Conclusion and Outlook

Kevin Fuchs

In this final chapter, we briefly summarize the INTUITEL project and the novelties it brings to the field of adaptive learning environments and learning analytics. INTUITEL itself is a system that operates on didactic knowledge and meta-knowledge respectively. However, it is not a tool for retrieving such knowledge. Therefore, in this section we give an outlook to current and future work to close this gap.

5.1 Summarizing INTUITEL

Adaptive learning environments seek to adapt the selection and presentation of learning content with respect to the learner’s individual characteristics. Amongst many more, this includes social, economic, cultural and ethical background, gender, physiological and psychological abilities and disabilities, the learning environment, the technical devices used, the learner’s prerequisite knowledge, her learning progress and physical or emotional states.

For example the impact of arousal and stress has been well explored and already utilized for the development of affective game designs [3, 84, 86, 106]. In the psychological field, there have been multiple suggestions to define learning styles and partly also how to derive appropriate course designs from them [43, 54, 55, 57, 68]. Large scale studies across Europe have disclosed significant correlations of learning success on the one side and factors like gender, cultural, ethnic, and socio-economic backgrounds on the other side [44, 45]. It is hence a worthwhile goal to enrich technology-enhanced learning with the capacity of detecting the above described factors and including them into decisions on how to guide a learner.

The INTUITEL system provides a novel and innovative way of transforming human-based didactic expertise into a formal, machine-processable representation. By using ontologies for the representation of both didactic knowledge and Didactic Factors we pattern human language and thinking.
Conclusion and Outlook

This way, no particular technical expertise is needed in order to create meaningful content. INTUITELE decouples technology from didactics entirely. The annotation of learning material is, therefore, intuitive and applicable to any domain of knowledge.

The result of INTUITELE-made deduction processes is a virtual tutor within a common Learning Management System like ILIAS or Moodle. However, INTUITELE does not enforce certain learning pathways to the learner. Instead INTUITELE only gives recommendations to the learner without obligation. The hypercube model and the calculation of cognitive distances between the learner’s state and recommended, predefined learning pathways makes INTUITELE constantly generate new recommendations regarding the learners actions.

As a theoretical foundation we may formulate four universal criteria a learning environment has to satisfy to be adaptive with respect to learning style, behavior and preferences of individual learners:

1. Indicators have to be found to detect factors of interest. These indicators have to be measured by or sent to the system constituting the adaptive learning environment.
2. The indicators must be aggregated to the desired factors and there must be a formal machine-processable representation of them.
3. In the context of the aforesaid formal representation, didactics and learning content must be associated with those factors that are supposed to have impact on learning behavior.
4. The learning environment must then deduce appropriate learner guidance from this formal representation.

5.2 Operationalization of Didactic Factors

With the INTUITELE project we developed an approach that is able to fulfill these requirements. However, the remaining problem is the identification and operationalization of influencing factors. This challenge is mostly located in the domain of social and educational science and forms a preliminary condition for the design of adaptive learning environments.

The INTUITELE system itself is capable of using such information for its recommendation process. Yet this information has to be provided for the system. INTUITELE does not generate the according data by itself. Therefore as a future task, we will introduce an approach on how to find and utilize influencing factors to use within an adaptive learning environment like INTUITELE.
5.2 Operationalization of Didactic Factors

Here, we particularly address time as an immanent dimension of learning. Time has a special meaning for learning in two aspects: first, learning always is a transformation of semantically connected content into a linear sequence along the time dimension. Second, many of the factors we supposed to have an impact on learning behavior and success are time-dependent functions. In a novel way, the following approach models learning histories of individual learners as spatio-temporal trajectories using techniques from the field of spatio-temporal databases. By this, we aim to provide a new way of learning analytics.

Let us briefly reflect on the foundations of INTUITEL. The formal representation of learning content, influencing factors and didactics is implemented by ontologies from which the INTUITEL system deduces learning recommendations. The set of these ontologies comprises the following parts:

- A pedagogic ontology based on the web didactics by Norbert Meder. This ontology organizes the learning content by courses, units, and atomic knowledge objects (KOs) [61, 65].
- A learner state ontology, describing the learner’s current state, progress as well as personal and environmental characteristics [25]. This ontology is also the part in which information about influencing factors is included.

