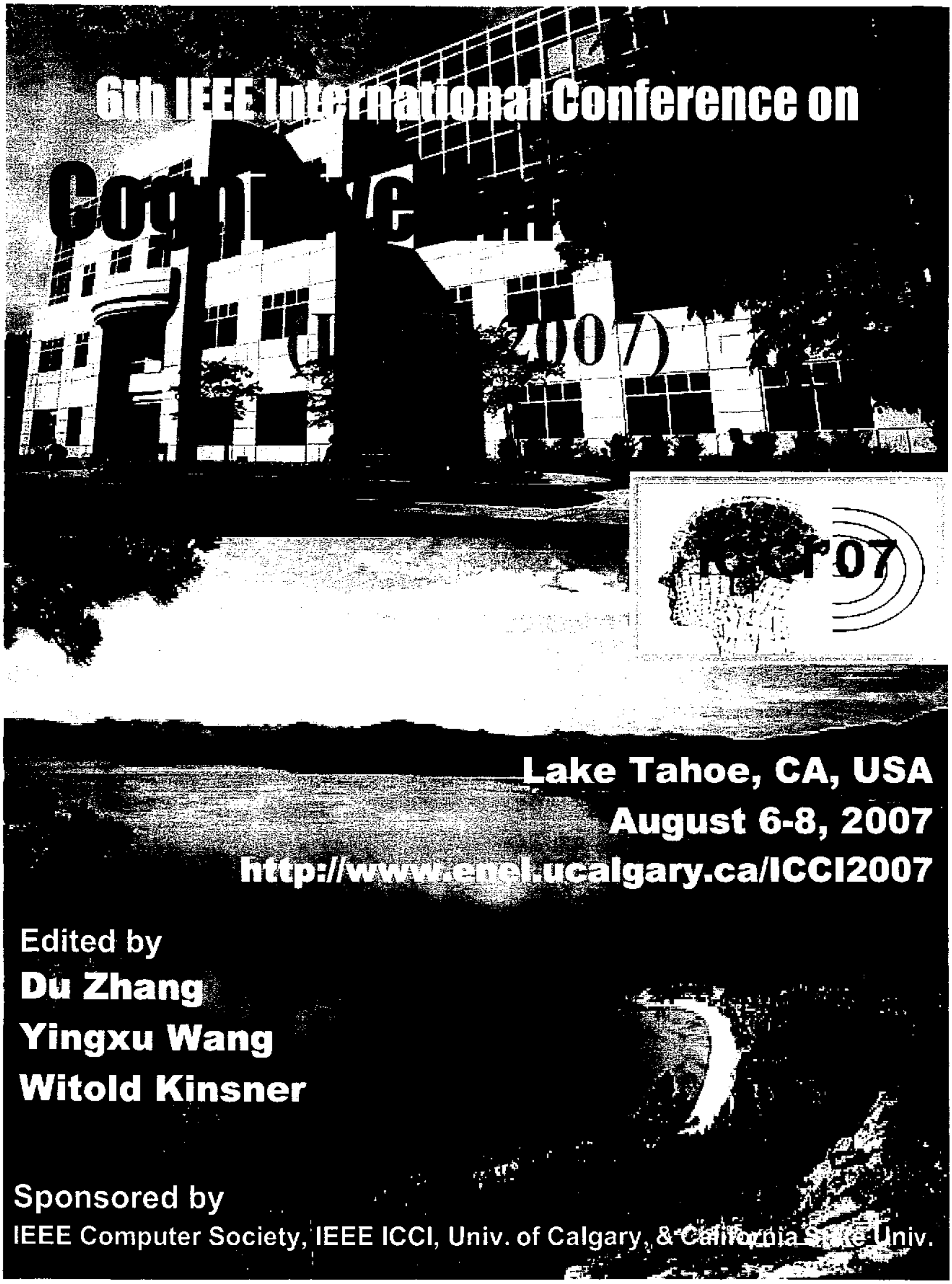

6th IEEE International Conference on
Lake Tahoe, CA, USA
August 6-8, 2007
http://www.enel.ucalgary.ca/ICCI2007
Edited by
Du Zhang
Yingxu Wang
Witold Kinsner
Sponsored by
IEEE Computer Society, IEEE ICCI, Univ. of Calgary, & California State Univ.

Proceedings of the Sixth IEEE International Conference on

Cognitive Informatics

ICCI 2007

August 6 – 8, 2007, Lake Tahoe, CA, USA

Edited by

Du Zhang, Yingxu Wang, and Witold Kinsner

Supported by

IEEE COMPUTER SOCIETY

IEEE ICCI   IJCINI

UNIVERSITY OF CALGARY

IEEE Sacramento Valley Section - Region 6

IEEE Canada

IEEE COMPUTER SOCIETY

Los Alamitos, California
Washington • Brussels • Tokyo

IEEE Computer Society Order Number **07EX1833**
ISBN **1-4244-1327-3**
Library of Congress **2007927978**

Additional copies may be ordered from:

IEEE Computer Society
Customer Service Center
10662 Los Vaqueros Circle
P.O. Box 3014
Los Alamitos, CA 90720-1314
Tel: + 1 800 272 6657
Fax: + 1 714 821 4641
http://computer.org/cspress
csbooks@computer.org

IEEE Service Center
445 Hoes Lane
P.O. Box 1331
Piscataway, NJ 08855-1331
Tel: + 1 732 981 0060
Fax: + 1 732 981 9667
http://shop.ieee.org/store/
customer-service@ieee.org

IEEE Computer Society
Asia/Pacific Office
Watanabe Bldg., 1-4-2
Minami-Aoyama
Minato-ku, Tokyo 107-0062
JAPAN
Tel: + 81 3 3408 3118
Fax: + 81 3 3408 3553
tokyo.ofc@computer.org

