Doctoral School of Business Informatics

Thesis Summary

Christian Weber

Creating a Concept Importance Measure for Domain Knowledge in the Context of Learning
Ph.D. thesis

Supervisor:

Réka Vas, Ph.D

Budapest, 2017

Department of Information Systems

Thesis Summary

Christian Weber

Creating a Concept Importance Measure for Domain Knowledge in the Context of Learning
Ph.D. thesis

Supervisor:

Réka Vas, Ph.D

Table of Contents

1. Research Motivations and Goals 1
1.1. Background and Focus of the Thesis... 2
1.2. Detailed Research Questions 3
2. Methodologies.. 5
2.1. Research methodology 5
2.2. An explorative development methodology.. 5
3. Research findings.. 7
3.1. Application System Description and Exploration 7
3.2. Preparing Data Collection: Integration of an Event-tracking Framework 8
3.3. Design of a Concept Importance Measure for a Domain Ontology 8
3.4. System Integration of the Concept Importance Measure 15
3.5. Evaluation of research results.. 18
4. Selected relevant references.. 19
5. Relevant publications of the author 21

1. Research Motivations and Goals

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.

---

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

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

[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.

knowledge” and more specialized relations. Furthermore, the concepts can be different 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 concepts are 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?

2. Methodologies

2.1. 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 are 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.

2.2. An explorative development methodology

Following Amaral et al. (Amaral et al., 2011), exploration and design oriented research methodologies in the field of computer science can be divided into five methodologies: Formal, Experimental, Build, Process and Model.

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 is 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.

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. 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.

Build Methodology: This methodology is fundamentally about the building of a software system, documenting the planning, the composition, and the final process of building. The final "built" should be compared against existing solutions, if applicable, or additional claims as speed, space requirements, and other measurable.

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.

This research builds 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. The discovery and reasoning will be of an explorative nature throughout the whole research.

3. Research findings

The literature analysis of the thesis sheds light on a comprehensive fundament of topics in the context of user modelling and adaption, organisational learning and learning theories 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. 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.

The final explorative research of the thesis, is informed by and based on the indications from the literature and is organized in five blocks which are addressed in the following.

3.1. Application System Description and Exploration

This realization of this research is embedded into a technology enhanced learning system, which also supports the experiments throughout the work. The experiments are conducted in a blended learning environment, which supports the seminar work and studies of bachelor students. The results and continuous system feedback supports 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, collected within the literature study. The system models the domain related knowledge as an ontology, offering the needed structure and semantic 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 algorithmic extension, implementing a concept-importance aware assessment, will be integrated into the STUDIO system. The system further hosts the experiments and data collection.

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

To enable a later stage integration of the concept importance measure and to track and enable the planned experiments within the STUDIO system, a stocktaking and survey of the tracked and potentially trackable assessment and domain ontology related variables within STUDIO is conducted. 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 were yet missing. To improve the understanding of the assessment and learning process in STUDIO, a new data-collection framework is built and integrated.

The new framework enables 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 in the existing system, while a second challenge is 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).

3.3. 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, as addressed in (Weber and Vas, 2016a). 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 1 shows the STUDIO-based visualization of the Management Information System domain. The visualization resembles a network structure and is based on the underlying domain ontology.

