On the Comprehension of Security Risk Scenarios

Ida Hogganvik, Ketil Stølen
SINTEF ICT, PO box 124 Blindern, N-0314 Oslo
{iho, kst}@sintef.no

Abstract

Methods for security risk analysis are often based on structured brainstorming (e.g. what [21] calls HazOp). A structured brainstorming gathers a group of different system experts and the idea is that they will find more risks as a team than one-by-one. The CORAS modelling language [20] has been designed to support the brainstorming process and to document security risk scenarios identified during these sessions. The language is graphical, based upon the Unified Modelling Language (UML) [22], and is recommended by OMG [9]. This paper reports the results from two empirical experiments concerning the CORAS language. Our results show (1) many security risk analysis terms are used in the daily language and therefore well understood, but the more abstract or less frequently used terms can be a possible source for misunderstandings in a security analysis, and (2) the language's graphical icons make diagram "navigation" faster, but the diagrams are not necessarily understood more correctly than those without graphical icons.

1. Introduction

The participants in a structured brainstorming are people with different expertise on the target system. The idea is that they see the system from different viewpoints and therefore in collaboration will identify more risks than through individual expert judgements. The participants may have little or no previous experience with security risk analysis and therefore they need a common and intuitive way of documenting the security risk scenarios they identify. In the following we call security risk analysis just *security analysis*. The CORAS language [20] is a graphical modelling language that has been developed to support brainstorming and help document the results. The CORAS language was developed within the EU project "CORAS" [3, 10, 12, 14-16] which gathered well known risk analysis techniques into one common methodology for security analysis. The CORAS language is built on UML [22] as a UML profile and is recommended by the international standardization consortium OMG [9].

The aim of this study is to investigate the *usability* of the CORAS language. We will look at **(1) the conceptual model**, on which the abstract syntax is based, and **(2) the external graphical representation**.

The underlying conceptual model incorporates terminology from several relevant standards [1, 2, 4-8]. The abstract syntax of the CORAS language is based on this conceptual model for security analysis. Clearly, this model must be intuitive for the CORAS language to be successful during brainstorming sessions.

The external representation of the CORAS language is recognized by its special graphical icons that symbolizes the different terms in the conceptual model. The icons are believed to make the security risk scenarios easier to read and understand (in the following security risk scenario is called risk scenario). Often graphical icons are seen upon as merely "decoration" just to make diagrams look nicer, but we believe that for people that are unfamiliar with system modelling, icons help understanding diagrams.

We report the results from two student experiments that have been conducted to investigate the issues discussed above. The first tested the understanding of the underlying conceptual model and showed that the interpretation of security analysis terms is influenced by how they are commonly used in daily language. The second experiment tested the use of graphical icons and the effect on the understandability of risk scenarios. The icons were found to increase the speed of risk scenario navigation (i.e. identification of different elements), but not the correctness of risk scenario interpretation. For a full description of the experiments we refer to the technical report [17].

The paper is structured as follows: Section 2 provides an introduction to structured brainstorming and the CORAS language. Section 3 describes the design of the two experiments and Section 4 reports the results from statistical analyses. In Section 5 we discuss these results and possible interpretations, Section 6 reports lessons learned and Section 7 concludes.

2. Structured brainstorming and the CORAS language

A structured brainstorming is a methodical assessment of a system where people with different roles and

Proceedings of the 13th International Workshop on Program Comprehension (IWPC'05)
1092-8138/05 $ 20.00 **IEEE**

competence participate. Through discussion they identify risks and suggest treatments under the guidance of an analysis leader. The participants need to have a clear understanding of the terminology the techniques is based on. Terms like frequency, probability, vulnerability and asset are commonly used, but the individual interpretation may differ depending on the background and experience of the person in question. Several of the concepts are often used in everyday language, but not with a unique, context independent interpretation.

2.1 Structured brainstorming

The participants in a structured brainstorming are potentially anyone with an interest in the target system. They may be system administrator, system developer, company security responsible, customer support personnel or represent other roles. They receive no special training in advance, only an introduction at the beginning of the brainstorming session. During the session the team assesses system models in a stepwise and structured manner, supported by guidewords that help them identify potential security risks. The findings are modelled in risk scenarios using the CORAS language.