Individual paper REPRINTS may be ordered at: reprints@computer.org

Editorial production by Yingxu Wang and Du Zhang

Cover art production by Yingxu Wang

Printed in USA

IEEE COMPUTER SOCIETY

IEEE

Table of Contents

ICCI 2007

Using Verbal Protocols to Assess the Influence of Import-Coupling on the Comprehensibility of OCL Expressions

Luis Reynoso
Department of Computer Science
University of Comahue, Argentina
Buenos Aires 1400 (8300) Neuquén
lreynoso@uncoma.edu.ar

Marcela Genero, Mario Piattini
Department of Computer Science
University of Castilla-La Mancha, Spain
Paseo de la Univ. 4 (13071) Ciudad Real
{Marcela.Genero}, {Mario.Piattini}@uclm.es

Abstract

The import coupling of an Object Constraint Language (OCL) expression increases when the expression operates within a large context of a Unified Modelling Language (UML) model. The expression would thus exhibit a decrease in its comprehensibility, and in consequence, the maintainability of the OCL expressions could be affected, due to the fact that modelers need to know all the artifacts upon which the expression relies. In order to explain the cognitive complexity of modelers dealing with OCL expressions with high import-coupling we have carried out a think aloud experiment which is described in detail in this paper. The experiment is part of a project which we are developing with the goal of obtaining valid measures through which to assess the influence of import coupling on OCL expression maintainability. The empirical results show evidence that: 1) reified objects, problem objects and their relationships constitute the main components of the cognitive complexity of modelers; 2) modelers use different cognitive approaches to comprehend an OCL expression within a UML/OCL model.

1. Introduction

Numerous attributes of software systems are highly dependent upon the comprehension and measurement of software structural properties [1], [21]. For this reason, the comprehension of structural properties has to be empirically studied by using cognitive informatics and theoretical software engineering methodologies [21],[6] as a base.

We believe that the comprehension of OCL expressions at early stages of software development is essential in order to understand the whole UML/OCL model to which the expressions are attached [15]. Furthermore, the increasing importance of OCL expressions in model-centric methodologies makes the comprehension of their major elements to be carefully assessed since the simplicity and intuitiveness of the OCL language can not guarantee the comprehensiveness of all written expressions.

It has been empirically proved that an important factor in a decreased software comprehension [2] is a high coupling of software artifacts. One concern that should be present in all modeling tasks [14] as a desirable property of quality is that of delivering a low coupling system. Coupling is, in fact, considered to be one of the most complex software attributes in object oriented systems [1]. Within an OCL expression, there is a concept (navigation) which is at the core of OCL language which defines coupling between the contextual instance and other objects within a model [22] and influences the OCL expression comprehension. Moreover, coupling within an OCL expression presents a special characteristic: the inner definition of OCL as a textual add-on to UML models allows us, in one expression, to refer to UML artifacts and not the other way round, so, the coupling' locus of impact of OCL expression is primary of import coupling. Import coupling [2] focuses on the client entity in a client-supplier relationship defining a coupling connection. Within this context, modelers should be aware of the fact that OCL expressions with high import coupling operate in a large context and in order to comprehend the meaning of the expression modelers need to know all the UML artifacts on which the OCL expressions rely.

Our hypothesis is that a high import coupling within OCL expressions may influence the cognitive complex-

Proc. 6th IEEE Int. Conf. on Cognitive Informatics (ICCI'07)
D. Zhang, Y. Wang, and W. Kinsner (Eds.)
1-4244-1327-3/07/$25.00 ©2007 IEEE

ity[1] of modelers, and high cognitive complexity leads OCL expressions to exhibit undesirable external qualities, such as a lower comprehensibility or a reduced modifiability. In order to assess the influence of import-coupling on the maintainability (comprehensibility and modifiability) of OCL expressions we defined a set of measures in [17]. However we now need to provide a plausible explanation of these measures from a cognitive psychology point of view, such as the understanding of the cognitive demands that software places on modelers [9], otherwise we shall only surface features of the software measured. In fact the consideration of multiple theoretical perspectives, including human cognition, provides a solid foundation upon which to derive an integrated model which relates the internal and external attributes of software quality [7].

Moreover, human cognition becomes more relevant if we take into consideration that in the last few years software engineering empiricists have begun to address the human role in software development in a serious way [19]. Seaman argues [19] that in order to delve into the complexity of the human role in software engineering rather than to abstract it away, qualitative methods should be used. Therefore, in this paper we describe a think aloud experiment in order to evaluate whether reified objects (which are not problem domain objects per se), problem objects and their relationships are the main constituents of the cognitive complexity of modelers dealing with OCL expressions. In this experiment we also evaluate which are the different approaches that modelers take in order to comprehend an expression within a UML/OCL model. The structure of the paper is as follows: the next section briefly explains related work, section 3 presents an overview of a qualitative method: verbal protocol analysis and section 4 describes the experiment. Finally, in section 5 conclusions are presented and future work is outlined.

