Synchronizing Domain Models with Natural Language Specifications

Mathias Landhäußer, Sven J. Körner, Walter F. Tichy
Model Creation from Natural Language

Problem area: Model Creation from Text
Challenge: Handling Modifications (on both sides)
The Whois client makes a text request to the Whois server. [...] 

The Thin client makes a text request to the Whois server. [...] 

[...] The Thin client is a Whois Client and has a Configuration.

initial build of the model

textual change leads to model update

model change leads to text change
Our Approach

Textual Specification

Text Element 1

Text Element 2

Text Element 3

Text Element 4

Text

Model (e.g. UML)

Internal Representation
Our Approach by Example: The Whois Server

The Whois client makes a text request to the Whois server; then the Whois server replies with Text content.

WhoisClient
makes(t: TextRequest, s: WhoisServer)

Whois server
replies(t: TextContent)

TextContent
Text content
The
makes
a
text request
to
the
then
the
Whois server
Phrase
Whois server
Whois client
replies
with
The
Whois server
TextRequest
The Whois Case Study

A `WHOIS_server` listens on TCP_port_43 for `text_requests` from `WHOIS_clients`.

The `WHOIS_client` makes a `text_request` to the `WHOIS_server` then the `WHOIS_server` replies with `text_content`.

All `text_requests` are terminated with `ASCII_CR` and then `ASCII_LF`.

[...]
The Whois Case Study

TCP_port_43

ASCII_CR

ASCII_LF

text_request

WHOIS_client

makes(in: text_request, rec: WHOIS_server)

WHOIS_server

listen(from: WHOIS_client, given: text_request, where: TCP_port_911)
replies(gives: text_content) ...
The WHOIS Case Study

A WHOIS_server listens on TCP_port_43911 for text_requests from WHOIS_clients.

The WHOIS_client makes a text_request to the WHOIS_server then the WHOIS_server replies with text_content.

All text_requests are terminated with ASCII_CR—and then ASCII_LF.

A Thin client is a WHOIS client.

[...]
Evaluating Random Modifications

<table>
<thead>
<tr>
<th>Case Study</th>
<th>Text Changes</th>
<th>Model Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text</td>
<td>Size</td>
<td>Updates</td>
</tr>
<tr>
<td>Cinema</td>
<td>153 Words</td>
<td>1/6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+ 5 irrel.</td>
</tr>
<tr>
<td>Timbered House</td>
<td>88 Words</td>
<td>3/3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WHOIS Protocol</td>
<td>100 Words</td>
<td>8/8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Findings, Limitations & Future Work

- Deletions can lead to text orphans (could be cleaned up automatically)

The is the to the WHOIS client that.
Findings, Limitations & Future Work

- Deletions can lead to text orphans (could be cleaned up automatically)

- Language synthesis is rudimentary (there are solutions for this to be integrated though)

- Current support for class diagrams only: Investigate whether (and how) other diagram types can be covered

- Studies with stakeholders and software architects need to show how sophisticated synchronization of text and models need to be
A WHOIS server listens on TCP port 43 for requests from WHOIS clients. The WHOIS client makes a text request to the WHOIS server, then the WHOIS server replies with text content. All requests are terminated with ASCII CR and then ASCII LF. The response might contain more than one line of text, so the presence of ASCII CR or ASCII LF characters does not indicate the end of the response. The WHOIS server closes the connection as soon as the output is finished. The closed connection is then automatically re-established.

```prolog
[ #! WHOIS_server(ASG(),NRECP()) Interno(ACT(),TCP_port_43|PAT|) for requests|HAB|] #to FROM

[ ] #The @WHOIS_client(ASG(nm),NRECP(nm)) makes|ACT(|#a text_request|HAB|) #to = |

[ ] #The WHOIS server(NRECP(nm)) |SUM| then |the |WHOIS_server|SUM| |

[ ] #The |WHOIS_server|SUM| |

[ ] #All @requests|PAT| #are terminated|ACT |with |

[ ] #ASCII CR and then ASCII LF |INSTEAD | |

[ ] #The |response|OMN| |

[ ] #might |contain|ACT|C| |P| |

[ ] #more_than_one |

[ ] #The |presence|ACT |of |ASCII_CR |OR |ASCII_LF |PAT | |characters |INSTEAD | |

[ ] #does_not_indicate|ACT|

[ ] #the_connection|HAB| |PAT| |as |soon |as | |

[ ] #the |subsequent|ASCII_CR |are |transformed|ASCII_LF| |

[ ] #the_connection|HAB| |PAT|

[ ] #the |connection|HAB| |PAT| |

[ ] #the |connection|HAB| |PAT| |
```

Landhäuser, Körner, Tichy | Synchronizing Domain Models with Natural Language Specifications
RAISE’12 – Workshop on Realizing Artificial Intelligence Synergies in Software Engineering
THANK YOU!
QUESTIONS?