

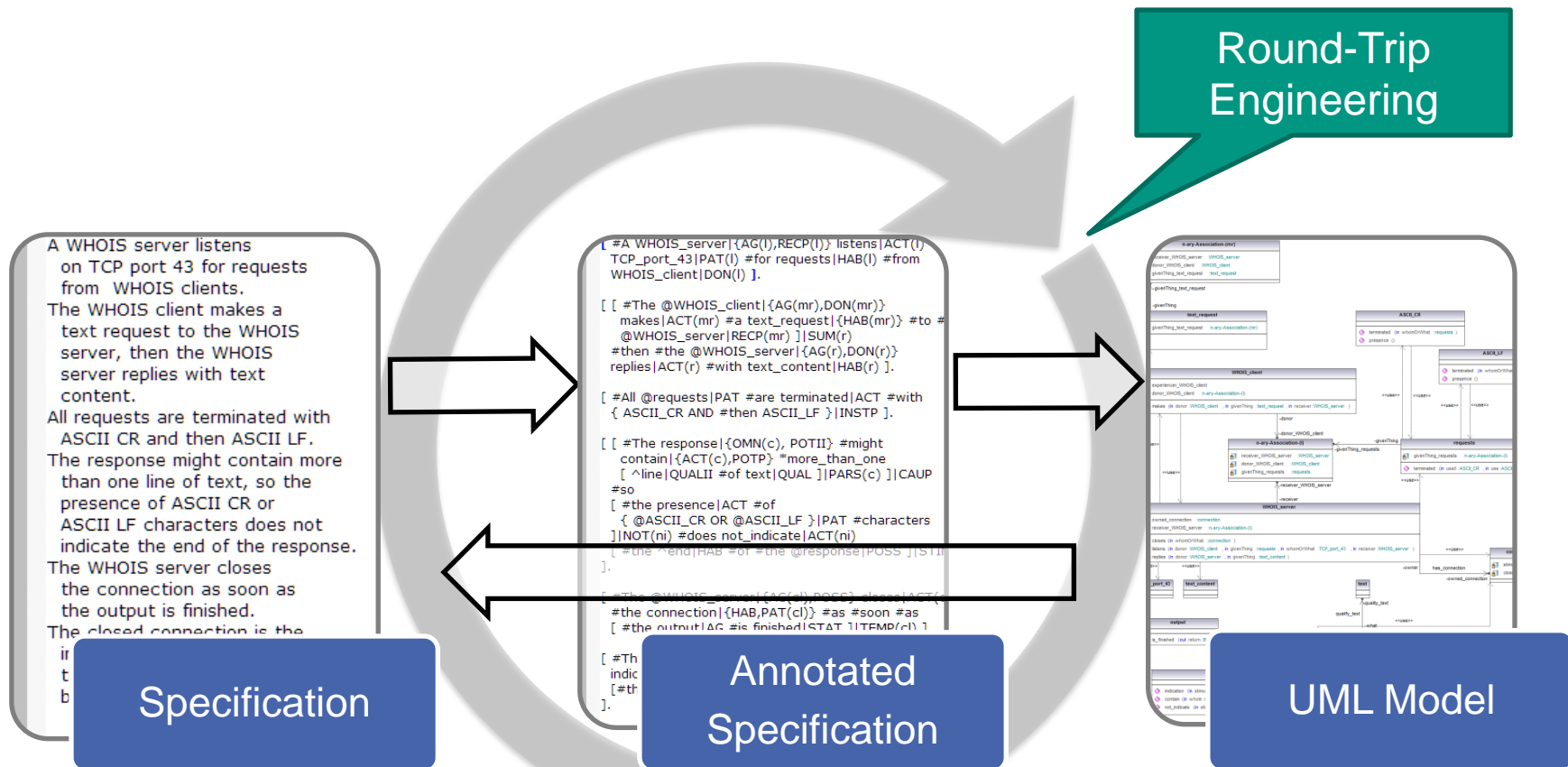
Synchronizing Domain Models with Natural Language Specifications

Mathias Landhäußer, Sven J. Körner, Walter F. Tichy

IPD Tichy, Department of Informatics



Model Creation from Natural Language



- Problem area: Model Creation from Text
- Challenge: Handling Modifications (on both sides)

The Challenges

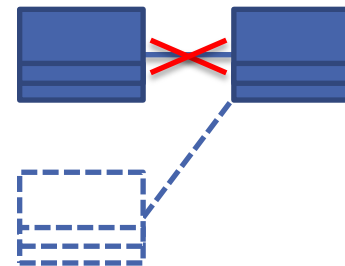
The whois client makes a text request to the whois server.
[...]