2. Related Work

As is previously mentioned, in [17] we have defined a set of measures to be used in assessing the influence of import coupling on the maintainability of OCL expressions. The definition of these measures was achieved by means of a method proposed in [3] which mainly consists of the following steps: measure definition, theoretical validation and empirical validation. However, we believe that in order to obtain valid measures we must understand the cognitive complexity of the modelers dealing with these measured artifacts. We have, therefore, applied a cognitive theory-based approach to give a more precise definition of a measure's rationale in relation to the modelers' mental load when they deal with OCL expressions [16]. We have based our reasoning about the comprehension of OCL expressions on two main topics: *mental models* and *cognitive models*. The former concept describes a subject's mental representation of the software artifact to be understood whereas a cognitive model describes the cognitive processes and temporary information structures in the subject's head which are used to form the mental model. Chunking and tracing, two cognitive techniques defined in the Cant el al.' cognitive model [4], were considered as a basis in the definition of the measures [17]. In addition, in [16] we defined a mapping between the concepts measured by the measures and the components of a mental model in order to explain the results obtained in a family of experiments. Nevertheless, the empirical validation carried out in [16] and [18] are based on quantitative methods, and it could be argued that human cognition is one of the few phenomena that is complex enough to require qualitative methods to study it [19]. Therefore, in this paper we decided to apply qualitative methods in the experimentation.

3. Overview of Verbal Protocol Analysis

The underlying principle of verbal protocol analysis, a qualitative method, is that any verbalization produced by a subject whilst solving a problem - known as concurrent 'think aloud' - will directly represent the contents of the subject's working memory [8]. Using this technique, a researcher can obtain an insight into the subject's cognitive process and use this to address a research question. For example, to investigate a subject's understanding of the problem space [10] or how the current state of the problem solution is evaluated, etc. Verbal protocol analysis is a simple technique to apply, and requires few special arrangements: The subject is asked to verbalize his/her thoughts concurrently at work on a task, recorded verbalizations are transcribed into protocols, the protocols are coded using a preestablished coding scheme, and then analysed in accordance with the relevant research question [10].

Verbal protocol analysis has been applied to a range of problem types, and the data obtained has been used to create, confirm, or refute hypotheses across a wide range of research domains. A thorough review of the literature used to appraise the application of the verbal protocols analysis technique to software engineering can be obtained in [10].

We shall now briefly summarize the stages of a verbal protocol, Figure 1 depicts the principal data used for the entire technique:

[1] Defined as the mental burden of a person dealing with a software artifact

1. **Preparing and Running the session:** We shall now describe the practical procedures which must be applied in experiments where subjects are asked to think aloud:

   (a) **Setting:** The experimental setting should be such that the subject feels comfortable. Some particularly important recommendations include the following: the room should be quiet, a glass of water should be at hand, it is important to explain that the data obtained will be handled in strict confidentiality [20], etc.

   (b) **Instructions:** The instruction related to thinking aloud should be quite simple. The essence of the instruction is: 'Perform the task and say out loud what comes into your mind'. Instructions should not be too long to avoid the situation of the subjects to make up their own interpretations about what is requested from them.

   (c) **Warming up:** Most subjects need a little training before the real experiment starts. It is important to give the subject an opportunity to practice thinking aloud. In general it is wise to look for a task which is not too different from the target task.

   (d) **Behaviour of the experimenter and prompting:** When the subject is working on the task, the role of the experimenter is a restrained one. Interference should only occur when it may be necessary to give the subject a short reminder to continue verbalizing if there is a period of silence. Then the experimenter should prompt the subject by simply saying: 'Please, keep on talking'.

   (e) **Recording:** Sessions are usually recorded on audio- and video-tape. It may be wise to include the instruction and practising phase, in order to be able to check afterward whether the procedure was performed correctly.

2. **Transcribing the Protocol:** The actions of transcribing the verbalizations recorded during the session (depicted as Raw Protocols in Figure 2) are performed without any difficulty. However this task is tedious and time-consuming, because transcribing a protocol usually means typing it out in a manner which is as verbatim as possible. In psychological research, in principle everything may be relevant and therefore basically everything should be typed out [20].

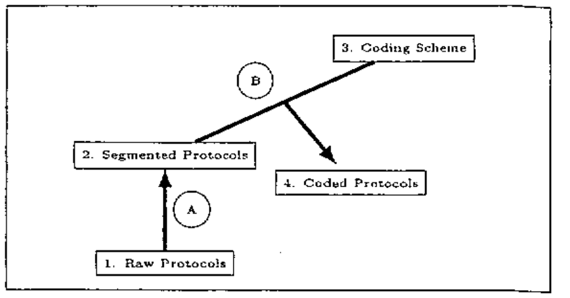


Figure 1. Data Transformation in Verbal Protocols

3. **Segmenting the Protocol:** The protocol is segmented, that is divided into segments (Figure 2, Activity A), each of which is given an identification number. Although there is no standard definition of a segment, most segmentation schemes are broadly similar, being based upon boundaries of phrases marked by pauses, an identifiable single unit such as a reference or assertion, etc. [20]. Segments are often combined into episodes, which are sequence of segments that corresponds to a single element in a cognitive model. However if the task analysis and the psychological model are very detailed then certain elements of the model may correspond directly to segments in the protocol.

4. **Constructing a Coding Scheme:** The coding scheme specifies how the elements of a cognitive model can be identified in the data and is based on our knowledge of the way in which cognitive process will be verbalized. Codes must be developed to correspond to a formalism which will be used to represent the knowledge [5]. This scheme is useful when comparing the protocol and the psychological model. Every process or component in the psychological or cognitive model is stated as how we expect that these processes or components will appear in the protocols.

   There are some categories in the coding scheme which are not directly derived from the model. These are the verbalizations which are not covered by the model, but may still be anticipated in the protocols; for instance: silent periods, actions, comments about oneself, etc. Sometimes the content of these categories in task performance is not relevant, but the moment at which they occur is.