2.2 The underlying conceptual model for the CORAS language

The most important terms of the CORAS conceptual model are: risk, unwanted incident, threat, vulnerability, treatment, consequence, frequency, asset, probability and stakeholder (Figure 1).

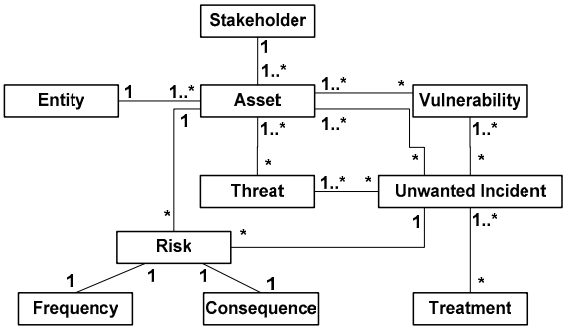

Figure 1 The CORAS conceptual model of security analysis terms

The associations between the elements have cardinalities that say how many instances of one element can be related to one instance of the other. Example: "a stakeholder has at least one and maximum infinitely many assets; and an asset belongs to only one stakeholder".

We explain Figure 1 as follows: the system or part of a system, assessed during a security analysis is called the **target of evaluation**. Anybody in contact with the target is a **stakeholder** of the system; system users, system maintainers and system developers are typical stakeholders. Different stakeholders often value the system differently; a system user who is dependent on the system will put a high value on it, while other stakeholders might not value the system equally high. The same **entity** may be assigned different values by different stakeholders. We refer to these entities with their values as assets. An **asset** is something to which a stakeholder directly assigns value and, hence, for which the stakeholder requires protection. An asset is therefore uniquely linked to a single stakeholder. A stakeholder wants to protect his/her assets from loosing value. Examples of assets are customer information, source code, company routines, critical system services etc. Target system stakeholders and their assets are normally identified early in the security analysis process. Figure 1 includes four important security analysis concepts related to asset: vulnerability, unwanted incident, threat and risk. A **vulnerability** is a weakness or lack, making an asset vulnerable to harmful actions. One may understand a vulnerability as something that is missing, e.g. if a company network lacks a firewall then this may be a vulnerability with respect to some assets in the network. An **unwanted incident** is an event that may harm the asset and something we want to prevent. An unwanted incident is the result of a threat exploiting a vulnerability. If the company network is an asset, then an unwanted incident is unauthorized access to the network by intruders. A **threat** is someone or something that wants to destroy, remove or interfere with the asset and a **risk** is the chance of this happening. With respect to the already mentioned company network, a threat may be a person who knows or discovers the vulnerability and wants to exploit it. First the company does not recognize the situation as a potential risk because nobody outside the company is aware of the security hole, but when an employee is fired, they suddenly realize that there is a risk for unauthorized network access by people familiar with the company infrastructure. The risk is characterized by a **risk value** (e.g. low, medium, high or other scales) which is based upon the estimated **frequency** for it to happen and its **consequence** in loss of asset value. If a risk is estimated to occur two times a year and the consequence is a loss of 200.000 dollars each time, the risk value could be "high" which means the risk should be treated. The **treatment** is applied either to the unwanted incident, the threat or the asset's vulnerability and the desired effect is reduced frequency and/or consequence, i.e. a reduced risk value.

2.3 Stereotyping with graphical icons in the CORAS language

A security risk scenario will typical illustrate how unwanted incidents relate to assets and treatments, it illustrates the threats and how they through various threat scenarios can harm the assets. A security risk scenario shows associations and relationships that are infeasible to describe with text only. Figure 2 illustrates one of the risk scenarios that were used as material in our icon experiment. The diagram illustrates three assets (Software token, hardware token and service availability) associated with two unwanted incidents (Token stolen and token missed) and two ways this can happen (the two threat scenarios). One of the threat scenarios is initiated by an attacker (Malicious person). Two treatments are identified, "educate user" will reduce the likelihood for token missed and "revoke token" will reduce the consequence of both token missed and token stolen.

