Tatiana

Trace Analysis Tool for Interaction Analysts

Gregory Dyke, Ecole des Mines de Saint-Etienne
Kristine Lund, umr ICAR, CNRS, Université de Lyon)
Jean-Jacques Girardot, Ecole des Mines de Saint-Etienne
Contents

1. Presentation of Tatiana
   1. Context - why is analysis difficult?
   2. Artefacts created by researchers
   3. Example of such artefacts in Tatiana
   4. Global overview of the model behind Tatiana

2. Two perspectives on analysis
   1. Social Sciences
   2. Computer Science
Multimedia and Multimodality

How do you record this?
How do you analyse the recording?
Complex data formats

Much better to understand once parsed?

<table>
<thead>
<tr>
<th>User</th>
<th>Message</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cedric</td>
<td>2 characters added near <em>he</em></td>
<td>10:20:40</td>
</tr>
<tr>
<td>Romain</td>
<td>3 characters added near <em>good</em></td>
<td>10:20:43</td>
</tr>
<tr>
<td>Cedric</td>
<td>1 character removed near <em>h</em></td>
<td>10:20:50</td>
</tr>
<tr>
<td>Romain</td>
<td>2 characters added near <em>goodbye</em></td>
<td>10:21:07</td>
</tr>
<tr>
<td>Cedric</td>
<td>2 characters added near <em>hi</em></td>
<td>10:21:22</td>
</tr>
<tr>
<td>Romain</td>
<td>1 characters added near <em>goodbye</em></td>
<td>10:21:42</td>
</tr>
<tr>
<td>Cedric</td>
<td>3 characters added near <em>hi all</em></td>
<td>10:21:51</td>
</tr>
<tr>
<td>Cedric</td>
<td>1 character added near <em>hi all</em></td>
<td>10:21:51</td>
</tr>
</tbody>
</table>

Maybe not!
Quantity of data
Analysis practices

- Gathered from
  - Case studies
  - Literature review (methodology)
  - State of the art (tools)
CoIAT (Avouris et al.)
Replayer (Morrison et al.)
Abstract (Georgeon et al.)
DREW replayer (Corbel et al.)
## Coding (rainbow)

<table>
<thead>
<tr>
<th>Time</th>
<th>User</th>
<th>Message</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>13/11/2001</td>
<td>drew-chat</td>
<td>salut</td>
<td>Interaction management</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>se déconnecte du grapheur</td>
<td>Interaction management</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>alors prête pour 1 h?</td>
<td>Interaction management</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>c'est parti!</td>
<td>Interaction management</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>alors que pense tu des ecm, (super question mais faut bien commencer)</td>
<td>Task management</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>se connecte au grapheur</td>
<td>Task management</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>crée la boîte 1</td>
<td>Argumentation</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>argument de la boîte 1 : ogm que pour l'argent</td>
<td>Argumentation</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>ils augment la pollution</td>
<td>Argumentation</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>supporte l'argument de la boîte 1</td>
<td>Opinions</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>crée la boîte 2</td>
<td>Opinions</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>déplace en 164,251 la boîte 2</td>
<td>Task management</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>je suis d'accord</td>
<td>Opinions</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>on ne va pas beaucoup avancer!!!</td>
<td>Task management</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>argument de la boîte 2 : cause de pollution</td>
<td>Argumentation</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>crée la boîte 3</td>
<td>Opinions</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>supporte l'argument de la boîte 3</td>
<td>Task management</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>déplace en 293,181 la boîte 3</td>
<td>Opinions</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>déplace en 151,251 la boîte 2</td>
<td>Task management</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>déplace en 293,250 la boîte 3</td>
<td>Task management</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>argument de la boîte 3 : détruit couche d'ozone</td>
<td>Argumentation</td>
</tr>
<tr>
<td>13/11/2001</td>
<td>grapher</td>
<td>je suis d'accord aussi!!!</td>
<td>Opinions</td>
</tr>
</tbody>
</table>
Transcription / Annotation
Analysis artefacts and the role they play
Log Data (1)

Writing units (2)

Transcription (3)

Visualisation (4)

Remote control (7)

Highlights are synchronised
Tatiana Replayables can be:
- Annotated / categorised
- Transformed / created by hand
- Transformed / created automatically
- Merged
- Visualised (as table / as graph)
- Synchronised
- Exported

External Replayables (Video/Tool replayer)
- Synchronised
Replayables can undergo…

- **Synchronisation**
  - Provides context
  - Not always enough

- **Transformation**
  - Import / Export
  - Filter / Search
  - Patterns
  - Statistics and indicators
  - Many other generic possibilities
  - Folder for new scripts

- **Visualisation**
  - Table
  - Scoresheet
  - Plugin structure for new ones…

- **Enrichment**
  - Codes / Annotations
  - Relationships
  - Plugin structure for new ones…
To sum up...

- Environment for manipulating *replayables*
- Versatile tool for researchers
  - Many data formats and extension possibilities
- Currently used to analyse many kinds of data
  - Three of our own data sets (collaborative note-taking, collaborative design, children’s explanations)
  - Argumentation (Switzerland)
  - Forums (Hong-Kong)
  - Blogs (Denmark)
  - LEAD project (G.B., Paris, The Netherlands)
  - Boundary object for discussion between epistemologies (series of workshops at ICLS, CSCL, Alpine Rendez-Vous)
  - French traditionnal dancing
Contents

- Computer Science in Tatiana
- Computer Science in Analysis
- Some problems
- Possible solutions
Computer Science in Tatiana

- Model the researcher’s practice
- Model of analysis artefacts which will support this practice
- Implementation of this model
  - Necessary?
  - How to evaluate?
- Most of the implementation is just software engineering
  - Good if this tool is actually supposed to help people
  - Bad for a young researcher
  - Actually a quite hard engineering problem
Who else could have done Tatiana?

- Need to:
  - Understand researcher’s activity
  - Model that activity
  - Produce an implementation
  - Have time to work on one project

- Software companies, some day?
  - Practices are too diverse, no money

- Big european project?
  - Still need someone to implement

- Social sciences researcher?
  - Maybe…
Computer Science in Analysis

- Data Mining
  - Find patterns
- Information Retrieval
  - Ask questions
- Natural language processing
  - Speech processing
  - Topic-extraction
  - Semantics and syntax
- AI
  - Classification
  - ITS
  - Machine learning
- HCI
  - Interfaces
  - Information visualisation
Problems

- What is a model? (what are acceptable answers in different research disciplines)
  - Subdisciplines of computer science tend to have a very narrow view of what appropriate results are
  - Human usage of technology is always Somebody Else’s Problem
- TEL … what about the learning?
- For interdisciplinarity to happen, both disciplines need to find research questions
- Implementations often need several sub-disciplines
- …and a good deal of software engineering
- Defining what *should* be implemented is hard -> most of the time you won’t get what you wanted
Solutions?

- Field of applied computer science
- Phds in pairs
- Pluridisciplinarity
  - Publications in multiple domains
  - Accept that these people lose in depth because of the breadth
  - Find me a job!