initial build of the model →



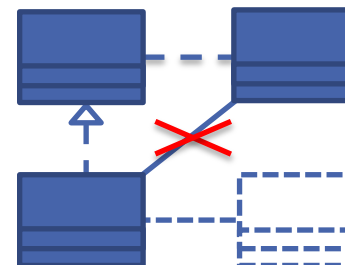
The **Thin client** makes a text request to the whois server.
[...]

textual change leads to model update →

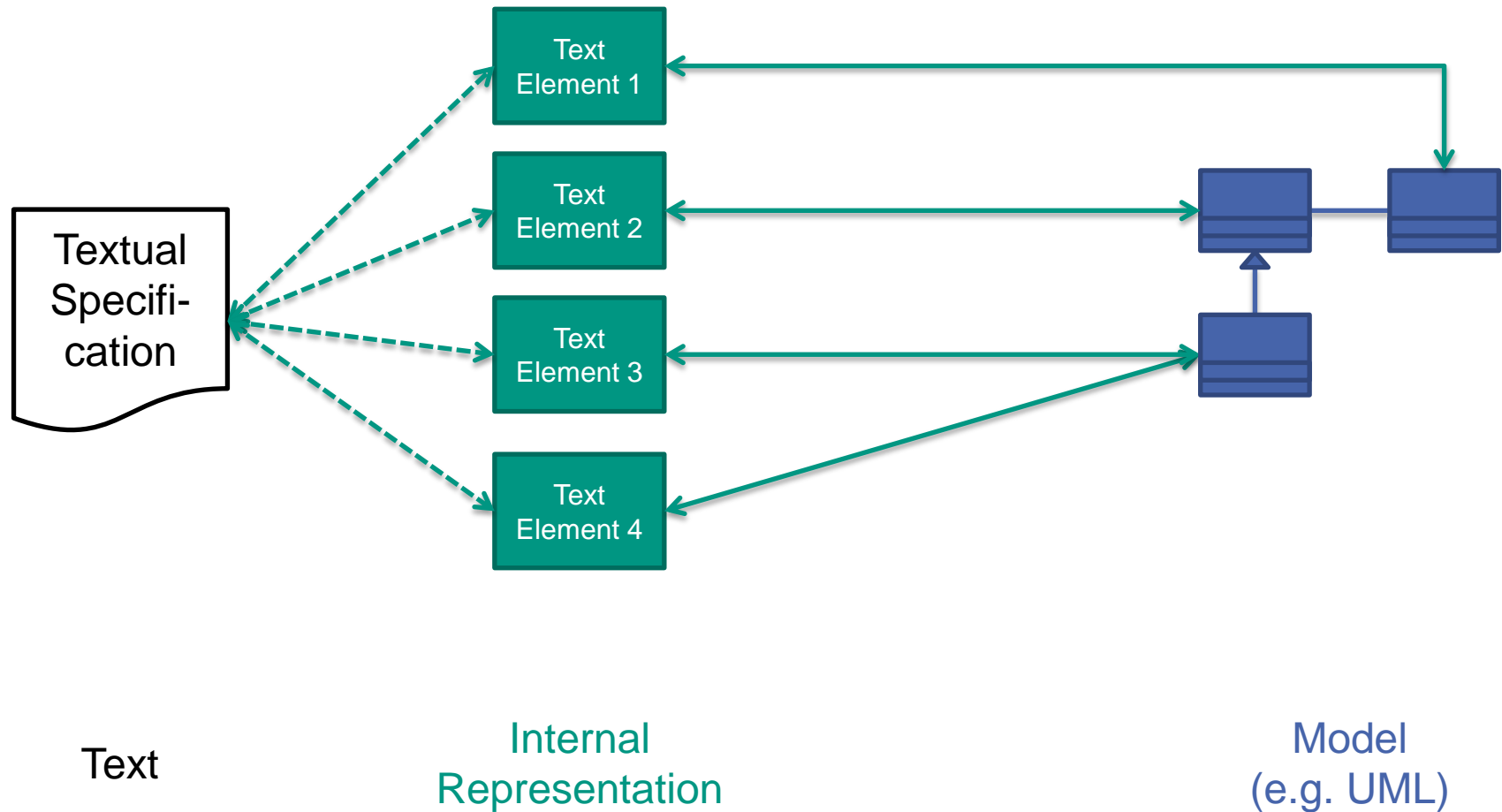


[...]
The **Thin client** is a **whois Client** and has a **configuration**.

