the concept importance measure and is supported by the visualization of the domain ontology. Section 8.4.6 summarized the feedback from the collected studies and compared the results for an evaluation, using an alternate visualization of the STUDIO domain ontology, extended by additional visualization features.

Figure 51 shows, for the domain of Human Resource Information Systems (HRIS), the STUDIO result graph visualization (left), visualizing all concepts in one size and in colour the information about passing and failing – informing the learner about the reached results and potentials for more learning. The right image shows an alternate visualization, visualizing by the size of the concepts the value of the concept importance indicator and with the colour the modularity class of a concept (as addressed in Section 8.4.6). While the latter image doesn't show information about the performance of the learner, it still indicates, based on the concept importance measure, which concepts are valuable to master. Similar, other visualization styles with different integrated information can help to explore the domain – to isolate more potentials or bottlenecks for learning or instruction. Some potentials are explored in the flowing, based on the centrality measures, introduced in Section 6.1.

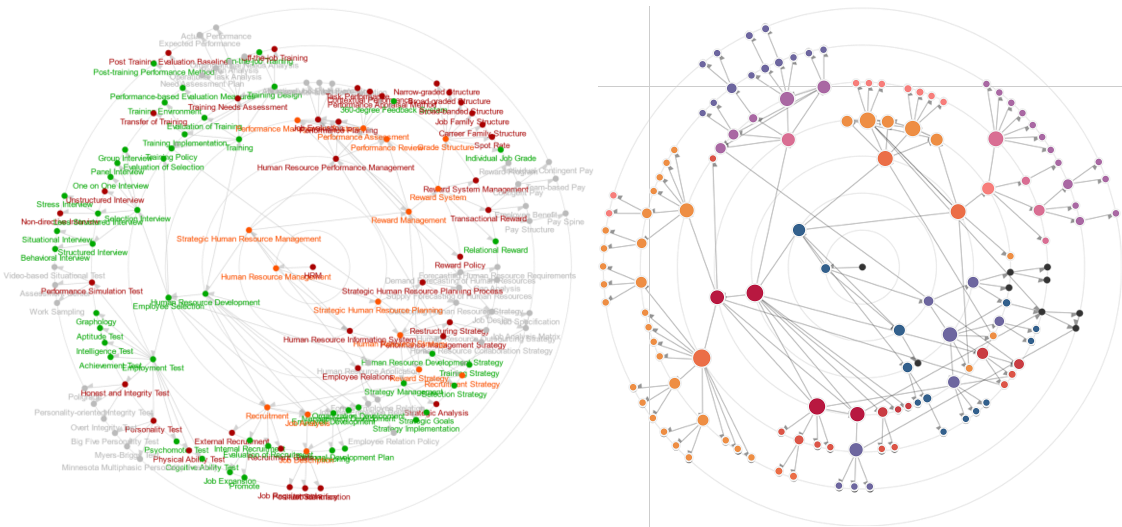


Figure 51: Comparing the STUDIO result visualization (left) and a concept importance measure based visualization (right) of the HRIS concept group.

Figure 52, visualizes the connectivity degree of each concept, considering incoming and outgoing relations. The degree in a straightforward manner, shows how well connected concepts are and in this regard, how well other concepts can be reached from the very same concept.

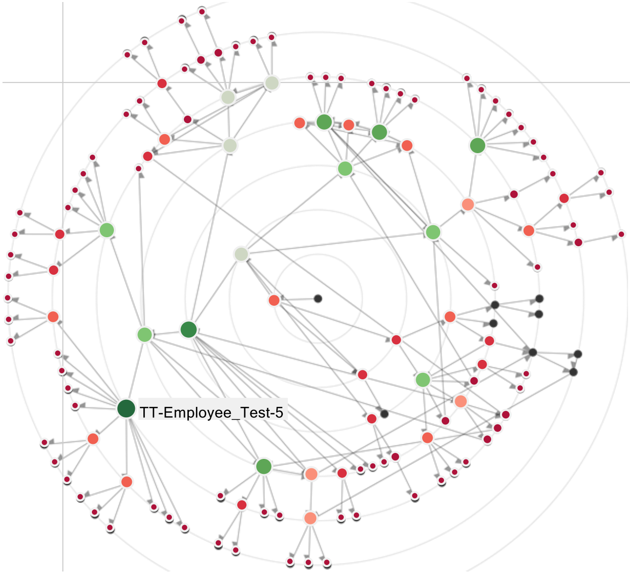


Figure 52: Degree centrality visualisation in the domain of HRIS.

The betweennes centrality, shown in Figure 53 tracks how often a concept is part of a shortest path between two concepts in the domain. As a result, betweennes rates concepts high which connect clusters of concepts and highlight bottlenecks and transition points – as e.g. in the transition in learning from one broader knowledge area to another and it also isolates concepts which are "needed" to connect areas of knowledge. As such the betweennes can be a structure-based extension as an importance dimension for the concept importance measure – as highlighted in Section 8.4.5.