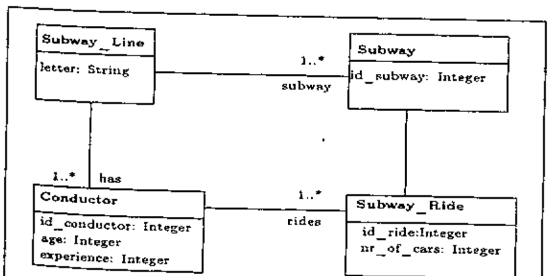
   The result of applying the coding scheme to a segmented protocol is a coded protocol (Figure 2, Activity B).

5. **Verifying intercoder reliability:** Whatever the origin of the coding scheme is, there should be confidence in the reliability of a scheme's application to a set of protocols. The consistency and reliability should be assessed, typically by means of a second encoder. The original encoding can be compared with that of the second encoder in order to obtain either the percentage agreement or the coefficient of agreement (Cohen's K). The accepted percentage figures are typically above 75% or above 0.8 K. The Kappa makes a correction for the correspondence that can be expected from the marginal frequencies. This is a conservative estimate of intercoder reliability [20].

6. **Analysing the code patterns:** The encoded protocol is analysed in order to answer the research questions. The type of analyses used typically reflect the type of research questions being asked. Examples include: enumeration of specific categories, relative distribution of the various activity categories identified, analysis of the pattern of activities, for example switching between two types of activity and even verbalization rates.

4. A Think Aloud Experiment

We have carried out a think aloud experiment through which to test a categorical model of the cognitive process of modelers dealing with OCL expressions. The categories concerned whether the utterances obtained by the students pertained to different coupling aspects: problem objects, relation between problem objects and reified objects (OCL collections of objects such as sets, bags, etc are reified objects, they are not problem object per se but they bundle problem objects together).



Context Subway_Line
inv: self.has->select(c | c.age > 50)->isEmpty()

Figure 2. Sample of an OCL expression

As we have previously explained the think aloud method involves the analysis of recorded verbal protocols that resulted from asking subjects to voice their thoughts when executing particular problem-solving tasks. Within the experiment the tasks we asked the subject to perform were to comprehend an OCL expression, and to express the meaning of the OCL expression. A expression consists of suitable short assertions that are not always easy to understand, specially when a lot of objects are coupled.

We carried out a think aloud experiment in order to test the following research questions:

- Do the modelers apply tracing and chunking cognitive techniques when comprehending an OCL expression?
- Do the subjects take a broad comprehension of the class diagram before comprehending an OCL expression?
- Are reified objects, the relationships between objects and problem objects the most important categories of the model of the subject' cognitive process when they deal with an OCL expression comprehension?

We gave a training course about OCL language to ten advanced students. We divided the participants into three groups according to their experience, and then three subjects (one from each group) were chosen to participate in a think aloud experiment. Their utterances were taped and video-recorded. The use of a relatively small number of subjects is typical of studies that collect and analyze verbal protocol data [11]. Think aloud protocols were collected while each subject comprehended each of the three tests. The three tests had different coupling complexity. The first test was the easiest test. The difficulty of the second test was due to the use of different rolenames whereas the difficulty of the third was due to the use of collection operations. Nevertheless, we believed that the last test would be more difficult to solve than the second test.

In order to set the experiment appropriately, we took several factors into account such as: we chose a comfortable and quiet room, we provided the subject with a glass of water, etc. In order to start the experiment and to give the subject an opportunity to practice thinking aloud, a warm up session took place before the real experiment.

Participants' verbalizations were transcribed and nine protocols were obtained, three protocols from each subject. A researcher checked each transcription against the video tape in order to make any necessary corrections or to adjust phrasing. The raw protocols were then broken up into segments which represented the subject' utterances. Once the verbal protocols were segmented, they were ready to be coded.

These coding categories consisted of whether the explanations and utterances used by the subjects pertained to reified objects (RO), the Relationship between problem objects (RBPO), the Problem Objects (PO), etc. Table 1 shows the coding categories and a description of each one.

The RO category was coded as being those utterances which were concerned with the collection of objects, the collection operations and the explicit iterator of collections. Within the RBPO category we included related problem objects such as the understanding of a relationship within the class diagram or the use of OCL navigations. Within the PO category concepts such as the contextual type, the contextual objects of the expression (or even their attributes) were included.

The SP category was used when the subjects' explanations were related to a statement of the problem whereas the TP category was related to how the subjects planned to perform the task. The TE category was used when the subject explained an OCL expression according to its type (the expressions included in the experiment were invariant expressions). The CWE category was used when the subject chunked the whole OCL expression and described its meaning. We also used special coding categories (SCC categories) to code some utterance that were not covered by the model, as we have described in the previous section (see the last five entries of Table 1).

The number of utterances ranged from 20 to 46. Table 2 contains a coded protocol from a subject to illustrate the coding. The utterances belong to a warm up session using the expression and the class diagram of Figure 2. We did not aggregate the protocol segments into episodes, because the grain size of segments is suitable for testing the hypotheses. We coded all the segments according to the coding categories of Table 1 but we were only interested in some specific categories (RO, RBPO, PO) that were related to coupling concepts. In these cases it would be more efficient to encode protocols directly from the audio-tape (i.e. direct encoding) instead of transcribing and coding the transcription. However, this is often not a method to be recommended [20].

The rest of the categories were used to describe a general process of comprehension used by the subjects. Special coding categories were also included because these remarks might be an indication of the level of difficulty of a sub-task.

In order to control the coding reliability, each verbal protocol was coded by two independent coders. Table 3 shows the Coehn' K across the nine protocols. As we explained in the previous section the values of Coehn' K are acceptable due to the fact that they are greater than 0.8 K.