In INTUITEL so called “Didactic Factors” as explained in Section 3.2 are of central meaning. They are identical with the influencing factors described previously. A Didactic Factor is aggregated by measuring one or more indicators. Indicators may represent non-nominal data. By particular transformation rules defined for each factor this non-nominal data is transformed into nominal data. This nominal data forms an identifier denoting an individual that is part of the INTUITEL ontologies. This way the factor exists in a formal and machine-processable form that we claimed before.

In order to calculate the learner’s position within the learning environment, INTUITEL uses the hypercube model as we elaborated in Section 3.3. Remember that each of the \( n \) dimensions of that hypercube represents a KO. Each of these dimensions is assigned a numeric value representing the state of progress a learner has performed on the according KO at a certain point in time expressed by a value from the interval \([0, 1]\). A learner’s position in this space is a vector \( L = (l_1, \ldots, l_n) \) evolving over time and thus drawing a trajectory in the \( n \)-dimensional space. The trajectory of any vector \( L \) is then located within that hypercube. Such trajectories must not be mistaken for the “learning pathways” that INTUITEL uses (see Section 3.3). Those learning pathways only consist of linear sequences of KOs with no explicit relation to the time dimension. However, a trajectory in the hypercube represents the entire learning history of an individual learner.
5.3 The Hypercube Database Project

The fact that a learner’s position is projected to a trajectory over time has not yet been utilized by INTUITEL nor has it in other systems known to the authors. Indeed, this is subjected by current and future work summarized under the “Hypercube Database” project [32].

INTUITEL provides the technology to use Didactic Factors for adaptivity in learning environments. However, those Didactic Factors still have to be operationalized. While the INTUITEL approach provides a universal process as well as the fundamental technology to associate Didactic Factors with learning content, it remains a challenging question how such factors are supposed to be identified and how they can be measured by sensor data.

The Hypercube Database project aims to combine learning analytics with the technology of spatio-temporal databases. Learning pathways of individual learners together with arbitrary indicators influencing learning behavior are measured and stored with a spatio-temporal database in the form of high-dimensional trajectories interpolated over the time dimension. For this purpose, we enhance the hypercube model of INTUITEL to utilize the trajectories of the learners’ positions together with measured sensor data that contribute to Didactic Factors.

5.3.1 The Advanced Hypercube Model

The hypercube model is enhanced by $k$ additional dimensions. This way the model describes an $(n + k)$-dimensional space with $n$ being the number of KOs in a learning environment and $k$ being the number of measured indicators. Each of the $k$ dimensions that stand for indicators is assigned a numeric value representing the value that is measured for this indicator at a particular point in time. A learner’s position in this space is now a time-dependent vector $M = (m_1, \ldots, m_n, m_{1^*} \ldots m_{k^*})$ forming a trajectory in the $(n+k)$-dimensional space. The $k$ dimensions may be normalized, which is not necessarily required.

Learners’ movements through this $(n + k)$-dimensional space define trajectories which we will model by the use of a spatio-temporal database that is – at the publication date of this book – yet in the process of development. The basic idea is to perform cluster-analysis solely on the basis of geometric relations between the trajectories of multiple learners in order to identify common learning pathways and learner groups. In subsequent analysis, we want to find out, which indicators correlate to these pathways in order to predict the learning behavior of individual learners.
5.3 The Hypercube Database Project

The major difference to common data analysis is the fact that all information – including the $k$ indicators – is transformed into purely geometric information and thus lifted to a highly abstract level. We only consider the geometric properties of and geometric relations between hyperpolylines. Arbitrary dimensions may or may not be included into analysis by simply performing projection. The model is entirely open to add and remove any kind of variables and dimensions respectively as long as they are numeric.

