A Computer Science Perspective on Learning and Knowledge Building

H. Ulrich Hoppe
Universität Duisburg-Essen / COLLIDE
http://www.collide.info
University of Duisburg-Essen (UDE)
founded by merger in 2013
11 Faculties

- Humanities
- Social Sciences
- Educational Sciences
- Economics
- Business Administration
- Mathematics
- Physics
- Chemistry
- Biology
- Engineering
- Medicine

- 39,343 students
- 442 professors
- 2,691 academic staff
- 1,375 employees technical support and administrative services (excluding Faculty of Medicine)
Our Department

... of Computer Science and Applied Cognitive Science in the Engineering Faculty (Duisburg)

17 professorships / groups (13 in CompSci)

Focus on Interactive Media and “Human-oriented Computing”

3 B.Sc./M.Sc. Study Programmes with approx. 1500 students
Welcome to the Collide Portal

Collaborative Learning Environments
Laborative Learning in Intelligent Distributed Environments

University Duisburg-Essen
Faculty of Engineering
Department of Computational and Cognitive Sciences
Building LF
Lotharstr. 63/65
47048 Duisburg, Germany
Fax +49 - 203 - 379 3557

Upcoming Events
- ICCS 2014 (10.06.14)
- ICLS 2014 (23.06.14)
- ICALT (07.07.14)
- ASONAM 2014 (17.08.14)
- CRMG 2014 (07.09.14)

Learning Analytics
Project History I

• **COLDEX** ("Collaborative Learning and Distributed Experimentation", 2002-05, coord. by Collide): collaborative modelling tools for various areas of science learning; creation of a globally accessible repository of learner-created models ("emerging learning objects").

• **SCY** ("Science Created by You", 2008-12, U. Twente): SCY-Lab environment facilitates learner experience through larger "missions"; sharing of emerging learning objects.
Project History II

• **Go-Lab** (2012-16, U. Twente): uniform access to a variety of remote laboratories, both virtual and physical; customisable web-based learning environments and learning analytics components.

• **JuxtaLearn** (2012-15, Open Univ. Milton Keynes): fostering understanding of core science concepts through video creation on the part of the students; collaborative production and discussion around videos.
Project COLDEX
("Collaborative Learning and Distributed Experimentation", 2002-05)

BioTube and Seismo Scenarios

The space planting scenario enables students to experiment with computer and robotic controlled plant growth chambers in the context of Advanced Life Support Systems.

Chemistry, biodiversity, seismology, astronomy ... these are the scientific topics which are combined in the COLDEX project. Originating from the pedagogical idea of "challenge based learning", we support student groups – from face-to-face groups up to international learning communities; they can have a realistic look inside scientific work. Various "digital experimentation toolkits" containing virtual and physical tools enable open-ended learning activities.

By using a synchronised "learning object repository" (LOR) learners can find people with similar interests. The LOR supports retrieval in a big pool of models and data, re-use of learning objects and building of learning communities between Europe and South-America.

Finding epicenter and hypocenter of earthquakes is the main task in the seismo scenario.
Project COLDEX
("Collaborative Learning and Distributed Experimentation", 2002-05)

The astro scenario contains remote control of telescopes, image processing and calculating moon crater heights.

Astro and Maze Scenarios
COLDEX – Results and „Lessons Learned“

• Notion of „emerging learning objects“ (ELOs)

• Provision of a general purpose modelling environment (*CoolModes / FreeStyler*) with different „palettes“ for various modelling languages (e.g., *System Dynamics*, *Petri Nets*) and special applications

• Support for sharing ELOs in larger communities through similarity based search

• Resistance of teachers against „open production“
Project SCY

Science Created by YOU

About SCY
Learn about the philosophy behind SCY and the outcomes

Impressions
Get a glimpse of the SCY Look and Feel

Partners
SCY is an international project with 12 institutions from Europe and Canada

News
SCY in the press, Publications etc.
In the SCY-Lab environment, students engage in collaborative inquiry learning on "missions" through sharing learner-generated objects (ELOs). Concept mapping is used as a tool for early conceptualisation activities in a new domain: students create concept maps from given resource documents.
Mission “CO2-Friendly House”

Scenarios and Learning Activity Spaces
SCY-Lab Environment
SCY – Results and „Lessons Learned“

• SCY-Lab as a powerful environment
  (yet – might be better delivered as a web application)

• Integration of feedback/scaffolding facilitated through multi-agent architecture

• „Mission Maps“ as navigation aids guide learners through complex inquiry processes

• „Complexity overkill“ with big socially relevant learning challenges
The Go-Lab Project

- Large scale use of online labs in education.
- The Go-Lab Portal offers students the opportunity to perform personalized scientific experiments.
- Teachers enhance their classroom activities with a personalized web-based environment.

http://www.go-lab-project.eu/
Experimentation with Remote and Virtual Laboratories

Remote lab: Students control real physical laboratories using a web-interface.