4.1. Analysis

The encoded protocols were analysed to answer the research questions. We performed an analysis of each subject performing the tasks. For each task of each subject we depicted the coded protocols. Two types of figures were obtained to facilitate the analysis. First, the sequence of coded utterances was represented in a graph which summarized each subject's protocol. In the graph, each dot represents a subject's utterance and the coding categories assigned to that utterance. Second, the number of related coded protocols for each protocol were depicted. We shall now describe the analysis of each subject. General conclusions are described in the following subsection.

Analysis of Subject 1. The number of utterances of each protocol obtained from subject 1 were: 22 (protocol 1), 20 (protocol 2), 33 (protocol 3). The time (in minutes and seconds) spent on each protocol was: 02:16, 02:24, 5:36. In Figure 3 (a graph of the subject's protocols) we can observe that the subject does not attempt to understand the class diagram before starting to comprehend the OCL expression. The technique applied by subject 1 is the same in the three protocols: after reading the statement of the problem (SP) he analyzes the context of the expression and starts to understand its meaning. Nevertheless, he uses the class diagram only to focus on those relationships and classes (and attributes) of the class diagrams when they appear within the OCL expression. We believe that the subject employs an as-needed strategy of the class diagram. Therefore, the understanding of the expression and of the class diagrams is the product of an intertwining activity.

The OCL expression is comprehended from left to right, and the subject starts to capture the meaning by obtaining different assertions and utterance of the meaning, until the subject chunks the whole expression. After the meaning is obtained the subject reviews the expression to be sure of his answer.

In several segments of protocol it is clear that the subject uses many expression such as 'now we are in the X class', 'now we are going to the Y class', etc. to reveal the object oriented perspective taken by the subject in understanding the expression. This cognitive perspective is similar to the identification of deictic words such as *here* and *there* in the work of Hutchis

Table 1. Coded Categories

Coding	Description
RO-CO	Reified Object: collection operations, etc.
RO-C	Reified Object: collection of objects
RO-EI	Reified Object: iterator variable within a collection operation.
RO-LOO	Looking for the definition of an operation collection in a list of operations
RBPO-CD	Relation between problem object: after reading the class diagram
RBPO-NR	Relation between problem object: Navigated Relationship
PO-CD	Problem Object(s) after reading the class diagram
PO-CC	Problem Object(s) a concept of a class from the OCL expression
PO-ACO	Problem Object(s): attribute belonging to contextual objects
PO-CA	Context Analysis and evaluation
EO	Elementary Operation
TE	Type of Expression
CWE	Chunking the whole OCL expression
TP	Task Planning
SP	Statement of Problem
SCC-NT	Special coding categories: Talking about not task related issues
SCC-EM	Special coding categories: Evaluation of the task or task-situation at a meta-level
SCC-CO	Special coding categories: Comments on oneself
SCC-SP	Special coding categories: Silent periods.
SCC-A	Special coding categories: Actions

Table 2. Example of a verbal protocol from a Pilot Subject

Id	Code	Utterance
1	SP	... what I'm going to do first is take a look at the class diagram
2	PO-CD	Subway Line ...
3	PO-ACO	I guess the identification would be to say which is the line
4	RBPO-CD	Associated with each subway line there are many subway train, one or more
5	PO-ACO	Yes, it is correct, the diagram show the year in which the train was bought...
6	PO-ACO	An identification for the subway train and another for the motor...
7	RBPO-CD	A subway train operates on more than one subway line ....
8	RBPO-CD	Of course, in a subway line there are many subway trains
9	RBPO-CD	A subway train can do many rides
10	RBPO-CD	Each ride with a number of cars
11	RBPO-CD	Err ...a ride is conducted obviously by a real conductor, and a real subway train
12	RBPO-CD	The subway line has many conductors
13	RBPO-CD	A conductor can drive in many subways lines
14	PO-ACO	A conductor has an identification, a age, and years of experience
15	RBPO-CD	A conductor drives in different rides
16	RBPO-CD	Within each ride there is no information about how many conductors there are?
17	TP	Well, having analysed the diagram I will start to understand the expression
18	TE	The context is Subway Line, an invariant expression
19	PO-CA	That means that the expression represents a constraint that always must be true
20	RBPO-NR	Self dot has (self.has) is talking about the drivers that the subway line has,
21	RO-CO	then we have a selection with a c variable of conductor
22	RO-EI	that means that c select the conductors of each subway line
23	RO-CO	c dot age (c.age) greater than fifty, that is select the conductors which have an age greater than fifty, and then the boolean operation 'is empty' is applied
24	EO	that is the operation should be true
25	CHU	The meaning of the expression is ... [writing] each subway line should verify that ...
26	CHU	[writing] none of its associated drivers are more than fifty years old.
27	SCC-NT	I've finished...

Table 3. Intercoder Agreement

	Subject 1			Subject 2			Subject 3		
Protocol number	1	2	3	1	2	3	1	2	3
Percentage of Agreement	0.75	0.85	0.81	0.88	0.75	0.81	0.85	0.86	0.85
Coehn' K (Kappa)	0.75	0.85	0.81	0.86	0.69	0.76	0.82	0.83	0.81

et al. [12] to reveal the perspective taken by the subject in solving the missionary and cannibals problem.

A dotted rectangle was depicted in Figure 3 to highlight the RO, RBPO, PO and EO coded utterances during the process of the OCL expression comprehension. As we can see few dots remain outside the limit of this rectangle.

During task 3 he had doubts about the meaning of a collection operation that appears in the OCL expression and he spent a little time using the list of collection operations in order to look for the meaning. In task 3 the subject made various attempts to chunk the whole expression before giving the final result.