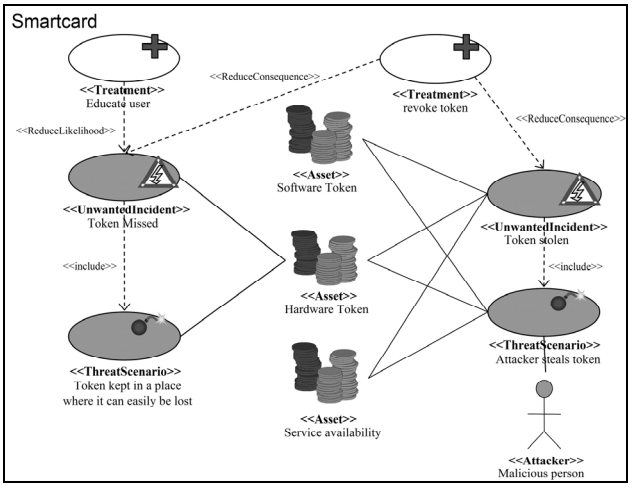

Figure 2 Security risk scenario for a smartcard

To make the scenarios easier to read, the concepts in the CORAS language are stereotyped with special icons and text-labels as illustrated in Figure 3.

Figure 3 Stereotyping with CORAS icons

Stereotyping is a technique offered by UML allowing you to mark normal UML elements with "labels" to emphasize that they are of a certain type, for more about this we refer to [11, 22]. Our focus is mainly on the graphical icon stereotyping and less on the stereotype name. We are not aware of other experiments with UML stereotyping, focusing on understanding, than [19]. The results from this experiment showed that icon stereotyping makes the models easier and faster to understand.

The treatment, unwanted incident and the threat scenario are all special cases of a UML use case, while asset and threat (here: "attacker") are UML actors. In our setting this implies that a risk scenario without icons will use normal UML symbols for use case and actor, illustrated in Figure 4.

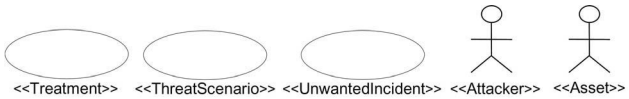

Figure 4 Stereotyping without CORAS icons

3. Description of the two experiments

Our experiments were conducted during the lectures in a university course at the institute for informatics, University of Oslo. The course has focus on the theoretic foundation for UML, refinement concepts in a UML context, modularity through contract-oriented specifications and model-based security analysis. The purpose of the two experiments was:

- to explore the comprehension of risk terms taken from the underlying conceptual model of the CORAS language (Terminology experiment).
- to investigate the effect of the graphical stereotyping used by the CORAS language (Icon experiment).

3.1 Material for the terminology experiment

The terminology experiment material was a questionnaire with 20 short statements. Each covered the relationship between 2-4 security analysis concepts that was either correct or incorrect. In addition to judging the statement, the subjects were asked to give a short explanation for their answer. The first section of the questionnaire asked some demographic questions about the subject's background and personal perception of the topic that should help us interpret the results from the analysis. To ensure that the questionnaire was completed on time, the subjects had to "tic off" an answer alternative for all the statements before writing their explanations. The questionnaire is presented in Table 1 and had three different answer alternatives: "correct", "wrong" and "I don't know" (originally in Norwegian):

Table 1 Terminology questionnaire

1. An unwanted incident can initiate new unwanted incidents by exploiting new vulnerabilities (correct).
2. A treatment will eliminate at least one unwanted incident (wrong).
3. The purpose of a treatment is to reduce risk, not necessary reduce vulnerability (correct).
4. An unwanted incident consists of consequence, frequency and a risk (wrong).

5. A treatment is always directed towards a vulnerability (wrong).
6. A threat is something that can initiate an unwanted incident (correct).
7. A risk is related to an asset's vulnerability (correct).
8. A threat can create a risk (correct).
9. There has to exist a threat before one think of an unwanted incident as a risk (correct).
10. A risk is something that will reduce the value of an asset (correct).
11. A vulnerability is the same as a risk (wrong).
12. A threat is something that has a potential to reduce the value of an asset (correct).
13. An unwanted incident cannot consist of multiple unwanted incidents (wrong).
14. An entity without value can also be considered as an asset (wrong).
15. A treatment can be directed towards a threat (correct).
16. A stakeholder is someone that wants to protect his or her assets (correct).
17. A threat can exploit a vulnerability (correct).
18. A successful treatment must reduce both consequence and frequency (wrong).
19. A risk and an unwanted incident is the same (wrong).
20. *One has to know both probability and frequency to calculate the risk value (wrong).*

A challenge in designing the questionnaire was which concepts to include, and which relationships to test. We selected 10 of the most common security analysis terms from CORAS' conceptual model in addition to "probability" which can be used to describe one kind of frequency. Some concepts are associated with more concepts than other, and this is reflected in the number of times they appear in different statements in the questionnaire. E.g. the term "risk" is included in more statements than "frequency" since it is associated with more concepts.

