Cognitive Modelling Goes To School

MODELLING CONVERGENT AND DIVERGENT PROCESSES IN SOCIAL EXPLORATORY SEARCH AND LEARNING

Tobias Ley
Tallinn University, Estonia

CELSA 2018 Keynote
21.10.2018
Research Group on

LEARNING ANALYTICS & EDUCATIONAL INNOVATION

Tallinn University, Estonia

http://ceiter.tlu.ee
Intelligent Tutoring Goes To School in the Big City

Kenneth R. Koedinger
Human-Computer Interaction Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213
USA
ken.koedinger@cs.cmu.edu
http://sands.psy.cmu.edu/ACT/people/koedinger.html

John R. Anderson
William H. Hadley
Mary A. Mark
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213
USA

Abstract. This paper reports on a large-scale experiment introducing and evaluating intelligent tutoring in an urban High School setting. Critical to the success of this project has been a client-centered design approach that has matched our client's expertise in curricular objectives and classroom teaching with our expertise in artificial intelligence and cognitive psychology. The Pittsburgh Urban Mathematics Project (PUMP) has produced an algebra curriculum that is centrally focused on mathematical analysis of real world situations and the use of computational tools. We have built an intelligent tutor, called PAT, that supports this curriculum and has been made a regular part of 9th grade Algebra in 3 Pittsburgh schools. In the 1993-94 school year, we evaluated the effect of the PUMP curriculum and PAT tutor use. On average the 470 students in experimental classes outperformed students in comparison classes by 15% on standardized tests and 100% on tests targeting the PUMP objectives. This study provides further evidence that laboratory tutoring systems can be scaled up and made to work, both technically and pedagogically, in real and unforgiving settings like urban high schools.
“This study provides further evidence that laboratory tutoring systems can be scaled up and made to work, both technically and pedagogically, in real and unforgiving settings like urban high schools.”

Koedinger, et al. (1997), pp.30
ADRESSING NEW CHALLENGES

- From Direct Instruction to Exploratory Learning
- Learning goals dynamically change in the learning process
- Learning is socially mediated
SOCIALY-MEDIATED SELF-DIRECTED LEARNING

Learning in Realistic Contexts

Individual Search and Sensemaking

Collaborative Knowledge Building

How working memory and long term memory interact to learn categories?

Role of Working Memory Capacity?

Collaborative Information Search on the Web

Discovery Learning in School

Modelling Memory Processes

Designing Support Mechanisms (Intelligent Learning Systems)
COLLABORATIVE KNOWLEDGE BUILDING & INDIVIDUAL LEARNING

- Individual and collective knowledge creation tightly linked

- Negotiation Processes in Social Software Environments
Collaborative Information Search in groups of university students (N=24)

CONVERGENCE IN THE GROUP STRENGTHENS INDIVIDUAL LEARNING

Semantic Stabilization in the Group leads to stronger memory traces for the used concepts.
Collaborative Knowledge Building

Collaborative Information Search on the Web

Role of Memory in Convergent and Divergent Group Processes
Individual Search and Sensemaking

Discovery Learning in Secondary School

Role of Working Memory Capacity
COLLABORATIVE KNOWLEDGE BUILDING

How to reach consensus?

How to deal with divergence?
How does consensus emerge?

Verbatim tag imitation:
Copying the word
(e.g. Dellschaft & Staab, 2008; Halpin et al., 2007; Rader & Wash, 2008)

Semantic imitation:
Tag-based Topic Inference
(e.g. Fu et al., 2009, 2010)

A MODEL OF SOCIAL TAGGING: FUZZY TRACE THEORY

(e.g. Brainerd et al., 2010)

Tag-based Search for Resources (at time n)

Perception of Tag t

Encoding

verbatim trace (direct access D) | gist-trace (reconstruction R)

Imitation of Tag t (time n+1)

explicit imitation \( p(t) = D \) | implicit imitation \( p(t) = R*J \)

\( J = \) Familiarity-based Judgement
AN EXPERIMENT USING THE RTTTT-PROCEDURE

(Brainerd et al., 2002)

Phase R: Incidental Learning of tags in a decision task

Phase T: Production of tags (previously learned or not)

N=39, age $M=32$
**MATERIAL USED:**
**PHOTOS AND TAGS**

<table>
<thead>
<tr>
<th>Fotografie 1</th>
<th>Fotografie 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
</tbody>
</table>

- **Aufgabe** lustig Unschuld Gasse
- **Moment lebhaft** Flaschen Mädchen
- **Jugend** Ecke lachen Neugier
- **Stadt spontan** Wein Augenblick

- ermüdet nass Stimmung scheitern
- **Jacke nach Hause** Sehnsucht
- **Emotion Mann Schutz** Herbst
- **Mangel überqueren** traurig
- Fußgänger leer
MEASUREMENT MODEL

Multinomial Processing Tree (adapted from Brainerd et al., 2010)
RESEARCH QUESTIONS

- What role do verbatim and semantic processes play when imitating tags?
  - D: Direct access to verbatim trace
  - R: Reconstruction of gist trace
  - J: Familiarity-based decision

- Can these processes be dissociated using practically significant variables?
  - Tag size
  - Semantic Layout of tag clouds (e.g. Lohmann et al., 2009; Rivadeneira et al., 2007; Schrammel et al., 2009)
  - Word connectivity: density of generated associations (Nelson et al., 1998)
Verbatim and Gist traces (for explicit and semantic imitation) are represented by independent parameters $D$ and $R$.