Virtual lab: Simulation of real-world processes.
Inquiry Learning at School

• Mapping of the inquiry cycle to virtual inquiry learning spaces.

• Open Social Apps support the activities in different phases.
Federation of Labs and Apps

Sharing of predefined inquiry learning spaces, apps and online labs in a lab repository.

Teacher assembles inquiry learning spaces for certain scenarios in the Go-Lab ILS platform.

Create url

Student view distributed by url.
Intelligent Student Support

- Supporting self-reflection through student dashboards.
Intelligent Student Support

- Intelligent feedback mechanisms.
**Learning Analytics**

*Learning Analytics* - a common denominator for all types of measurement, collection, analysis and reporting of data about learners and their learning contexts with the aim of understanding and optimising learning and learning environments.

- **big data? (MOOCs?)**
- *interest in algorithms & computational methods*
- *relation to educational data mining?*
Spectrum of Learning Analytics Topics and Applications

• Prediction of students at risk from academic records
• Monitoring of course participation on learning platforms (e.g. Moodle)
• „Open Learner Modeling“ as a tool for reflection support
• Discourse and argumentation analysis
The Methodological „Trinity“ of LA

Activity analysis:
- process-oriented
- sequence analysis (e.g. „process mining“)
- action patterns

Artefact analysis:
- product-oriented
- information / text mining
- „semantic richness“

Network analysis:
- social or actor-artefact relations
- network measures (centrality, cohesion, …)
- no inherent time
Enhancing LMS with analytics: LeMo

* Beuster/Elkina/Fortenbacher/Kappe/Merceron/Pursian/Schwarzrock/Wenzlaff [IDAACS 2013]
Example #1:

Analyzing student created concept maps (from project SCY)
Scaffolding of CM Construction

- More extreme weather
- Acidification of oceans
- CO2 increases
- Global warming
  - Increased temperature
  - Increased sea level rise
  - Fossil fuel increase
- Melting glaciers

Concept Suggestions:
- Maybe you consider adding the following:
  - Fossil fuel
  - Sea level rise
  - Temperature
- Concept already exists!
Agent Architecture

- SCYMapper
- Notifier
- domain ontology
- commands
- actions
- SQLSpaces
- ConceptMap Enricher (coordination)
- ConceptMap Modeller
- ConceptMap Proposer
Ontology-based Support / Matching

keyword extraction (LDA)

ontlogy augmentation

Global warming refers to the increase observed in recent decades in the average temperature of the lower atmosphere and the oceans and their expected future warming. Between 1906 and 2005, the average temperature near the ground by 0.74 °C (+/- 0.18 °C) has increased. The year 2000-2009 was by far the warmest on record, followed by the 1990s, which in turn were warmer than the 1980s. According to current scientific understanding of this, it is very likely the reinforcement of the natural greenhouse effect of human influence. The man-made warming caused by burning fossil fuels, deforestation, and world-embracing agriculture and livestock. As a result, the greenhouse gas carbon dioxide (CO2) and other greenhouse gases like methane and nitrous oxide accumulates in the atmosphere so that less heat radiation emitted by the Earth’s surface into space. By far the largest part of the base of the growing anthropogenic warming is due to the recent and still growing increase of the greenhouse gas carbon dioxide. Through strong feedback processes is the direct heating effect of the carbon dioxide but with significantly less likely than expected, resulting from the heating. Warming also secondary effect. By the year 2100, depending on future CO2 emissions and the actual response of the climate system to ensuring a warming of 1.1 to 6.4 °C. This would have a number of consequences. Increased glacier melt, sea level rise, ocean acidification, changes in precipitation patterns, more extreme weather events. Including with regard to the projected rise in sea level, many questions remain unanswered. The measured bandwidth of the rise in sea level by the end of the 21st Century is, depending on the scenario, 18 to 59 centimeters. The scenarios take into account not only the polar ice dynamics (e.g., the collapse of large ice sheets in Antarctica) and uncertainties in climate-carbon cycle feedbacks. In a warmer climate will reduce the capacity of the surface and the ice cover of the oceans. Sea levels rise, ice caps descend, and remove the salt at higher latitudes...
Quality Assessment of CMs

Comparison to expert map(s)

Generic approach using background knowledge (a domain ontology) and information extraction (LDA) + structural (graph-theoretical) measures
Concept Mapping Study

Participants:
37 high school students aged between 16-18

Procedure:
45 min - introduction to concept mapping
  (+ tool “SCY Mapper”)
45 min - reading/highlighting of a two page text on global warming followed by CM construction

experimental group received adaptive suggestions
control group could only access a dictionary
Expert Assessment

Four experts assessed the concept maps produced by the students …

based on two criteria (inspired by Marra, 2002): completeness regarding the concepts used connectedness regarding the relations introduced (adequacy, missing rel’s)
First Results

No significant differences between scaffolded and non-scaffolded conditions.