3.2 Material for the icon experiment

The icon experiment material was a set of security risk scenarios modelled with the CORAS language (example in Figure 2) and a related questionnaire. The risk scenarios were taken from a report on security threat modelling project resulting from a research collaboration between Microsoft Research in Belgium and SINTEF ICT [13]. The questionnaire focused on general understanding of notation making domain knowledge less important. Nevertheless, the models we used illustrated general security issues related to smartcard (credit card sized storage medium) and web services (software that makes itself available over the Internet using standardized XML) that could be known to some of the students.

Half of the subjects received risk scenarios stereotyped with both CORAS icons and stereotype names while the other half had standard UML use case icons and stereotype names. The questions and risk scenarios were the same for both groups. This experiment did not test stereotyping in general since we kept the stereotype name, but rather the graphical icon stereotyping. The questionnaire was divided into three parts:

Table 2 Experiment material

	Max time	Type of task	# Scenarios, # Questions	Max score	Variables measured
Part1	3 min	Training task for model navigation	1 risk scenario, 7 questions	7p	Score, # questions answered
Part2	1.5 min	Diagram navigation	1 risk scenario , 10 questions	10p	Score, # questions answered
Part3	15 min	Diagram under-standing	3 risk scenarios, 11 (4+5+2) questions	11p	Score, # questions answered, time used pr. scenario

Part 1 and 2 had questions that tested whether the subjects were able to quickly identify specific elements like treatment, risk, threat agent etc in a risk scenrio (diagram navigation). The questions were of the type "*How many assets are explicitly modelled in this diagram?"* Part 3 focused on how the risk scenarios are understood and interpreted. An example of the type of question used here is "*Which unwanted incidents affect the asset "Web server" in this diagram?"* Since part 3 included three different sets of risk scenarios with belonging questions, the students had to record the start time on each of the sets to make us able to analyse the time spent (in minutes) on each sets.

3.3 Execution and practical considerations

The master students attending the experiments had not been given information in advance since they did them as exercises during the lessons. At this level of education it is natural to assume that the students know some UML.

For the terminology experiment, we faced the situation that the subjects' previous knowledge about security analysis could vary from 0-4 hours. We believed this could influence the result and therefore we asked about this in one of the demographic questions. A short introduction to the type of task was given and the students were given 20 minutes to fill in the questionnaire.

The icon experiment took place in the same setting two weeks later. Before the experiment the students had a lesson with an introduction to the CORAS language. Also in a real structured brainstorming the participants will be given an introduction to the notation in advance, but then considerably shorter. The questionnaires were handed out in a randomized manner, 13 subjects received questionnaires with CORAS icons and 12 received the

one with standard UML use case icons.

3.4 Analysis models

The two experiments measured different variables, and had therefore different analysis models. In the terminology experiment we measured score for each statement looking for particular easy or difficult concept-relationships.

The icon experiment measured both score and time used per question /number of questions accomplished. The two groups of subjects were compared to see whether the one with CORAS icons had a higher score than the group without. In the same manner we analyzed whether one group managed to complete more questions than the other group or used less time pr. model-question-set. Since the sample set was expected to be relatively small and possible not normally distributed, we chose to use a one-tailed, non-parametric Mann-Whitney test with significance level 0.05.

4. Results from the two experiments

4.1 Statistics from the terminology experiment

31 subjects participated in the experiment. The shaded numbers in the table indicates statements that obtained a high percentage of either "correct answers", "incorrect answers" or "I don't know answers".

