Analysis and Assessment of Computer-Supported Collaborative Learning Conversations

Stefan Trausan-Matu
Politehnica University of Bucharest
stefan.trausan@cs.pub.ro
http://www.racai.ro/~trausan
Overview

1. Computer-Supported Collaborative Learning
2. Polyphony
3. Automatic chat analysis
4. Results
5. Conclusions
Computer Supported Collaborative Learning (CSCL)

A new paradigm in learning with computers (Koshmann 1999):
- **from** Intelligent Tutoring Systems (cognitive)
- **to** Computer Supported Collaborative Learning (socio-cultural)

Knowledge is constructed socially (Vygotsky)

The spread of forums, **chats**, blogs, wikis and folksonomies → learning in (on-line) virtual teams and/or communities
Experiments with chat-based CSCL

- K-12 students solving mathematics problems both individually and collaboratively in the VMT project at Drexel University, Philadelphia, USA
- Computer Science students at Bucharest “Politehnica” University, Romania at
  - Human-Computer Interaction course in Romanian and French – role playing and debate
  - Algorithm Design – problem solving
Virtual Math Teams
Drexel University, Philadelphia, USA

- Extend the Math Forum’s “Problem of the Week (PoW)” in mathematics
- Groups of 3 to 5 students
- Grades 6th to 11th
- 60-90 minutes moderated chats
- Non-routine mathematical problems
K-Teams,
Politehnica University of Bucharest

Groups of 3 to 5 students

Classes + about 1-2 hours of non-moderated chats
- each participant has a role in a chat debate on the subject of the course
- algorithm design
LTfLL - EU FP7 Project, 2008-2011

Language Technologies for Lifelong Learning

Netherlands, France, United Kingdom, Germany, Romania, Bulgaria

Language technologies considered:

- Chat (conversation) analysis
- Latent Semantic Analysis
- Ontologies (semantics)
- Semantic Social Networks
- Corpus linguistics
Dear 137, davidsf, Jason, and agnon,
It seemed to us that you had a very productive first session exploring the given pattern of sticks and squares. We were especially interested in the variety of strategies you used, such as constructing the next step of the pattern on the whiteboard, verifying the pattern in horizontal and vertical lines (other teams did that as well) and deriving a formula for that line.

As for the working time-related items, you built up each other’s ideas and tried to work with them in interesting ways. We find it very important that everyone felt comfortable asking the team to explain in detail the reasoning for the work completed (e.g., section 2.05, 05:47-48 PM EST), as they had not been told initially, and that as a team you provided that explanation. It looked helpful to us when your group tested together the formula you used and one question that was left unanswered was whether a pattern is found above, below, the number of stick and square group. Someone offered that yes, a possibility but you opted for using a summation notation. We notice when ideas or questions are stated in a group but not discussed. What do you think about that situation and how should deal with it?
Problems

How to assist teachers in evaluating students’ work in chats

Offer assistance to students
- Abstraction tools
- Automatic feedback

Natural Language understanding is very difficult, especially for conversations
The Key Role of Natural Language in CSCL

- **Sfard**: “rather than speaking about ‘acquisition of knowledge,’ many people prefer to view learning as *becoming a participant in a certain discourse*” (2000)
- **Wertch**: Lotman - text is a „thinking device” (1981)
- **Stahl**: “to learn is to become a skilled member of communities of practice .... and to become competent at using their .... speech genres” (2006)
- **Koshmann**: “the voices of others become woven into what we say, write, and think” (1999)
- **Wegerif**: teaching thinking skills by inter-animation: “meaning-making requires the inter-animation of more than one perspective“ (2005)
Dialogism – Mikhail Bakhtin