← model change leads to text change

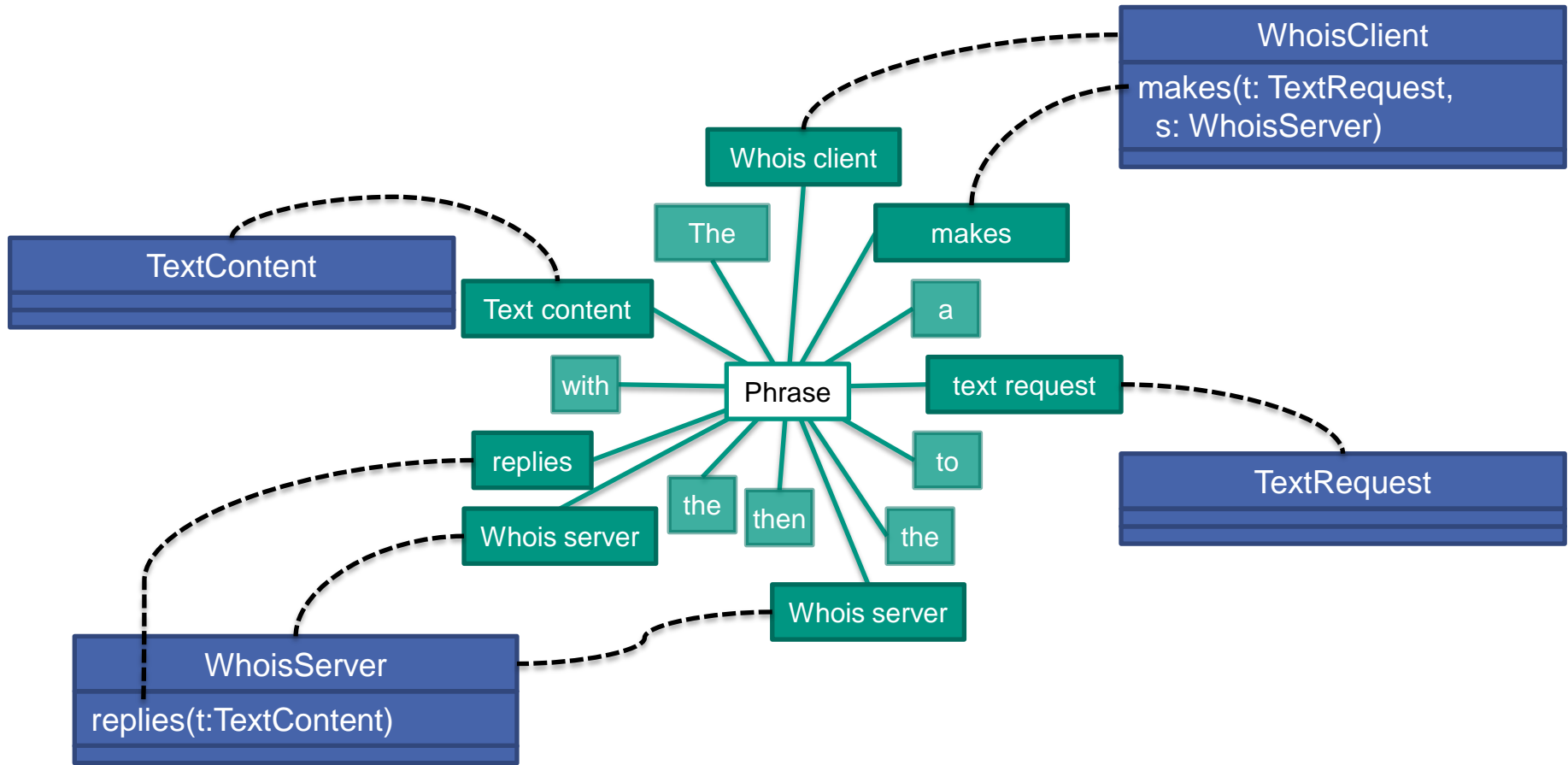


Our Approach



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.