Table 3 Descriptive statistics of the statements

Statement (terms used)	Correct answers		Incorrect answers		"I don't know" answers	
	#	%	#	%	#	%
1 (unwanted incident, vulnerability)	23	74.2	3	9.7	5	16.1
2 (treatment, unwanted incident)	12	38.7	16	51.6	3	9.7
3 (treatment, risk, vulnerability)	20	64.5	9	29.0	2	6.5
4 (unwanted incident, consequence, frequency, risk)	9	29.0	19	61.3	3	9.7
5 (treatment, vulnerability)	14	45.2	9	29.0	8	25.8
6 (threat, unwanted incident)	25	80.6	3	9.7	3	9.7
7 (risk, asset, vulnerability)	23	74.2	4	12.9	4	12.9
8 (threat, risk)	19	61.3	8	25.8	4	12.9
9 (threat, unwanted incident, risk)	13	32.3	10	41.9	8	25.8
10 (risk, asset)	13	41.9	18	58.1	0	0.0
11 (vulnerability, risk)	25	80.6	4	12.9	2	6.5
12 (threat, asset)	25	80.6	6	19.4	0	0.0
13 (unwanted incident)	22	71.0	3	9.7	6	19.4
14 (entity, asset)	16	51.6	9	29.0	6	19.4
15 (treatment, threat)	19	61.3	8	25.8	4	12.9
16 (stakeholder, asset)	29	93.5	1	3.2	1	3.2
17 (threat, vulnerability)	28	90.3	0	0.0	3	9.7
18 (treatment, consequence, frequency)	16	51.6	11	35.5	4	12.9
19 (risk, unwanted incident)	23	74.2	4	12.9	4	12.9
20 (probability, frequency)	5	16.1	20	64.5	6	19.4

The distribution of data is also illustrated using histogram in Figure 5.

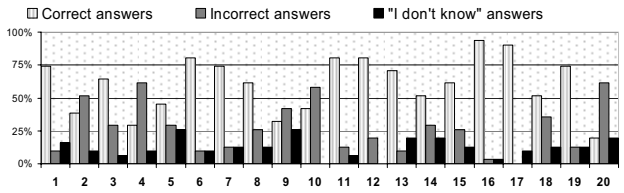

Figure 5 Histogram of data from terminology experiment

Not all concepts are used equally many times in the statements. Terms like risk, treatment, threat, vulnerability, unwanted incident and asset were used in 25% or more of the statements, while frequency, probability and consequence were used in less than 25% of the statements.

4.2 Statistics from the icon experiment

25 subjects participated, 13 of them received material with CORAS icons, 12 got material with standard UML use case icons. Table 4 contains descriptive statistics from the icon experiment. The results are grouped according to which part of the questionnaire they represent. **Score** denotes how many points the students scored, **# Answers** are the number of answers the students managed to complete within time and **Time** is the minutes used pr set of questions in part 3. We report mean and standard deviation for each of the groups.

Table 4 Descriptive statistics

		CORAS icons		UML icons	
		mean	st.dev	mean	st.dev
Part 1	**Score**	4.5	1.38	4.77	1.09
Part 2	**Score**	2.88	0.91	2.58	0.95
	#Answers	4.75	1.54	3.69	1.18
Part 3	Score, set 1	3.04	0.96	3.38	0.92
	Score, set 2	3.62	0.68	3.7	1
	Score, set 3	0.63	0.8	0.27	0.6
	Score, set1-3	7.29	1.41	7.36	1.82
	Time, set 1	5.58	1.44	6	1.53
	Time, set 2	5.3	1.34	6.7	1.25
	Time, set 3	2.71	0.95	-*	-*
Total	**Score**	14.67	2.11	14.7	2.51

*only one subject finished

4.3 Statistical tests for the icon experiment

The results from the Mann-Whitney test is presented in Table 5, the critical value in column five is taken from [23]. Both for #Answers (part 2) and Time, set 2 (part3) the results showed a significant difference between the two groups.

Table 5 Mann-Whitney test

		n1	n2	critical value	u
Part 1	Score	12	13	<=47	71.5
Part 2	Score	12	13	<=47	67.5
	# Answers	12	13	<=47	46
Part 3	Score, set 1	12	13	<=47	62
	Score, set 2	12	13	<=47	70.5
	Score, set 3	12	13	<=47	56.5
	Score, total	12	13	<=47	77.5
	Time, set 1	12	13	<=47	71
	Time, set 2	10	10	<=27	21.5
	Time, set 3	7	1	-	-
Total	Score	12	13	<=47	76

5. Discussion

This section discusses both experiments, first the terminology experiment, then the icon experiment.