5.3.2 Example Applications

We now sketch two examples to illustrate how such data analysis can contribute to the improvement of adaptive learning environments.

i. Discovery of Unknown Didactic Factors: Within an experimental learning situation, arbitrary indicators are measured. The resulting data is transferred into the above described system and the data is converted into persistent learning histories together with their indicators. Using factor analysis, new Didactic Factors can be identified together with indicators that are represented by these factors. In a second step, the original set of indicators can be reduced to a smaller one, restricted to indicators that are easy to measure in a non-experimental learning environment.

ii. Real-Time Learning Pathway Prediction: Like in the previous example, learning histories as well as influencing indicators are stored with the advanced hypercube model. In a first stage, the learning histories are subjected to a cluster analysis in order to identify common classes of learning histories. In the second stage, taking these clusters on the one side and the measured indicators on the other side, one can perform for example either a discriminant analysis or a logistical regression. As a result, we can determine which variation of indicators of a specific learner will probably lead to a specific learning history. Built on this knowledge and measuring these indicators in the learning environment, i.e. in an LMS, we can predict the learner’s future learning history and recommend according learning pathways and KOs.

5.3.3 Implementation of the Hypercube Database

There are various approaches of existing spatial, temporal and spatio-temporal databases. Purely temporal databases are for example the ARCADIA database for clinical applications [19], Calanda for time series with financial data [89], ChronoLog running on top of a standard Oracle database [8], HDBMS [18],
TDBMS [101] and TimeDB for general purpose which is based on the ATSQL2 query language [15, 16, 95, 100].

The field of spatio-temporal databases is mostly dominated by Geographical Information Systems (GIS), Network and Facility Management, Land Information Systems (LIS) and Image Processing [1]. For example GRASS GIS [70] and GeoToolKit [5] are Geographical Information Systems while the CONCERT database focuses on management of raster images [81, 82]. The SECONDO database – developed at the University of Hagen – is a multipurpose system for spatio-temporal data [41, 105]. Due to the nature of their subject these systems mostly provide support for only two or three spatial dimensions. The DEDALE database is capable of dealing with higher dimensions and is based on a constraint database technique [36–39, 83].

All databases dealing only with two or three spatial dimensions are not an option for the Hypercube Database due to its high-dimensional space. The DEDALE system appears to be an interesting candidate because of its constraint approach that can be exploited for any number of dimensions. However, the constraint database approach is most appropriate for querying geometric objects containing infinite point sets whereas it is less suitable for querying continuous trajectories. We will therefore develop our own database but we will use the temporal database TimeDB as its back end and build the spatio-temporal functionality upon it.

At the publication date of this book, the implementation of the Hypercube Database is still in progress. Therefore the following elaborates current and future work. We describe the architecture of the software as well as the data structures and algorithms we intend to use. Figure 5.1 shows the fundamental parts of the system which is subdivided into the Vector Module, the Hypercube Module and the Database Access Module.

The system is written in Java and uses the temporal database TimeDB as a back end which is managed and accessed by the Database Access Module. TimeDB itself provides temporal support only for database tuples but not attribute-wise. The Database Access Module built upon TimeDB provides an interface for the Vector Module with which temporal support for single attributes is achieved.

The task of the Vector Module is the transformation of single measuring points (with respect to indicators) into temporal vectors and storing it via the Database Accessor Module. Consider an individual (a learner) for which we want to measure the values of m variables over time. For each variable we measure values at arbitrary points in time. For each variable we regard the last measured value as valid until a new value is measured. Alternatively,
we can also interpolate between two measured values. This way, we get an $m$-dimensional time-dependent vector for each individual. The listing below describes the transformation of measuring points into vector representation. We illustrate the subsequent insertion/deletion of measuring points and the evolution of the vectors for a trajectory with three variables $a_1$, $a_2$ and $a_3$. At the beginning all variables have an initial value, e.g., 0 from start to eternity.

\[
\begin{align*}
  a_1 &= 0 \text{ for } t \in [0, \text{ forever}) \\
  a_2 &= 0 \text{ for } t \in [0, \text{ forever}) \\
  a_3 &= 0 \text{ for } t \in [0, \text{ forever})
\end{align*}
\]