The Whois Case Study

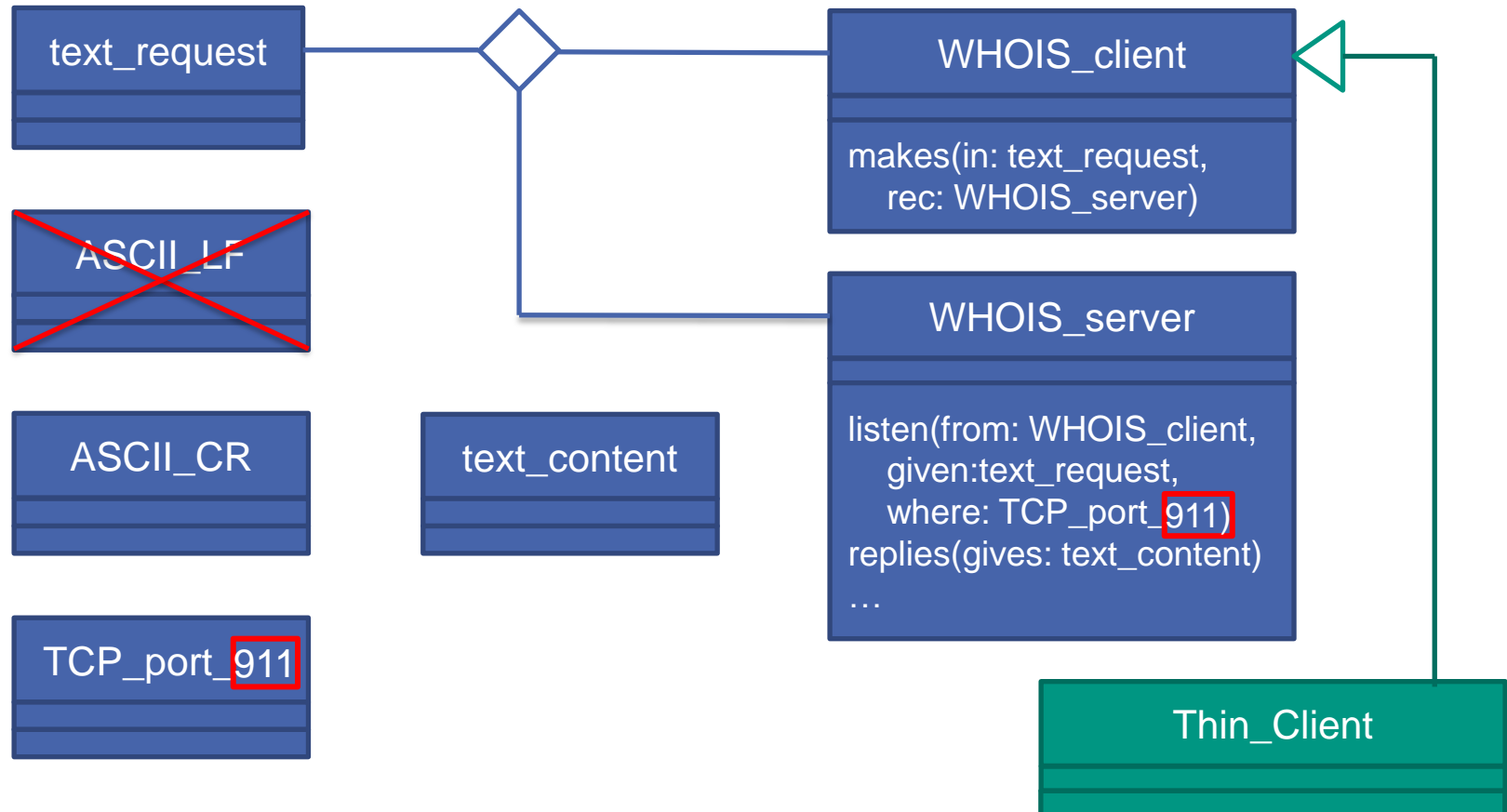
A `WHOIS_server` listens on `TCP_port_43` for `text_requests` from `WHOIS_clients`.

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



The Whois Case Study



whois_mod.txt - Microsoft Word

File Home Insert Page Layout References Mailings Review View

```

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.

[...]
  