Predictive power of simple structural measures cannot be much improved by including of semantic features (ontology).
Correlations

**Correlation of structural measures with expert judgments**

<table>
<thead>
<tr>
<th></th>
<th>Criterion 1 (completeness)</th>
<th>Criterion 2 (connectedness)</th>
</tr>
</thead>
<tbody>
<tr>
<td># nodes</td>
<td>$r = 0.79$, $p = 0.000$</td>
<td>-</td>
</tr>
<tr>
<td># edges</td>
<td>$r = 0.74$, $p = 0.000$</td>
<td>$r = 0.63$, $p = 0.000$</td>
</tr>
<tr>
<td>density</td>
<td>$r = -0.53$, $p = 0.001$</td>
<td>$r = -0.26$, $p = 0.11$</td>
</tr>
</tbody>
</table>
A mathematical explanation ....

• In any graph **density** is **average degree** divided by #nodes.

• In „scale-free networks“ the **average degree** tends to be stable.

• Hence: If CMs evolve like SFNs the density is highest for the **smallest maps**!
The Evolution of Scale-free Networks:
Preferential Attachment

A SCALE-FREE NETWORK grows incrementally from two to 11 nodes in this example. When deciding where to establish a link, a new node (green) prefers to attach to an existing node (red) that already has many other connections. These two basic mechanisms—growth and preferential attachment—will eventually lead to the system’s being dominated by hubs, nodes having an enormous number of links.

A Network Perspective on Concept Maps

Concept maps are scale-free networks!

=> evolution of concept maps is governed by “preferential attachment”;

density decreases with growing networks

The Litmus test

Aggregated degree distribution (from 37 maps)
left: number of nodes per degree value, right: log-log graph
Ifenthaler, Masduki and Seel (Instructional Science, vol. 39, 2011) compare a number of general graph theoretic measures to identify changes in cognitive structures using concept maps from five consecutive stages of a learning process.
Example #2:

Extraction of semantic networks from textual artefacts created by students (from ongoing project JuxtaLearn)
Background:

EU Project JuxtaLearn (2012-2015)

- Provoking student curiosity and understanding of knowledge and technology through creative performance (concretely: film making – editing – sharing activities)

- Teaching and learning support is guided by threshold concepts

- First step: identifying such concepts and their appropriation by conducting face-to-face workshops with teachers and students

→ Learning Analytics techniques are used to extract representations of the underlying conceptual relations
Initial Workshops

- Teacher-student workshop
- Six A-level students
- Role reversal: students teach the teachers
- Topics: Chemistry - moles, Biology - alleles, Physics - potential energy

→ transcripts and summaries analysed using the AutoMap/ORA toolset for Network Text Analysis
Network Text Analysis

- Use of the AutoMap and ORA software tools for „network text analysis“ developed by the CASOS group at CMU

- Workflow:
  • collection of workshop transcripts (textual artefacts)
  • pre-processing and analysing with AutoMap
  • visualization through ORA-Netscenes

- Result:
  generation of multi-modal concept maps
  (with categories: actor, domain concepts, pedag. concepts ...)
Network Text Analysis - Process