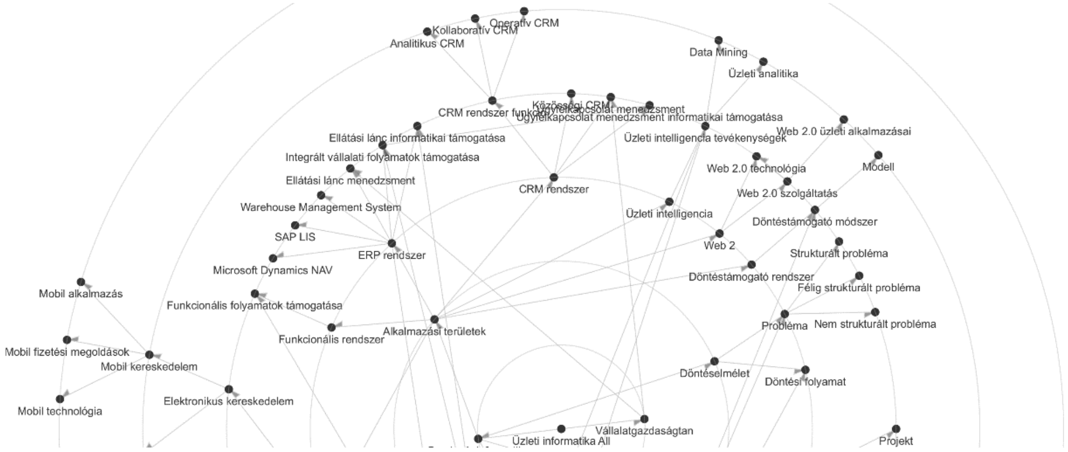

Figure 1: STUDIO based network representation of the Business Informatics concept group.

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 1 is based on a semantic hierarchy, which is grounded on the domain ontology. 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: connectivity, (semantic) complexity and a semantic-based hierarchy

Connectivism as an Interpretation for the Importance of Concepts for Learning

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. Taking into account the ontology perspective of STUDIO, three main pillars can be extracted from the concept of connectivism, coined by (Siemens, 2005). The pillars, following Siemens, are: 1) *learning … is derived as a competence from forming connections*, 2) *learning is motivated by connectivity, connecting experiences but also external information* 3) *connections which offer the learner to learn more are more important*.

Pre-Study: Analysing Connectivity and its Influence on the Assessment Performance of Students

The rationale of the 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.

Within two experiments of the study, 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 (Weber and Vas, 2016b). In a broader neighbourhood, this trend couldn’t be traced.

A Concept Importance Measure for a Domain Ontology

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.

The STUDIO domain ontology integrates a semantic model. Three features are 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; 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 connectivity of a concept is straightforward to translate into a measure, the translation of the semantic of concepts and their relations needs an interpretation (the *semantic complexity* and *semantic hierarchy*). To tackle the interpretation, two considerations are explored about the semantic of concepts and relations: 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?).

The Importance Dimensions of 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.

But, what parts of the ontology should be considered to map the semantic of concepts and relations into values? Concepts within the STUDIO ontology, as shown in Figure 2, 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.

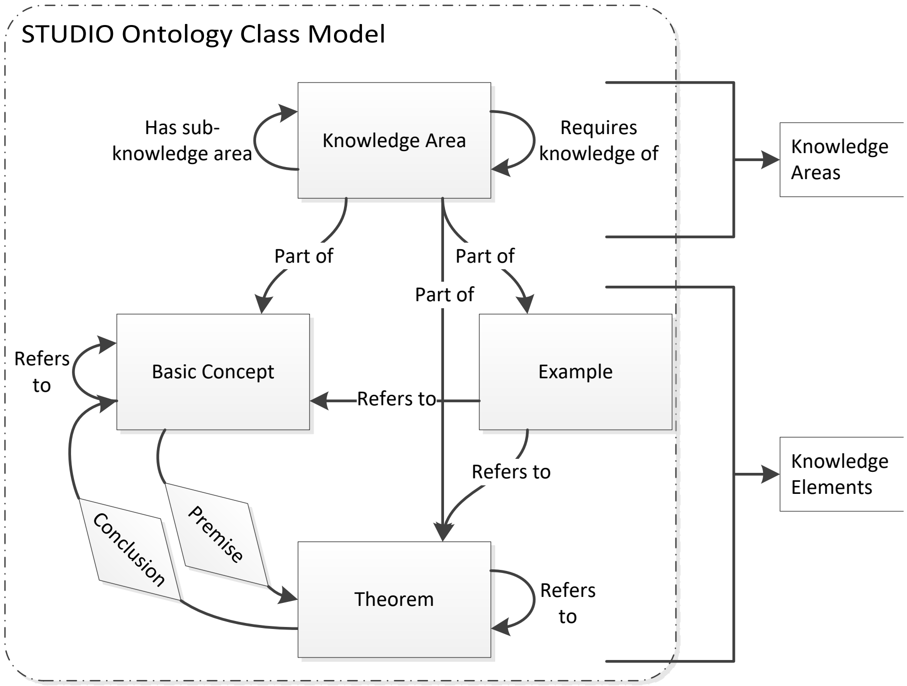

Figure 2: Grouping of concepts, visualized on the STUDIO domain model.