The composition of the coded utterances for each protocol is depicted in the first line of Table 4 (see appendix A). The coded utterances related to coupling were grouped into three groups RO, PO and RBPO (reified objects, problem objects and the relationships between them). We also grouped the special coding categories codes in the SCC group. TP and SP were depicted together due to the fact that they refer to the problem or to the way in which the subject plans to tackle it.

From Table 4 we see that the quantity of coupling coded categories increases at the same time as the object coupling within the expression increases.

If we subtract the number of SCC coded utterances (special coded categories which are not relevant to the model itself) from the total number of utterances (of each protocol), the coupling coding categories (RO, RBPO, PO and EO) represent 0.65%, 0.77%, and 0.83% of protocols 1, 2 and 3 respectively. This means that coupling coding categories seem to be a significant proportion of the cognitive process of subject 1.

We conclude that the RO category represents a significant proportion of the utterance in the three tasks; the quantity of RBPO and PO coding categories is higher in the last two protocols; the highest quantity of RBPO coding protocol is visible in Protocol 2 (the complexity of the expression in test 2 relies upon using different rolenames).

Analysis of Subject 2 The number of utterances of each protocol obtained from subject 2 are: 26 (protocol 1), 45 (protocol 2), 36 (protocol 3). The time (in minutes and seconds) spent on each protocol (1, 2 and 3) was: 02:52, 08:04, 05:55. The fact that, in protocol 2, the subject spent the longest time in understanding an expression is depicted in the graph of the subject's protocols (see Figure 4). We believe that the second test was more difficult for the subject to verbalize than the third one.

From Figure 4 we can also observe that the subject attempts to understand the class diagram before starting to comprehend the OCL expression. The same technique was applied in the three protocols: the subject focuses on the relationship between the classes while chunking the diagram. Moreover, in the last two protocols the subject looks at the contextual type of the expression in order to start to see the relationships that this contextual type has in the diagram before starting to comprehend the expression. We have depicted this regular situation with a rectangle (with solid lines) which contains all the coded utterances that correspond to the comprehension of the class diagram. The subject only focused on relations between classes, and said nothing about the classes attributes.

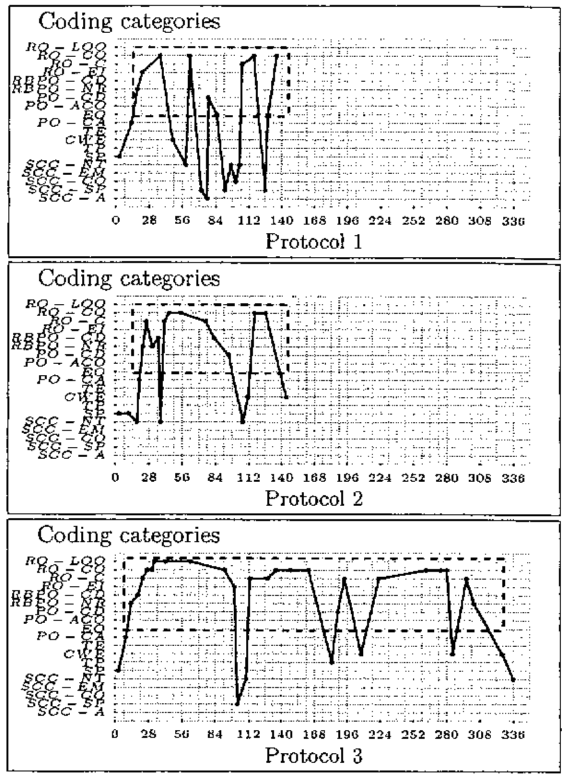


Figure 3. Coded Protocols of Subject 1

A dotted rectangle in Figure 3 was depicted to highlight the RO, RBPO, PO and EO coded utterances during the process of the OCL expression comprehension.

The composition of the coded utterances is shown in the second line of Table 4 (see appendix A). From Table 4 it is clear that the RO and RBPO coding categories are significant portions in protocol 2 and 3 respectively,

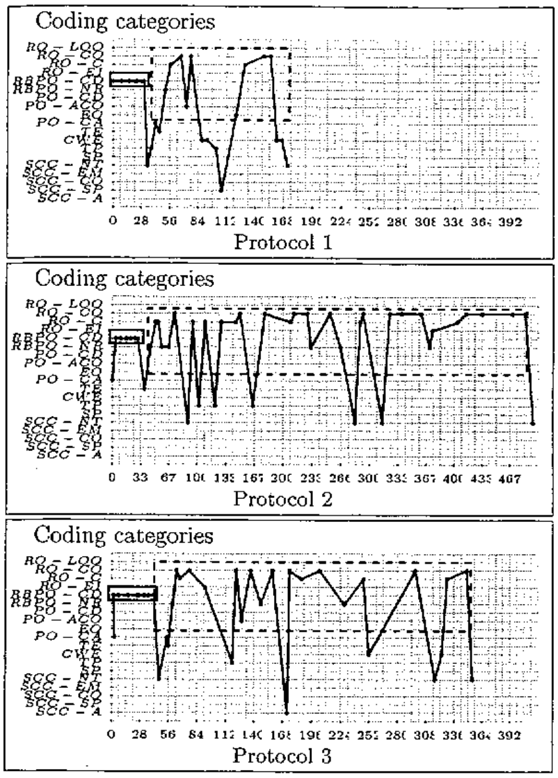


Figure 4. Coded Protocols of Subject 2

representing almost half of the coupling categories. We observe that the quantity of coupling coded categories is higher in protocol 2 than in protocol 3. Nevertheless, its ratio (number of coupling categories divided by the number of the protocol utterances) is 0.82 and 0.77. If we do not consider the SCC coded utterance from the total number of protocol utterance, this ratio is 0.90 and 0.87 respectively. The ratio of protocol 1 is 0.58 (considering the total number of utterances) and 0.66 (without taking into consideration the SCC coded protocols). This means that coupling coding categories seem to be a significant proportion of the cognitive process of subject 2.