• “... Any true understanding is dialogic in nature” (Voloshinov-Bakhtin, 1973)
• Real life dialog should be the considered, not only written text (as Saussure recommended)
• Utterances (not sentences) should be the unit of analysis
• Carnival
• Speech genres
  ❖ Inter-animation of voices
  ❖ Polyphony
<table>
<thead>
<tr>
<th>Nr</th>
<th>Ref</th>
<th>Time</th>
<th>User</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td></td>
<td>10.26.25</td>
<td>tim</td>
<td>You discussed about a <strong>topic separation</strong></td>
</tr>
<tr>
<td>18</td>
<td>15</td>
<td>10.26.37</td>
<td>adrian</td>
<td>First of all, the <code>reply</code> method is cumbersome</td>
</tr>
<tr>
<td>19</td>
<td>17</td>
<td>10.26.50</td>
<td>john</td>
<td>yes... because we did not like the way the topics were presented in concert chat</td>
</tr>
<tr>
<td>20</td>
<td>18</td>
<td>10.26.56</td>
<td>john</td>
<td>yes!</td>
</tr>
<tr>
<td>21</td>
<td>20</td>
<td>10.27.04</td>
<td>john</td>
<td>i hate <code>double-clicking</code>!</td>
</tr>
<tr>
<td>22</td>
<td>20</td>
<td>10.27.18</td>
<td>tim</td>
<td>and how can we find topics?</td>
</tr>
<tr>
<td>23</td>
<td>18</td>
<td>10.27.26</td>
<td>adrian</td>
<td>What bothers me is the <strong>linear presentation</strong> of the discussion</td>
</tr>
<tr>
<td>24</td>
<td>23</td>
<td>10.27.43</td>
<td>john</td>
<td>Yep</td>
</tr>
<tr>
<td>25</td>
<td>18</td>
<td>10.27.46</td>
<td>adrian</td>
<td>and <strong>double-clicking</strong> too</td>
</tr>
<tr>
<td>26</td>
<td></td>
<td>10.27.54</td>
<td>tim</td>
<td>You mean u want something like a chat forum? :D</td>
</tr>
<tr>
<td>27</td>
<td>24</td>
<td>10.27.58</td>
<td>john</td>
<td>and the <code>reply-to-facility</code> is supposed to help you</td>
</tr>
<tr>
<td>28</td>
<td>18</td>
<td>10.28.15</td>
<td>adrian</td>
<td>i'd like a <code>tree presentation</code> more</td>
</tr>
<tr>
<td>29</td>
<td>18</td>
<td>10.28.38</td>
<td>adrian</td>
<td>or maybe multiple chat columns, for each chat sub-thread</td>
</tr>
<tr>
<td>30</td>
<td>27</td>
<td>10.28.58</td>
<td>john</td>
<td>but it is really difficult to use in real-time, because there are so many topics discussed which intertwine each other</td>
</tr>
<tr>
<td>31</td>
<td>28</td>
<td>10.29.18</td>
<td>john</td>
<td>i subscribe to a <code>tree-like presentation form</code></td>
</tr>
<tr>
<td>32</td>
<td>P 30</td>
<td>10.29.20</td>
<td>adrian</td>
<td>yes, that's why a clear separation of topics is needed</td>
</tr>
<tr>
<td>33</td>
<td>31</td>
<td>10.29.47</td>
<td>adrian</td>
<td>this is easy to implement, no problem here!</td>
</tr>
<tr>
<td>34</td>
<td>30</td>
<td>10.29.49</td>
<td>tim</td>
<td>You need also a clever visual representation</td>
</tr>
<tr>
<td>35</td>
<td>30</td>
<td>10.30.05</td>
<td>tim</td>
<td>you'll need also a clever visual interface</td>
</tr>
<tr>
<td>36</td>
<td></td>
<td>10.30.22</td>
<td>tim</td>
<td>Who decides the topics?</td>
</tr>
<tr>
<td>37</td>
<td>33</td>
<td>10.30.33</td>
<td>john</td>
<td>i suppose you are referring to the <code>tree representation</code>, right?</td>
</tr>
<tr>
<td>38</td>
<td>37</td>
<td>10.30.45</td>
<td>john</td>
<td>What i would like is a clever way to separate the topics. :)</td>
</tr>
<tr>
<td>39</td>
<td>38</td>
<td>10.30.59</td>
<td>john</td>
<td>not just doing of myself, manually</td>
</tr>
<tr>
<td>40</td>
<td>37</td>
<td>10.31.00</td>
<td>adrian</td>
<td>Yeah</td>
</tr>
<tr>
<td>41</td>
<td>39</td>
<td>10.31.44</td>
<td>adrian</td>
<td>When you start a new thread (a new message, non-related to other message), the app can assume a new topic</td>
</tr>
<tr>
<td>42</td>
<td>39</td>
<td>10.31.46</td>
<td>john</td>
<td>i would like the application to be able to detect w <code>topic change</code> all by itself</td>
</tr>
<tr>
<td>43</td>
<td>42</td>
<td>10.32.01</td>
<td>tim</td>
<td>That right</td>
</tr>
</tbody>
</table>
Explicit vs. implicit links