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. 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” as every relation expresses the “need” and “detail” of a connected concept.

The final measure then accounts for all relations a single concept shares and enables to compare individual concepts. As such, the measure implicitly integrates the connectivity of each concept by considering all relations of a concept at once – 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 that the connectivity can be a weak predictor for performance in a structured learning and assessment approach.

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

This work uses an elementary approach, motivated by but not following the approach of AHP, as AHP models a complex decision process while the concept importance measure targets to measures individual concepts. The assignment of numbers is expert modelled, translating the semantic of the relations to numeric values. The assignment of numbers 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 influence a measure for rating the importance of concepts for the goal of learning.

The domain ontology incorporates directed relations. To consider the connectivity degree of concepts, while equally considering the semantic type of the relation of an individual concept, this work considers incoming and outgoing relations simultaneously. To ease the modelling process and the human understanding of the translation of the importance dimensions of a relation to numbers, a relation is rated for a specific importance dimension on a scale of [0,100].

The translation 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 an 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.

The Numeric Interpretation of the Importance Dimensions

Multiple approaches (manual, semi-automatic, automatic) are applicable to define numeric values for each concept, reflecting the semantic of the relations and thus concepts in the context of the importance dimensions “need” and “detail”. 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.

The expert modelled mappings for the chosen importance dimensions of “need” and “detail” are shown in Table 1. 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 1: Collection and comparison of the expert modelled mapping values for the relations in the importance dimensions "need" and "detail”.

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**

Fusing Importance Dimensions to a Concept Importance Measure

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 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, as pictured in the matrix in Figure 3. So, to combine the importance dimensions into one measure, representing the concept importance, requires to define how to weight them.

To do so the measure considers three factors: **Dimension mapping:** the importance dimensions to translate for one individual relation; **Dimension weight:** the specific weight of the importance dimension for the context of learning; **Sum of relations per concept**: the different relations of an individual concept.

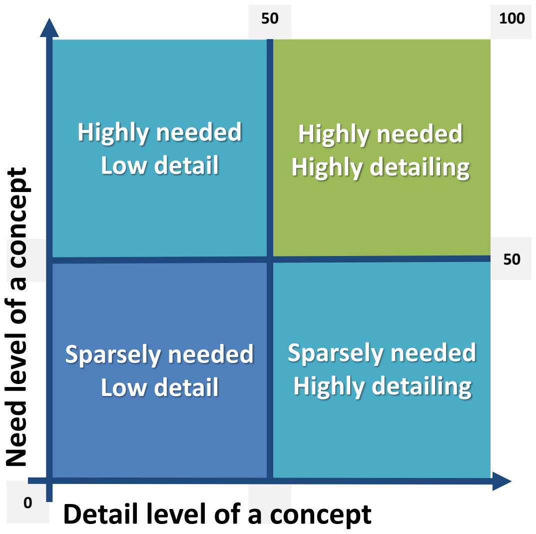

***Figure 3:** The concept importance matrix. More needed and more detailed concepts result in a higher value of the importance dimensions and the combined measure.*

To enable a direct and intuitive modelling approach for weights, weights are defined on an interval of [0,100] for each dimension and are set into relation to each other in the interval [0,1] by dividing each weight through the sum of weights. 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 $\sim 2\,/\,3$. 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:

$$\mathrm{Imp}_{\mathrm{Concept}} = \sum_{i=0}^{n} \sum_{j=0}^{m} w_j * \, Rel_{dim_{ij}}, \qquad for\ m\ dim.\ and\ n\ rel. \tag{1}$$

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 the implementing system.

3.4. System Integration of the Concept Importance

Measure

To integrate the measure, the STUDIO system implements three main pillars:

1. **Concept importance measure module:** the system integrates a new module to derive and manage the concept importance measure and to enable a system-wide seamless use of the measure.
2. **Path based assessment:** To utilize the concept importance measure, a new path-based assessment is created. The path-based assessment enables to create paths through the tailored domain ontology. From the created paths, concepts can be selected on demand, following different assessment strategies.
3. **Path-based evaluation:** A concept importance algorithm follows paths but selects the next concept to test within the same path and based on the highest value of the measure for a specific concept. The algorithm will make use of assessment paths and implements a tailored evaluation strategy.