5.1 Terminology experiment

The results from our terminology experiment show that most of the concepts and relationships from the CORAS' conceptual model are well understood. The easiest statements seem to be the ones which include terms used in the daily language, like *stakeholder*, *asset*, *threat* and *vulnerability*.

We tested several relationships and in the following we discuss those that obtained high percentages of correct (Figure 6), incorrect (Figure 7) and "I don't know" answers (Figure 8). In the remaining we refer to the statements using their number: #3 means statement 3 in the questionnaire (in the figures the "#" is not used).

The single most intuitive relation is the asset-stakeholder relationship (#16). A large majority of the subjects (93.5%) believe "a stakeholder is someone that wants to protect his or her assets" is correct. 90.3% of the subjects agreed in "a threat can exploit a vulnerability" (#17). The relation between asset and threat (#12) obtains over 80% correct answers and over two third agree in "a risk is related to an asset's vulnerability" (#7).

Our impression is that the subjects find it easier to deal with concrete concepts, i.e. an unwanted incident is something real, as opposed to the more abstract term "risk". As an example the statement "a threat is something that can initiate an unwanted incident" (#6) obtained a high level of correct answers (80.6%), but in a similar statement (#8), where we substituted unwanted incident with risk, we received considerably less correct answers (61.3%).

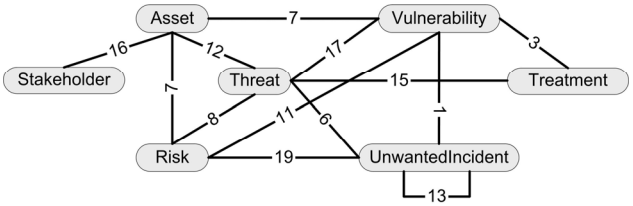

Figure 6 Well understood relations

The subjects seem to be unaware of the fact that the purpose of a treatment often is to *reduce* frequency or consequence to an acceptable level. More than half of the subjects (51.6 %) believe a treatment always will *eliminate* at least one unwanted incident (#2).

In the CORAS conceptual model a risk is defined to consist of an unwanted incident, a frequency and a consequence. This means that risk in a way is an abstract concept, something that seems troublesome for the subjects. The terms "unwanted incident" and "risk" are often confused, 71 % of the subjects were wrong or uncertain on whether a risk includes an unwanted incident or vice versa (#4).

In statement #10: "a risk is something that will reduce the value of an asset", more than half of the subjects disagreed (58.1%). They disagreed in the use of "will" in the statement, claiming that a risk with frequency near zero "*may* reduce...", not "*will* reduce...*"*. We agree in their argumentation and will reformulate the statement in further work with the conceptual model.

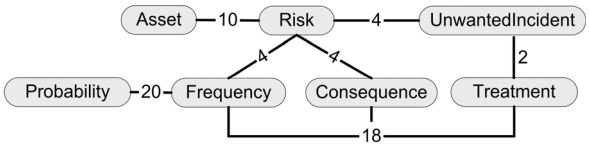

Figure 7 Misunderstood relations

The single most difficult statement in the questionnaire was "one has to know both probability and frequency to calculate the risk value" (#20). A risk value is a function of consequence and frequency, but few seem to be aware of that probability and frequency are two measures for the same thing. A large group (83.9%) either gave incorrect answer or did not know the answer on this statement. Neither frequency nor probability is particularly common in the everyday language and this is probably why we get a high percentage of incorrect or uncertain answers. It is possible that just by avoiding the exact terms and rather

ask “how often?” or “how many out of 100?” is enough to reduce the uncertainty among the participants.

More than half of the subjects (54.8%) either think a treatment will always be directed towards a vulnerability or they are uncertain on the answer (#5). The subjects did not see the possibility of applying a treatment to an unwanted incident or a threat.

According to CORAS’ conceptual model an unwanted incident is always associated with one or more threats, but a threat is not necessary associated with an unwanted incident. The statement “There has to exist a threat before one think of an unwanted incident as a risk” (#9) confused the subjects. A majority either disagreed (41.9%) with this interpretation or were uncertain (25.8%), and they did not accept this as an intuitive explanation for the relation between threat, unwanted incident and risk. We believe that some of this confusion arises from a misunderstanding that an unwanted incident in itself is a threat (comments made by the subjects in the questionnaire).