Explicit links
- ConcertChat

Implicit links
- Linguistic markers
- Inter-animation patterns
  - Adjacency pairs
  - Repetitions
  - Difference-making
Overview

1. Computer-Supported Collaborative Learning
2. Polyphony
3. Automatic chat analysis
4. Results
5. Conclusions
Polyphony
Polyphony and counterpoint

- Concept derived from classical music
  - “These are different voices singing variously on a single theme. This is indeed 'multivoicedness,' exposing the diversity of life and the great complexity of human experience. 'Everything in life is counterpoint, that is, opposition,' “ (Bakhtin, 1984)

- Multiple voices – each utterance contains multiple voices

- Voices inter-animate in an unmerged way:
  - “a plurality of independent and unmerged voices and consciousnesses”
Polyphony

“only the difference between difference and unity as an emphatic difference (and not as a return to unity) can act as the basis of a differential theory (which dialectic merely claims to be) – is the methodical point of departure for the distinction between polyphony and non-polyphony.” (Mahnkopf, 2002).
Polyphony

- A merge of:
  - Melody – longitudinal
  - Harmony – transversal, vertical ("the structure of music with respect to the composition and progression of chords", WordNet; *Dissonance is not excluded, it is very important!*)

- Unity vs. Difference
- Inter-animation of voices – inter-animation patterns
Inter-animation patterns

- Longitudinal
  - Adjacency pairs
  - Repetitions
  - Elaboration
  - Convergence
  - Cumulative talk
  - Repair
- Transversal, differential
  - Dissonance
Other theories

- Discourse analysis (Tannen)
- Conversation analysis (Sacks, Jefferson, Schegloff)
- Accountable talk (Resnick)
- Transactivity (Teasley, Berkowitz & Gibbs, Joshi & Rose)
- Polyphony (Trausan-Matu et al.)
- Inter-animation (Wegerif, Trausan-Matu)
Polyphonic support for inter-animation

- Encourage multiple threads (chat allows them, in contrast to f2f dialog)
- Explicit threading (Wessner)
- Automatic evaluation
- Visualization – diagrams
- Summarization: knowing what came before in clear summaries would help people to respond and carry on the melody
Overview

1. Computer-Supported Collaborative Learning
2. Polyphony
3. **Automatic chat analysis**
   Contributed also Traian Rebedea and Mihai Dascalu
4. Results
5. Conclusions
Identification of Chat Topics

- XML or HTML chat logs
- Tokenization
- Stop-words, emoticons and usual abbreviations (:) , :D , brb, thx, ...) are eliminated
- Misspells are searched using the Google API
- WordNet for identifying synonyms
- Pattern (cue phrases) analysis
## Participants
- azemel
- wmt
- Moderator_Nen
- stefan
- Jsaenni
- gerry
- wees
- murat
- terry
- oprose
- Jim_Waters
dennys

## Topics

<table>
<thead>
<tr>
<th>Index</th>
<th>Topic</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>voices &lt;&lt;&lt; vocalisation, vocalisation, vocalism, phonation, phonation, articulation, spokesperson, interpreter, representative, part, sound, vocalise, vocalise, vocalise</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>voices &lt;&lt;&lt; vocalisation, vocalisation, vocalism, phonation, phonation, articulation, spokesperson, interpreter, representative, part, sound, vocalise, vocalise, vocalise</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>individual &lt;&lt;&lt;</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>role &lt;&lt;&lt; function, office, part, character, theocene, role, persona, purpose, use</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>metaphor &lt;&lt;&lt;</td>
<td>14</td>
</tr>
<tr>
<td>6</td>
<td>text &lt;&lt;&lt;</td>
<td>14</td>
</tr>
</tbody>
</table>