Analysis of Subject 3 The number of utterances of each protocol are: 28 (protocol 1), 43 (protocol 2), 47 (protocol 3). These quantities include SCC coding categories. The time (in minutes and seconds) spent on each protocol (1, 2 and 3) was: 03:07, 06:41, 06:27. From Figure 5 we can also observe that the subject attempts a broad understanding of the class diagram before starting to comprehend the OCL expression. The subject applied the same technique in the three protocols: he focused on the relationship between the classes while reading the diagram. He also focused on the classes involved and their attributes. Again, this regular situation was depicted with a rectangle which includes the coded utterances related to the understanding of the relationship between the classes in the diagram as well as the understanding of the classes and their attributes.

The composition of the coded utterances for each protocol is depicted in the third line of Table 4. The ratios of coupling coding categories from the total number of utterances are 0.79, 0.79 and 0.83. Similar values are obtained if we subtract the SCC coded protocols from the total number of utterances: 0.81, 0.88, 0.89. This means that coupling coding categories seem to be a significant proportion of the cognitive process of subject 3. The quantity of the RBPO category is similar in protocol 2 and 3. Nonetheless, the quantity of PO is slightly higher in protocol 2 than in 3 (it should be remembered that the second protocol uses an expression in which the complexity is due to the use of different rolenames). The third protocol according to the complexity of the third OCL expression demands more RO utterances from the subjects than the first and second protocols.

4.2. Results of the Verbal Protocols

According to each hypothesis, the conclusions are:

1. The scope of comprehension of the class diagram taken by the subjects was different in each case. Nevertheless, the technique that each subject applied was the same in the three protocols. Subject 1 was able to understand the OCL expression correctly without much understanding of the class diagram before he started to comprehend the OCL expression. However subject 2 focused on the relationship between classes in the diagram before he attempted to comprehend the OCL expression, whereas subject 3's breadth of familiarity with the class diagram was the widest. Subject 3 focused on relationships, classes and attributes before he started to comprehend the OCL expression. This difference was depicted by using a rectangle (with solid lines) in the graph which depicts the coded utterances (see Figures 4 and 5). According to the situation of subject 1 the rectangle (with solid lines) is missing from Figure 3 due to the fact that there was absolutely no comprehension of the class diagram before the subject started to chunk the expression.

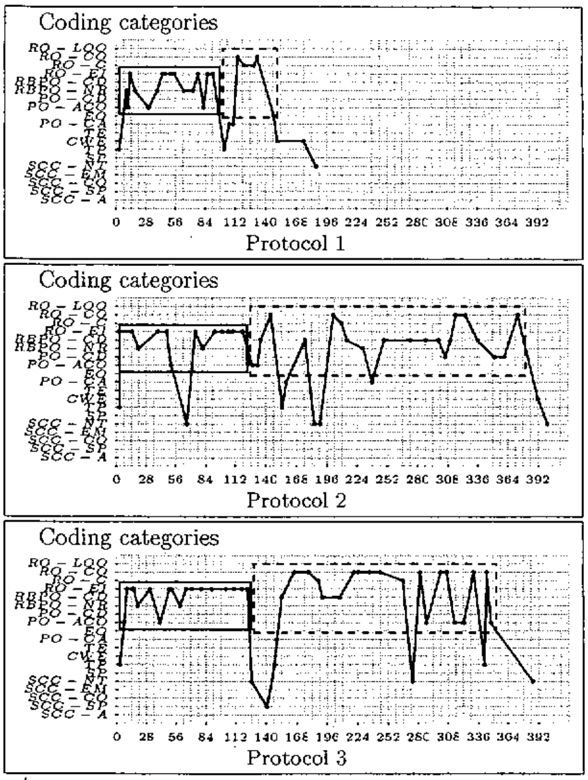


Figure 5. Coded Protocols of Subject 3

2. Tracing and chunking techniques are applied during the comprehension of OCL expressions. Nonetheless, it is not possible to produce a graphical model (such as a landscape model [4]) of how the subject switched between these cognitive techniques from the verbal protocol data. In order to model a landscape model we need to use a coarser-grain information during comprehension such as that likened to tracking eye movements from the UML diagram to the OCL expression (and vice versa). The eye fixations can support us to indicate what information is being heeded at any moment in time [5]. Nevertheless, the verbal protocol provides us with evidence that tracing and chunking are applied: (1) During the comprehension of the OCL expression the subject traces the class diagrams. This includes those subjects who attempt a more systematic comprehension of the class diagram before chunking the expression (for example subject 3). This is confirmed by the various verbal utterances related to the relationships (RBPO-CD), and the problem objects (PO-CD) which include aspects of multiplicity of relationships, etc. (2) From the recorded video it is possible to visualize that the subject sometimes follow some navigations by using a pencil which is moved over the classes in the diagram.

3. RO, RBPO and PO are a significant part of the mental model of subjects when they deal with OCL expression comprehension. This was measured according to the ratios of coupling coded utterances by the number of utterances of each protocol. The range of the ratios obtained for all the subjects was between 0.7 and 0.9.