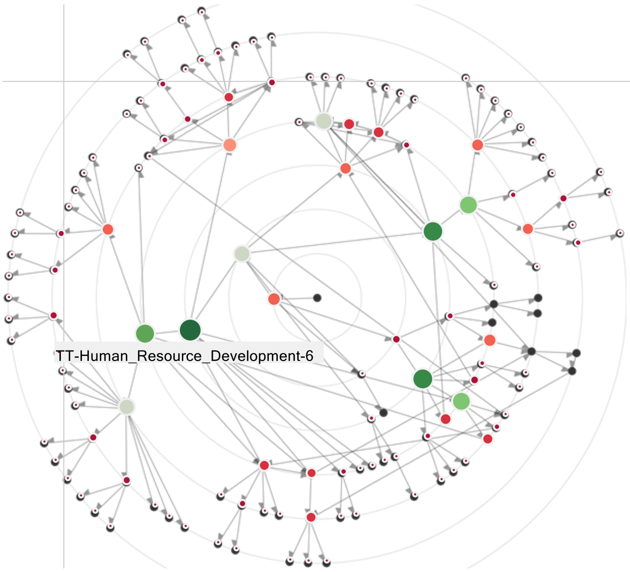


Figure 53: Betweennes centrality visualisation in the domain of HRIS.

Closeness, as shown in Figure 54, rates high concepts which are often on paths leading to "near" nodes and highlight concepts in the centre of clusters, which in the frame of learning may respond to concepts which are representing the surrounding concepts well.

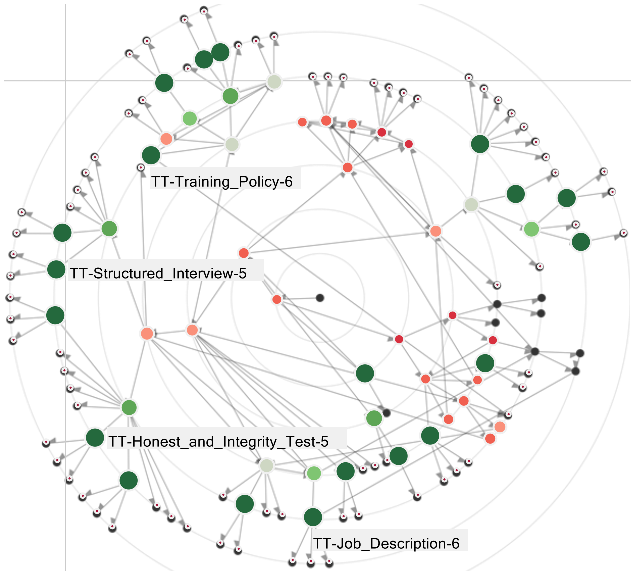


Figure 54: Closeness centrality visualisation in the domain of HRIS.

The eigenvector centrality in Figure 55, considers to how many other concepts a concept is connected. It is based on an iterative process, considering the whole concept neighbourhood, and converging over time. In the context of learning it can highlight alternate start-concepts which enable to reach "well" different parts of the domain.

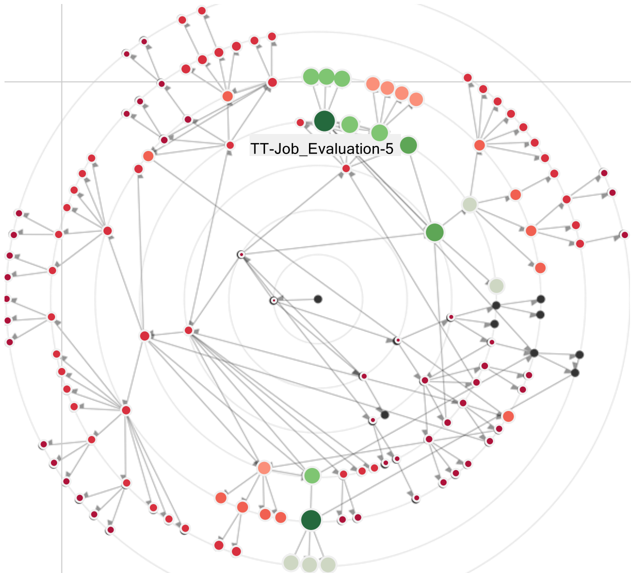


Figure 55: Eigenvector centrality visualisation in the domain of HRIS.

The PageRank centrality (Page et al., 1999), visualized in Figure 56, extends the idea of the eigenvector centrality and furthers boosts concepts which are point to by many other concepts or by a few boosted other "important" concepts. As such, concepts which have few incoming connections but are pointed to by many important concepts become important as well. In this regard, PageRank can point out the most influential concepts in a given learning domain and can support and motivate a reorganisation of the domain in terms of teaching and instructional design.

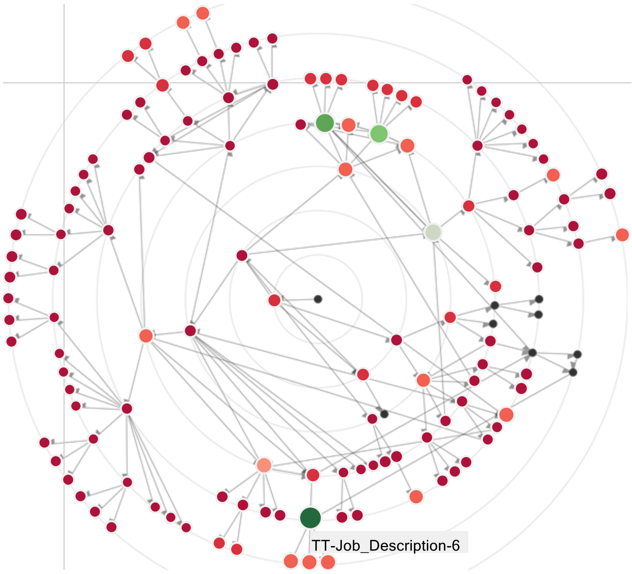


Figure 56: PageRank centrality visualisation in the domain of HRIS.