A Path-based Exploratory Knowledge Assessment

To create a different solution which explores the structure freely – as needed for a concept importance measure implementation – the system is extended with a new open logic. The extension will act as a potential fundament for different smart knowledge exploration and evaluation solutions. A new knowledge exploration framework is conceptualized, which enables the implementation of various assessment and evaluation approaches in STUDIO. A challenge is to cope with the complexity of the semantic domain model and the assessment paths and how to translate the handling into the existing system. The new approach solves specific algorithmic problems as: loop-prevention and conceptual problems as defining the criteria for terminating the assessment.

A 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 to the start-element. With this new approach the tree shaped knowledge structure, which is extracted for each assessment from the domain ontology, can be interpreted as a set of paths from one knowledge-element to the start-element. To prevent loops in the path creation, the algorithm makes use of a strategy to black-list visited nodes, combined with a backtracking algorithm to create and explore alternative paths.

A Concept Importance Based Knowledge Assessment Algorithm

The integration, implementation and utilization of the concept importance measure is based on the path-based assessment and evaluation. The path-based assessment introduces paths to provide a trade-off between the exploration and the complexity of the domain and the underlying domain model. Every path represents one possible “walk” through the domain, starting with the start-concept and ending in another concept. The path can be considered as a window for the algorithm and the selection of the concepts on the path, while the path length is a flexible window size.

The integration of the concept importance based assessment selects 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.

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. The measure exploits the semantic and the connectivity of the domain ontology. 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. Only concepts which can be connected by a path of passed concepts to the start-concept are considered for evaluation, reflecting that a concept is understood in the context of the domain. The failing of a concept will change how the domain is explored. It will change the selection, as the algorithm selects concepts from each path by “moving” based on the measure and parts of the path may be omitted if they include failed concepts. The list of next concepts is then automatic rearranged, following the concept importance measure, as a part of the path isn’t available anymore, limiting the search space for assessment and latter learning and further addresses the desired adaptivity.

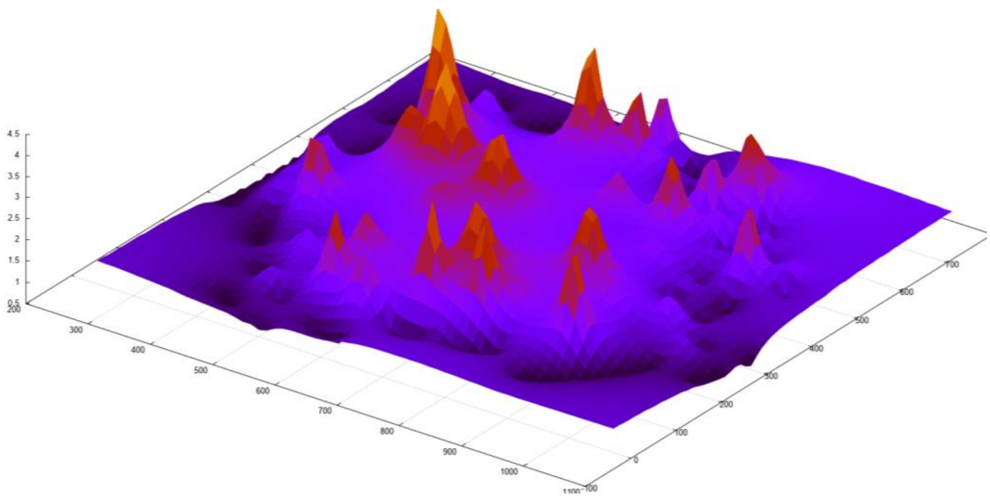

Figure 4: The measure converts the domain into a topology for learning.

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 4. 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.

3.5. Evaluation of research results.

In this subchapter I summarize the results of the research.