## Utterances Information

<table>
<thead>
<tr>
<th>Index</th>
<th>Text</th>
<th>Voices</th>
<th>Abs value</th>
<th>Miscellaneous</th>
</tr>
</thead>
<tbody>
<tr>
<td>65: 1</td>
<td>Yes, Baldwin's theory is that every text contains a series of voices</td>
<td></td>
<td>0.134</td>
<td>(AGREEMENT)</td>
</tr>
<tr>
<td>66: 1</td>
<td>Baldwin's &lt;&lt;&lt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66: 2</td>
<td>&lt;&lt;&lt; test &lt;&lt;&lt; text, textbook, test, edition, schoolbook, school_text</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66: 3</td>
<td>&lt;&lt;&lt; container &lt;&lt;&lt; incorporates hold, beam, carry, control, hold_inチュ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66: 4</td>
<td>&lt;&lt;&lt; serial &lt;&lt;&lt; serial, serial, publication</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66: 5</td>
<td>&lt;&lt;&lt; voices &lt;&lt;&lt; vocalisation, vocalisation, vocalism, phonation, phonation, articulation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>67: 1</td>
<td>I'm with Nen here. I wonder if this isn't a slightly loose metaphor, In p</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>67: 2</td>
<td>&lt;&lt;&lt; meter, metre, molarity, molar, concentration, thousands, one, thousands</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Graphical Representation of the Conversation

- For each *participant* in the chat, there is a separate horizontal line in the representation.
- Each *utterance* is placed in the line corresponding to the issuer of that utterance, according to the emission time.
  - The *explicit references* among utterances are depicted using blue connecting lines.
  - The *implicit references* (deduced by the system) are represented using other colour (red or green).
- The *strength* of each utterance is represented as a bar chart.
Implicit Links Discovering

Text mining techniques:
- Pattern (cue phrases) analysis
- Co-reference analysis
- Lexical chains
- Heuristics
The method

- List of patterns that consist of a set of words (expressions) and a local subject called the referred word.
- If an utterance matches one of the patterns, we determine what word in the utterance is the referred word (e.g. “I don’t agree with your assessment”).
- We search for this word in a predetermined number of the most recent previous utterances.
- If we can find this word in one of these utterances, then we have discovered an implicit relationship between the two utterances, the current one referring to the identified one.
- During the identification process, the synsets of the words are used.
Utterances’ Strength

The importance of an utterance in a conversation can be computed by computing the amount of useful information.

Another approach (social): an utterance is important if it influences the further evolution of the conversation.

An important utterance – referenced by many further utterances.

Thus, the importance can be considered as a measure of the strength of the utterance.

The utterance is strong if it influences the rest of the conversation (like a breaking news at TV).

Computed recurrently.
The method

- The length and the number of key (important) words.
- The influence on the subsequent evolution of the conversation, considering the explicit and implicit links
- Graph algorithms
- Heuristics
Learners’ Evaluation in Conversation Based on the Polyphonic Model

- Natural Language Processing + Social Network Analysis
- Implicit and explicit reference factors,
- Bonuses for agreement,
- Penalties for disagreement,
- Minimum value for a chat utterance,
- Penalty factors for utterances that agree or disagree with other utterances as these utterances have less originality than the first ones.
Computing the Contributions - NLP

At the start of the conversation, each participant has a null contribution.

For each utterance in the chat, the value of the contributions are modified accordingly:

- The participant that issued the current utterance receives the its score, eventually downgraded, if it is an (dis-)agreement;
- All the participants that are literally present in the current utterance are rewarded with a percentage of its value;
- The participant that issued the utterance referred by the current one is rewarded for an agreement and penalized for a disagreement, with a constant value;
- The participant that issued the utterance referred by the current one and is not a (dis-)agreement is rewarded with a fraction of the value of this utterance;
- If the current utterance has a score of 0, the issuer will receive a minimum score (for participation).
Contributions’ Graphics

- Oy axis – Value of contributions
- Ox axis – The number of the utterance
Evaluation Based on the Social Network Analysis (SNA)

- Total number of characters
- Number of characters / Utterance
- Degree
  - InDegree
  - OutDegree
- Centrality
  - closeness
  - graph
  - betweenness
  - stress
  - eigenvector
- Rank

14 November 2008
OUNL, Heerlen
Participants Evolution

- SNA statistics, plus
- Thread evolution
- Statistical annotation
  - Length
  - Level
  - Length of keywords
  - Type of utterance

\[
mark_{empirc} = \frac{\text{length}}{6} + \frac{\text{MClength} \times 5}{6} \times \left(1 + \frac{1}{10} \times \sum_k \log_2 (\text{no\_occurrences}_k + 1)\right)
\]

\[
mark_{final} = mark_{previous\_utterance} + \text{coefficient} \times mark_{empirc}
\]
Overview

1. Computer-Supported Collaborative Learning
2. Polyphony
3. Automatic chat analysis

4. Results
5. Conclusions
Testing Data (1)

Analysis of 3 collaborative chat conversations in HCI:

- Groups of 4 students
- Chat system: ConcertChat
- Competitive topic: which technology is better for web collaboration: chat, forum, wiki or blog?
- Collaborative topic: means of integrating all the technologies into a single product
Testing Data (2)

2 chat discussions are positive example considering both the content and the collaboration process
- Chat logs 4 and 36

1 chat discussion is a negative example considering the content and, to a lower degree, the collaboration process
- Chat log 34
Testing methodology

Feedback and grading:

- Two evaluators using only the chat logs in HTML format (STM and VP)
- Two evaluators using Polyphony Analysis and the chat logs in HTML format (TR and DM)

Separate grading using ChAMP

Analysis of improvements in grading, feedback, consumed time
## Results – Chat Log 4

### A positive example

<table>
<thead>
<tr>
<th>Chat 4</th>
<th>Iacob Liviu</th>
<th>Andreea Enache</th>
<th>Dragos Diaconu</th>
<th>BARIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>STM</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>VP</td>
<td>10</td>
<td>9</td>
<td>7</td>
<td>6.5</td>
</tr>
<tr>
<td>DM</td>
<td>8</td>
<td>8.5</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>TR</td>
<td>9.5</td>
<td>10</td>
<td>6.5</td>
<td>8</td>
</tr>
<tr>
<td>Average</td>
<td>9.125</td>
<td>8.875</td>
<td>7.125</td>
<td>7.875</td>
</tr>
<tr>
<td>Polyphony</td>
<td>10</td>
<td>8.23</td>
<td>6.50</td>
<td>8.17</td>
</tr>
<tr>
<td>ChAMP</td>
<td>10</td>
<td>5</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>
## Results – Chat Log 36

A positive example

<table>
<thead>
<tr>
<th>Chat 36</th>
<th>Florin</th>
<th>Bogdan</th>
<th>Raluca</th>
<th>Elena</th>
</tr>
</thead>
<tbody>
<tr>
<td>STM</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>VP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM</td>
<td>9.5</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>TR</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>8.5</td>
</tr>
<tr>
<td>Average</td>
<td>8.83</td>
<td>8.66</td>
<td>9.33</td>
<td>8.5</td>
</tr>
<tr>
<td>Polyphony</td>
<td>7.80</td>
<td>9.51</td>
<td>10</td>
<td>8.14</td>
</tr>
<tr>
<td>ChAMP</td>
<td>7</td>
<td>7</td>
<td>10</td>
<td>6</td>
</tr>
</tbody>
</table>
A negative example

Grades marked with * were penalized for plagiarism

<table>
<thead>
<tr>
<th>Chat 34</th>
<th>Delia</th>
<th>Madalin</th>
<th>Cristian</th>
<th>Marian</th>
</tr>
</thead>
<tbody>
<tr>
<td>STM</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>VP</td>
<td>7</td>
<td>7*</td>
<td>4*</td>
<td>6</td>
</tr>
<tr>
<td>DM</td>
<td>8</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>TR</td>
<td>7</td>
<td>6*</td>
<td>5.5*</td>
<td>6</td>
</tr>
<tr>
<td>Average</td>
<td>7.5</td>
<td>6.25*</td>
<td>5.875*</td>
<td>6.25</td>
</tr>
<tr>
<td>Polyphony</td>
<td>6.81</td>
<td>7.97</td>
<td>10</td>
<td>6.47</td>
</tr>
<tr>
<td>ChAMP</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>8</td>
</tr>
</tbody>
</table>
Time Analysis – per Chat

- **STM**: 50 minutes for content, 30 minutes for highlighting each participant, 20 minutes for topics detection
- **VP**: 30-60 minutes
- **TR**: 20-35 minutes
- **DM**: 15-25 minutes

More than 30% improvement for the time needed to analyze each chat
Overview

1. Computer-Supported Collaborative Learning
2. Polyphony
3. Automatic chat analysis
4. Results
5. Conclusions
Conclusions

- The polyphonic model may be used for analysing chats
- Polyphony and ChAMP provide useful information for:
  - a better understanding
  - a faster and more efficient ranking of the participants in a collaborative chat
- Strong points:
  - Reduced time for chat analysis (more than 30%)
  - The automatic grading system looks promising
Conclusions (2)

Weak points:
- Semantics
- Maximum grade in a chat
- Plagiarism

Further work:
- Design a new analysis tool that uses both Polyphony Analysis and ChAMP
- Improvements
- New features
Thank You!

Questions?