A good amount of tag-processing and production is implicit ($R$ and $J$).
Semantic imitation constant at around 13%

Verbatim imitation depends on conditions (20% vs. 8%)

Other influencing factors

Semantic Layout of Tags → R
Size of Tags → J
Connectivity of Tags → D
FROM A MEASUREMENT TO A COMPUTATIONAL MODEL

3Layers - A tag recommender based on verbatim and semantic processing

Based on ALCOVE – a connectionist model of category learning

Encode the resource in terms of topics or categories

Match the pattern to all previously encountered examples

Activate the tags from those examples

Draw tags depending on their activation
CAN THE ALGORITHM GUESS WHICH TAGS PEOPLE WILL USE?

- Using Wikipedia pages tagged in Delicious
- Algorithm recommends tags from previous tags of that user
- Reinforces convergence
In Social Tagging Research, **Convergence** is usually highlighted (consistent indexing of resources) and **Divergent Search** is important for effective problem solving (Lazer & Bernstein, 2010).
DIVERGENT PROCESSES IN INFORMATION SEARCH

- Trade-off: Fluency (Flat Associative Hierarchy) vs. Consistency (Steep Associative Hierarchy)

- Research Questions
  - Evidence for the trade-off?
  - Supporting information search with recommender?

- Study: Individual and collaborative information search at the workplace (one month, N=18)

COLLABORATIVE INFORMATION SEARCH MAKES PEOPLE MORE CREATIVE...

Collaborative search leads to higher fluency of associations …
... BUT LEADS TO LOWER LEVELS OF CONSISTENCY

... weaker relationship between topic and tags used
HOW TO SUPPORT CONSISTENCY DURING CREATIVE SEARCH?

- Recommender 1: *Most Popular* (MPT)
- Recommender 2: *Search of Memory* (SoMe)
  - Similar to the 3Layers Recommender
  - Based on Episodic Memory Model MINERVA2
Some has higher acceptance in collaborative search.
CONVERGENCE & DIVERGENCE

Collaboration leads to higher levels of creativity
But also to lower levels of consistency
Need a recommendation paradigm based on semantic processing
Individual Search and Sensemaking
Discovery Learning in Secondary School
Role of Working Memory Capacity
SELF-DIRECTED SEARCH AND SENSEMAKING
DISCOVERY LEARNING IN SCHOOL: SEARCH AND SENSEMAKING

Reflective Search Model

Information → Goal

Search

Attention → Control

Find

Information → Sensemaking

Nii näeb välja loom, kellel on on tiivasarnased nahkribad, ta on suur, ja tal ei ole pealuust väljaaluuvat harja.

Milline on sinu arvates tema nimi? Kliki vastaval nimel!

- Saurolophidae
- Isthodactylus
- Dorygnathus
- Pteranodon
- Ornithoscelida
- Scaphognathus
DISCOVERY LEARNING IN SCHOOL: SEARCH AND SENSEMAKING

- 8th and 9th grade pupils (N = 109)

- Independent Variable
  - Working Memory Capacity (WMC): high vs. low

- Dependent Variable
  - Learning Curves
  - Search and sensemaking in DinoNimi
Performance Difference between high and low Working Memory Capacity condition …

Group x Practice: F(6,318) = 2.24, p < .05

Group: F(1,55) = 4.41, p < .05

\( \chi^2(4) = 15.58, p < .01 \)

… is due to better retrieval of previously learned categories (sensemaking) …

… and more strategic search.

LEARNING CATEGORIES IN SUSTAIN

Inhibition of interference (Unsworth & Engle)
Empirical Pattern

Simulated Pattern
TAKE AWAYS
1. THE SOCIAL & INDIVIDUAL

- Coupling of social and individual learning
  - Convergence in group important for individual learning
  - Importance of social cues for creative thought
  - Mapping of internal and external (e.g. fluency and tagging)

- Cultural production of patterns and individual learning
  - Formal model of Sociocultural Learning Theory

2.

- Coupling is mediated by our memory processes
- Working memory and long term memory as a sensemaking machine
- Executive functions control attention and shield search and retrieval
- Cognitive Models mimic human exploratory learning
- … and allow diagnosis and support in open ended learning tasks
- “Learning Analytics” made theory-driven rather than data-driven
REFERENCES


07-11 January 2019
Tallinn, Estonia

http://winterschool.tlu.ee

Educational Innovation – Getting ready for the future
Educational Robotics in Preschool and Primary Education
Tobias Ley
Professor for Learning Analytics and Educational Innovation
Tallinn University, Estonia

Twitter @tobold
http://ceiter.tlu.ee
http://tobiasley.wordpress.com
Input Layer

Latent topics of new post

Semantic matrix $M_S$

Posts

$P_i$

Latent topics

Hidden Layer

Output Layer

Tag combination corresponding to pattern in $I$

Educational Sciences
HOW DOES CONSENSUS EMERGE?

Verbatim tag imitation: Preferential Attachment
(e.g. Dellschaft & Staab, 2008; Halpin et al., 2007; Rader & Wash, 2008)

\[ p(t) = p(a) \cdot p(o) \cdot \frac{N(t)}{\sum_{i} N(i)} \]

Semantic imitation: Tag-based Topic Inference
(e.g. Fu et al., 2009, 2010)

\[ p(t) = p(c \mid r) \cdot p(t \mid c) \]