Learning – the process how new knowledge is acquired – is affected by several factors, including but not limited to the semantics and the structure of the given domain, the (ICT) infrastructure of the learning environment, the goals and modes of learning, and of course the motivation of the student. At the center of this research, the main question is "*how the learning effectiveness can be increased in a formal, blended-learning environment*"? How is it possible to find specific learning paths to facilitate the student in the process of the iterative self-assessment and learning?

In this research, the STUDIO system was used to store the ontological representation of the domain specific knowledge. STUDIO was further used as an e-learning system in the blended learning environment of the research. During the research, I found a connection between the most recent learning theory, namely the connectivism, and the network based interpretation of the domain model. How these network-like models can be travelled through also strongly resonates with the idea of connectivism. I found that those points of the network, which have a higher connectivity, are indeed more important related to learning.

Other qualitative attributes of the relations are also important, like dependencies and other characteristics which can be described with (semantic) complexity and semantic hierarchy. Based on these attributes I defined the importance of concepts (nodes, knowledge elements) in the thesis (first research question).

In order to answer the second research question, I implemented the theoretical construction of the concept importance measure into the STUDIO system. I also made experiments, in order to compare the importance based assessment and learning with the system's native adaptive testing algorithm. The results (numeric data etc.) of the experiments can be found in my publications. The experiments confirmed, that the created algorithm gives at least as good results as the original drill-down algorithm. The theoretical background of the concept importance measure was also validated with these experiments.

It was also confirmed that my method is becoming more and more advantageous as the volume of the domain knowledge, the complexity and the semantic structure broaden. Based on these findings I concluded that the experiments to validate the methods have to be continued, and controlled conditions have to be created. It is another task to draw in the motivation and the behaviour of the students to the set of test criteria.

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 to address the connectivistic learning idea – while its implementation is well-suited to support a flexible learning for rapidly changing requirements of the labour market.

4. Selected relevant references

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

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.
Bednar, A.K., Cunningham, D., Duffy, T.M., 1991. Theory into practice: How do we link. Constructivism and the technology of instruction: A conversation 17–34.
Brusilovsky, P., Kobsa, A., Nejdl, W. (Eds.), 2007. The Adaptive Web, Lecture Notes in Computer Science. Springer Berlin Heidelberg, Berlin, Heidelberg.
Carr, B.P., Goldstein, I.P., 1977. Overlays: A Theory of Modelling for Computer Aided Instruction.
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.
Coffield, F., Moseley, D., Hall, E., Ecclestone, K., 2004. Learning styles and pedagogy in post 16 learning: a systematic and critical review. The Learning and Skills Research Centre.
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.
Duffy, T.M., Jonassen, D.H., 2013. Constructivism and the Technology of Instruction: A Conversation. Routledge.
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.
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
Gabor, A., Ko, A. (Eds.), 2016. Corporate Knowledge Discovery and Organizational Learning, Knowledge Management and Organizational Learning. Springer International Publishing.
Gruber, T.R., 1995. Toward principles for the design of ontologies used for knowledge sharing. Int. J. Hum.-Comput. Stud. 43, 907–928. doi:10.1006/ijhc.1995.1081
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.
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
Kő, A., Gábor, A., Vas, R., Szabó, I., 2008. Ontology-based Support of Knowledge Evaluation in Higher Education, in: Proceedings of the 2008 Conference on Information Modelling and Knowledge Bases XIX. IOS Press, Amsterdam, Netherlands, pp. 306–313.
Kuckulenz, A., 2007. Continuing Training, in: Studies on Continuing Vocational Training in Germany, ZEW Economic Studies. Physica-Verlag HD, pp. 9–52.
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.
Lord, F.M., 2012. Applications of Item Response Theory To Practical Testing Problems. Taylor & Francis.
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
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.
Piaget, J., 1954. The construction of reality in the child. Routledge.
Siemens, G., 2005. Connectivism: A learning theory for the digital age. ITDL.
Singhal, A., 2012. Introducing the Knowledge Graph: things, not strings. Official Google Blog.
Skinner, B.F., 1974. About behaviorism. Vintage.
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.
Triantafillou, E., 2007. Computerized adaptive test–adapting to what?, in: Proceedings of the Informatics Education Europe II Conference IEEII. pp. 379–385.
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
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., 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.
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.

5. Relevant publications of the author

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 publishing). 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.

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.