```

Words: 35 English (U.S.) 170%

Evaluating Random Modifications

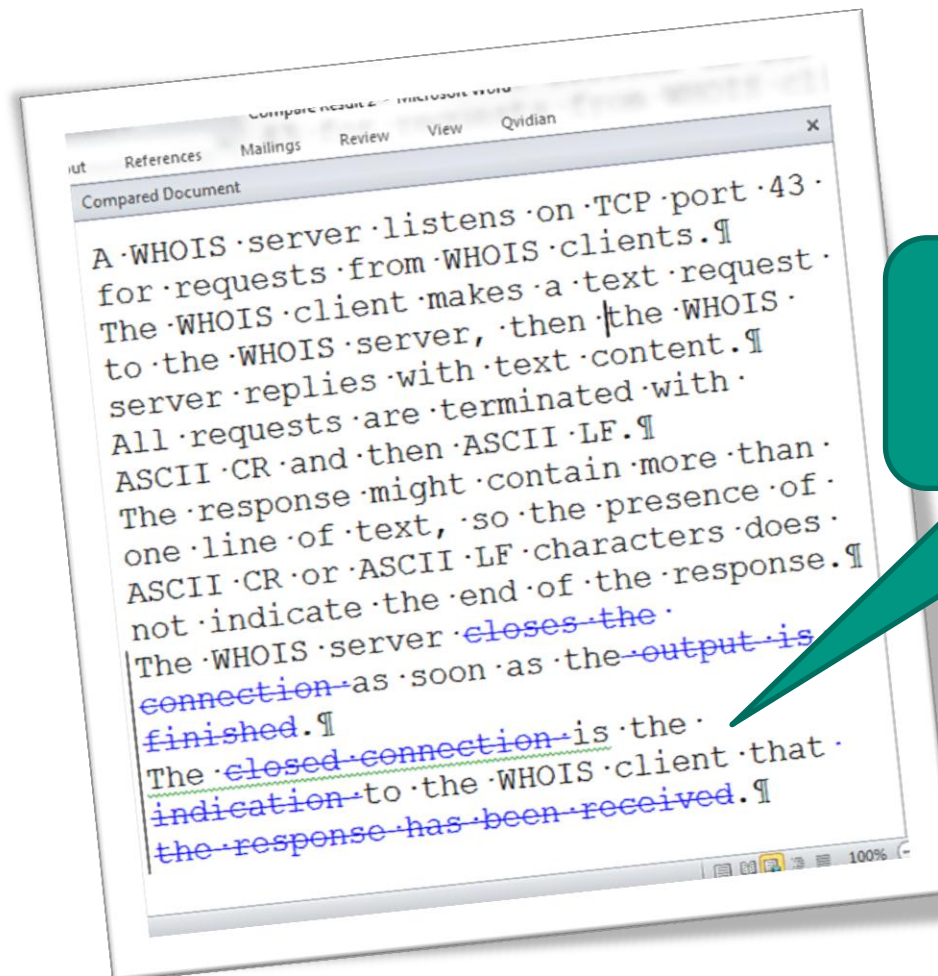
Case Study		Text Changes		Model Changes	
Text	Size	Updates	Deletions	Updates	Deletions
Cinema	153 Words	1/6 + 5 irrel.	6/7 + 1 irrel.	4/4	7/7
Timbered House	88 Words	3/3	2/7 + 5 irrel.	3/3	7/7
WHOIS Protocol	100 Words	8/8	2/2	5/6	2/3
				+ 1 incorr. crea/del	

Irrelevant changes: Text parts that are not modeled.

Incorrectly detected model change: Change masked by deletion.

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

Round-Trip Engineering

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 the

Specification

```

1 #A WHOIS_server|{AG(l),RECP(l)} listens|ACT(l)
   TCP_port_43|PAT(l) #for requests|HAB(l) #from
   WHOIS_client|DON(l) ].

2 [ [ #The @WHOIS_client|{AG(mr),DON(mr)}
   makes|ACT(mr) #a text_request|{HAB(mr)} #to #
   @WHOIS_server|RECP(mr) ] ]SUM(r)
   #then #the @WHOIS_server|{AG(r),DON(r)}
   replies|ACT(r) #with text_content|HAB(r) ].

3 [ #All @requests|PAT #are terminated|ACT #with
   { ASCII_CR AND #then ASCII_LF } ]INSTP ].

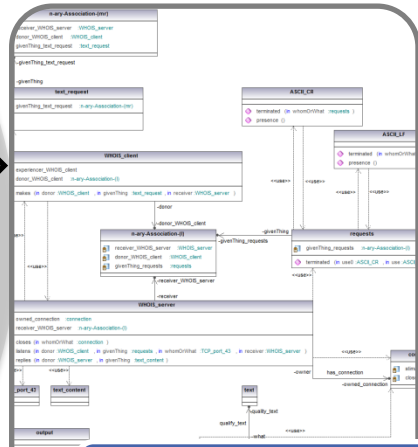
4 [ [ #The response|{OMN(c), POTII} #might
   contain|ACT(c),POTP} *more_than_one
   [ ^line|QUALII #of text|QUAL ] ]PARS(c) ] ]CAUP
   #so
   [ #the presence|ACT #of
   { @ASCII_CR OR @ASCII_LF } |PAT #characters
   ] ]NOT(ni) #does not_indicate|ACT(ni)
   [ #the ^end|HAB #of #the @response|POSS ] ]STII
   ].

5 [ #The @WHOIS_server|{AG(mr),RECP} #closes|ACT(
   #the connection|{HAB,PAT(c)}] #as #soon #as
   [ #the output|AG #is finished|STAT ] ]TEMP(c) ].

6 [ #Th
   indic
   [#th
   ].

```

Annotated Specification



UML Model

THANK YOU!
QUESTIONS?