5. Conclusions

The hypothesis that import coupling influences OCL expression maintainability is being assessed through the definition of valid measures for OCL expressions [17, 18, 16]. However in order to obtain reliable measures we need to study the cognitive complexity of the modelers dealing with these measured artifacts. Qualitative methods need to be employed in order to fully undertake a solid cognitive explanation of the concepts being measured, and it was for this reason that we carried out a think aloud experiment. The findings obtained in the experiment described in this paper should be considered as preliminary, Nevertheless, empirical evidence is gathered about:

- Problem objects, relations between problem objects and reified objects are a significant part of the mental model of subjects when they comprehend OCL expressions.
- Subjects take different scope of comprehension of the class diagram before starting to comprehend an OCL expression. The range varies in a continuous form which extends from those subjects who made absolutely no attempt to comprehend the diagram to those who attempted to systematically comprehend the class diagram.

We are planning further empirical studies to confirm these findings by considering studies of the human effort needed to perform OCL tasks and the difficulty in understanding the OCL expressions. In this context, cognitive informatics play an important role in providing the underlying theories [21], [13].

Acknowledgements

This research is part of the COMPETISOFT project (506AC0287) supported by CYTED (Programa Iberoamericano de Ciencia y Tecnología para el Desarrollo), the MECENAS project (PBI06-0024) financed

by 'Conserjería de Ciencia y Tecnología de la Junta de Comunidades de Castilla La Mancha' and the ESFINGE Project: 'Evolución de Software Factories mediante INGeniería del software Empírica'.

References

[1] L. C. Briand, C. Bunse, and J. W. Daly. A Controlled Experiment for Evaluating Quality Guidelines on the Maintainability of Object-Oriented Designs. *Software Engineering*, 27(6):513–530, 2001.

[2] L. C. Briand, J. W. Daly, and J. Wüst. A Unified Framework for Coupling Measurement in Object-Oriented Systems. *Fraunhofer Institute for Experimental Software Engineering*, 1996.

[3] C. Calero, M. Piattini, and M. Genero. Method for Obtaining Correct Metrics. *Proc. of the 3rd International Conference on Enterprise and Information Systems (ICEIS'2001)*, pages 779–784, 2001.

[4] S. N. Cant, D. R. Jeffery, and B. Henderson-Seller. A Conceptual Model of Cognitive Complexity of Elements of the Programming Process. *Information and Software Technology.*, 7:351–362, 1992.

[5] M. Chi. Quantifying Qualitative Analyses of Verbal Data: a Practical Guide. *Journal of Learning Sciences*, 1997.

[6] V. Chiew and Y. Wang. From Cognitive Psychology to Cogntive Informatics. In *Proceedings of the 2nd IEEE International Conference on Cognitive Informatics*, pages 4–5. IEEE Computer Society, 2003.

[7] D. P. Darcy and S. A. Slaughter. The Structural Complexity of Software: An Experimental Test. *IEEE Trans. Softw. Eng.*, 31(11):982–995, 2005. Member-Chris F. Kemerer and Member-James E. Tomayko.

[8] K. A. Ericsson and H. A. Simon. *Protocol Analysis: Verbal Reports as Data.* MIT Press, Cambridge, Mass., 1993.

[9] D. Glasberg, K. Emam, W. Melo, and N. Madhavji. Validating Object-Oriented Design Metrics on a Commercial Java Application., 2000.

[10] J. Huges and S. Parkes. Trends in the Use of Verbal Protocol Analysis in Softw. Eng. Research. *Behaviour Information Technology*, 22(2):127–140, 2003.

[11] B. C. Hungerford, A. R. Hevner, and R. W. Collins. Reviewing software diagrams: A cognitive study. *IEEE Trans. Softw. Eng.*, 30(2):82–96, 2004.

[12] E. L. Hutchins and J. A. Levin. Point of View in Problem Solving. *CHIP Tech. Rep. No. 105, Umv. Califorma at San Dingo.*, 1981.

[13] S. Misra and A. K. Misra. Evaluating Cognitive Complexity Measure with Weyuker Properties. In *ICCI '04: Proceedings of the Third IEEE Int. Conference on Cognitive Informatics (ICCI'04)*, pages 103–108, 2004.

[14] I. Nunes. An OCL Extension for Low-Coupling Preserving Contracts. In *UML*, pages 310–324, 2003.

[15] OMG. UML 2.0 OCL Final Adopted Specification. OMG Document ptc/2003-10-14, 2003.

[16] L. Reynoso, M. Genero, and E. Manso. Measuring Object Coupling in OCL Expressions: A Cognitive Theory-Based Approach. In *Proceedings of the IEEE Instrumentation and Measurement Technology Conference. Sorrento, Italy.*, pages 1087–1092, 2006.

[17] L. Reynoso, M. Genero, and M. Piattini. Measuring OCL Expressions: An Approach Based on Cognitive Techniques. *Imperial College Press, UK.*, 2005.

[18] L. Reynoso, M. Genero, M. Piattini, and E. Manso. Assessing the Impact of Coupling on the Understandability and Modifiability of OCL Expressions within UML/OCL Combined Models. In *11th IEEE International Symposium on Software Metrics (METRICS 2005)*, page 14, 2005.

[19] C. B. Seaman. Qualitative Methods in Empirical Studies of Software Engineering. *IEEE Transactions on Software Engineering*, 25(4):557–572, 1999.

[20] M. Someren, Y. Barnard, and J. Sandberg. *The Think Aloud Method: a Practical Guide to Modelling Cognitive Processes.* Academic Press, London, 1994.

[21] Y. Wang. Psychological Experiments on the Cognitive Complexity of Fundamental Control Structures Systems. In *Proceedings of the 4th IEEE Int. Conference on Cognitive Informatics*, pages 4–5, 2005.

[22] J. Warmer and A. Kleppe. *The Object Constraint Language. Second Edition. Getting Your Models Ready for MDA.* Addison-Wesley. Massachusetts., 2003.

A Coding Categories

Table 4. Coding Categories from Subjects