Now, we insert the measuring points $a_2 = 3$ at the time point $t_1$, $a_2 = 5$ at $t_2$ and $a_1 = 7$ at $t_3$

\[
\begin{align*}
  a_1 &= \begin{cases} 
    0 & \text{ for } t \in [0, t_3) \\
    7 & \text{ for } t \in [t_3, \text{ forever})
  \end{cases} \\
  a_2 &= \begin{cases} 
    0 & \text{ for } t \in [0, t_1) \\
    3 & \text{ for } t \in [t_1, t_2) \\
    5 & \text{ for } t \in [t_2, \text{ forever})
  \end{cases}
\end{align*}
\]
The Hypercube Module finally is responsible for transforming these vector data into spatio-temporal trajectories as described by the advanced hypercube model. Within the Hypercube Module we will implement indexing and querying functionality in order to access the trajectories efficiently and to cluster them by spatio-temporal characteristics.

There are multiple indexing techniques for spatio-temporal data. Many of them are based on the R-Tree family [42] for multidimensional spatial indexing. [40] lists, e.g., the 3D R-tree, the HR-tree, the RT-tree and the MR-tree. Moreover – for indexing moving objects with respect to the current time and the near future – [40] refers to TPR-trees, multilevel partition trees, kinetic B-trees and kinetic external range trees.

The usefulness of those indexing techniques strongly depends on the kind of the data and the kind of queries to be performed. In the case of the Hypercube Database we are less interested in querying point sets like “select the geographic region that was covered by the storm between 5 am and 7 pm”. Such a query would be useful within a Geographical Information Systems and would return a point set as a geometric object altering over time. But in our case we are mostly interested in queries referring to entire trajectories like “select all trajectories close to trajectory \( x \)”.

Appropriate methods for indexing and querying trajectories are for example the Spatio-Temporal R-tree (STR-tree) and the Trajectory Bundle Tree (TB-tree). Both index structures are appropriate for performing point, range and nearest-neighbor queries as well as trajectory-based queries [40]. However, using such indexing techniques like the STR-tree or the TB-tree are problematic for the high dimensional hypercube space. In particular, the performance of indexing rapidly decreases with the growing dimensionality. Addressing this particular issue, the X-tree is another index tree structure that is a variation of the R-tree and that is designed especially for high-dimensional data [7].

Another promising candidate is the grid index [17]. This index structure represents its bounding boxes in the form of static cells that are organized in a grid instead of a search tree. This structure is especially interesting for trajectories as they grow monotonically along the time dimension with few or no modifications after insertions. The grid index treats spatial and temporal indexing separately. This means that for temporal indexing any other method may be used. This way the grid index is a perfect candidate to be combined with the above described temporal index structure based on TimeDB.
5.4 Conclusion

We have defined four universal criteria a learning environment has to satisfy to be adaptive with respect to learning style, behavior and preferences of individual learners. Firstly, Didactic Factors have to be retrieved by measuring correlated indicators. Secondly, these factors have to be transformed into a machine-processable form. Thirdly, the Didactic Factors have to be annotated to learning content, together with didactic relations between pieces of learning content. Fourthly, the learning environment deduces the according instructional design from this formal representation.

INTUITEL satisfies the second, third and fourth of these requirements. With the Hypercube Database project we aim to close the gap to the first requirement, designing and developing a research tool for the analysis of learning histories. We model learning histories as spatio-temporal trajectories treating the time dimension as an immanent part of learning. Besides the learning content itself, the concept of the advanced hypercube also includes arbitrary additional data that may result from measured indicators. By this – inside the space of the advanced hypercube – data is lifted to a highly abstract level, mapped to purely geometric information.

This leads to a compact representation allowing us to analyze a wide range of data solely on the grounds of hyperpolylines, their spatio-temporal characteristics and their relations to each other. Not only is this a new application of a spatio-temporal database. It also offers a new approach for finding common learning pathways and Didactic Factors correlating with them. By this, we can predict learning pathways by observing a learners’ current actions and retrieving the according Didactic Factors, which constitutes the enhancement of adaptive learning environments in the future.

References


Conclusion and Outlook


References


Conclusion and Outlook


Conclusion and Outlook


Conclusion and Outlook