Analysis Workflow

Text Extraction

Text Cleaning
- Delete List
- Stemming

Concept Identification
- Concept List
- Bigrams
- Named Entity recognition

Concept Generalization
- Generalization Thesauri

Concept Classification
- Meta-Matrix Thesauri

Network Analysis

Union Network
- Statistical Reports
- Visualization
- Measure Charts

Meta-Networks
- Over Time Measures/Visualization

Analysis of Dynamics
Concept Generalization

-> Generalization Thesaurus

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Teachers</td>
<td>teacher</td>
</tr>
<tr>
<td>3</td>
<td>Teacher</td>
<td>teacher</td>
</tr>
<tr>
<td>4</td>
<td>teachers</td>
<td>teacher</td>
</tr>
<tr>
<td>5</td>
<td>teacher</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Head of Technology</td>
<td>head_of_technology</td>
</tr>
<tr>
<td>7</td>
<td>Head of technology</td>
<td>head_of_technology</td>
</tr>
<tr>
<td>8</td>
<td>head of technology</td>
<td>head_of_technology</td>
</tr>
<tr>
<td>9</td>
<td>Head Technology</td>
<td>head_of_technology</td>
</tr>
<tr>
<td>10</td>
<td>Head technology</td>
<td>head_of_technology</td>
</tr>
<tr>
<td>11</td>
<td>head technology</td>
<td>head_of_technology</td>
</tr>
<tr>
<td>12</td>
<td>Head of science</td>
<td>head_of_science</td>
</tr>
<tr>
<td>13</td>
<td>Head of science</td>
<td>head_of_science</td>
</tr>
<tr>
<td>14</td>
<td>head of science</td>
<td>head_of_science</td>
</tr>
<tr>
<td>15</td>
<td>Head Science</td>
<td>head_of_science</td>
</tr>
<tr>
<td>16</td>
<td>Head science</td>
<td>head_of_science</td>
</tr>
<tr>
<td>17</td>
<td>head science</td>
<td>head_of_science</td>
</tr>
</tbody>
</table>

Concept Classification

- actor
- country
- general_concept
- (domain) knowledge
- pedagogical_concept
- role
- technical_concept
- tools_and_technologies

-> Meta Thesaurus
Teacher-Student Workshops - Results
„External“ Learning Analytics in JuxtaLearn

Extraction of information from *video comments* about:

- associations of concepts
  (adequate or inadequate from a scientific point of view)
- identification of concepts that are frequently addressed in questions as indicators of possible origins of comprehension problems
- associations between concepts often used in answers as indicators for missing relations in students’ mental models (⇒ misconceptions or “stumbling blocks”)
## Case study: Khan Academy

**Scaffolded questions & answers**

<table>
<thead>
<tr>
<th>Questions</th>
<th>Tips &amp; Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>I understand the process, but I don't see why the water has a better chance of going IN than going OUT. I mean, if the sugar molecule is blocking from the inside, doesn't that mean the water from outside won't be able to squeeze in either? Why is it any different from either side of the membrane? Either way, the sugar is blocking...</td>
<td></td>
</tr>
</tbody>
</table>

Show all answers · Answer this question
Artefact Analysis - the JxL Approach

- Textual annotations of videos as indicators for students’ understanding and potential misconceptions
- In addition to domain concepts “signal concepts” (difference between X and Y, help on X needed, etc.) indicate special relations

- Benefits for…
  - Researchers: insight into patterns of learning
  - Teachers: monitoring learning processes around videos
  - Students: system recommends videos or peer helpers
Data Selection / Extraction

- 1.284 comments from educational videos

<table>
<thead>
<tr>
<th></th>
<th>chemistry</th>
<th>biology</th>
<th>physics</th>
</tr>
</thead>
<tbody>
<tr>
<td># users</td>
<td>328</td>
<td>309</td>
<td>86</td>
</tr>
<tr>
<td># questions</td>
<td>279</td>
<td>184</td>
<td>70</td>
</tr>
<tr>
<td># answers</td>
<td>362</td>
<td>312</td>
<td>77</td>
</tr>
</tbody>
</table>

- Video topics:
  - Chemistry: *Mole* and *Avogadro's Number*
  - Biology: *Diffusion* and *Osmosis*
  - Physics: *Voltage* and *Electrical Potential*
Multimodal networks of
• domain concepts
• signal concepts

Examples:
• *I don’t understand* diffusion, but ….  
• What’s the *difference between* osmotic_pressure and solution??
Multimodal Analysis - disambiguated

Introduction of *combination nodes between domain concepts and signal concepts*
### Highlighting Results in Context

- **I finally understand osmosis. Thanks Khan!!**
  - 1: do\_understand osmosis thanks khan

- **How do I know if the membrane will allow sugar to diffuse or not? Plz any body reply.**
  - 1: explanation i\_know if membrane be allow sugar diffusion not plz any body reply

- **Khan Academy helped me to review a unit on OSMOSIS AND DIFFUSION in my BIOLOGY class!**
  - 1: khan\_academy help review unit on osmosis\_diffusion in biology class

- **Hey is it possible for you to create a video explaining the pressure flow theory of phloem transport?**
  - 1: hey be possible creation video explanation pressure flow theory phloem transport

- **Still confused about osmotic pressure / wasted a bit of time..**
  - 1: still confusion about osmotic\_pressure / waste bite time

- **How do the sugar molecules get out?**
  - 1: explanation sugar\_molecule get out
Embedment into JxL Process

Stimulate and support reflection & feedback by

– providing awareness of one’s own performance
– providing awareness about the learning process (students’ progress, group structure)
Short Summary / Outlook

- *Artefact analysis* can help to identify problems of understanding and misconceptions.

- Network perspective on knowledge artefacts facilitates new theoretical approaches.

- For learning analytics combinations of the perspectives (*artefact – activity – network*) are needed!
QUESTIONS – COMMENTS ?