The subjects were uncertain on how an entity relates to an asset (#14) and gave inconsistent answers. The reason may be that “entity” is a term that hardly exists in the daily language and we think this created more uncertainty than the actual relationship to asset.

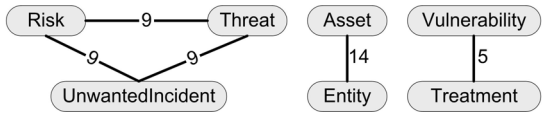

Figure 8 Relations creating uncertainty (answer=”I don’t know”)

It was not identified correlation between data from the initial demographic questions and score. Surprisingly enough, attendance to the previous lesson did not affect the overall score on the questionnaire. This indicates that many security analysis concepts are already known to people, but often with different interpretations.

5.2 Icon experiment

What we can see from the icon experiment is that by stereotyping with CORAS icons the subjects managed to answer more questions than the group without CORAS icons, but it did not affect the “correctness” of their answers. Both risk scenario sets kept the stereotype name; otherwise it would be impossible to distinguish the UML icons. If the subjects had based themselves upon reading the text under the icons there had not been any difference between the two groups, but looking at the results it is possible to assume that the stereotype name is ignored. We did ten statistical tests of the data from the icon experiment (not including the test on *time used for set 3* since the number of subjects was too small). For two of the tests we obtained a statistical significant difference between the two groups of subjects at a significance level of 0.05.

Part 1 and 2 had questions that aimed to investigate whether the subjects were able to differentiate between the elements in the model. They were asked to count the number of different security analysis concepts, some were in fact represented in the models and some were not. Part 1 was a training task with no time pressure; therefore the differences between the groups are insignificant both with respect to correctness and number of answers. Part 2 was similar to part 1 but with less time available and here we obtained a significant difference in the number of questions answered between the CORAS icon subjects and the other group. The subjects with CORAS icons managed to answer in average 4.75 questions, while the result for the group without CORAS icons is 3.69. There was insignificant difference with respect to the correctness of the answers, the group with CORAS icons scored in average 2.88 points while the others scored 2.58 points.

Part 3 of the questionnaire had a different style than the previous parts. This section focused more on interpretation of risk scenarios, and less on navigation. We used three risk scenarios with belonging sets of questions (Set 1, Set 2 and Set 3). All sets were measured for score and time used pr set. The results show that there was no significant difference between the groups in Set 1. We see this set as a training set since the type of questions was different from the previous ones and therefore both groups used some time to get used to the questions. For Set 2 we obtained a significant difference with respect to the time used. The group with CORAS icons used less time to complete the questions (average: 5.3 min) than the group using standard UML use case icons (average: 6.7 min). But also here there were no difference in score between the two groups. It was not possible to statistically analyse the difference between the two groups in Set 3 since seven from the CORAS icon group finished it, but only one from the group using UML icons. We can only assume that this indicates that the CORAS icon group performed better also in this set.

[19] reports the results from a similar experiment investigating the understanding of stereotyped UML class diagrams. The experiment tested icon-stereotyping vs. normal class stereotyping (roughly sketched in Figure 9**)**. They substituted traditional stereotyped classes (Alt.1) with special stereotype icons (Alt.2) and found this to improve both the speed of navigation and the correctness of risk scenario interpretation. According to their findings we should expect to find an improvement in both areas but we only found the difference in the navigation speed. A possible reason is that many of the security risk concepts do not have obvious graphical representations and that our icons may be ambiguous. Also the interpretation of a risk

scenario may depend on other factors than icons like naming of elements, modelling style and familiarity with the system or domain that is modelled. At this point our results support the belief that icons affect how fast the subjects navigate through the risk scenario to identify different elements, but less on how they interpret the scenario.

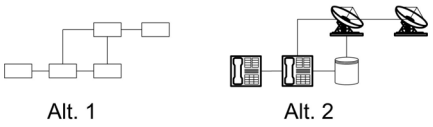

Figure 9 Class diagram with and without icons

5.3 Threats to validity for the experiments

Experiments will often have some threats with respect to the validity of its results. In the following we explain the ones we have identified.

For the terminology experiment:

- The questionnaire design: developing a good questionnaire using natural language is difficult. There is always possible to write text that is interpreted different from its intention. Even though our questionnaire was both pilot tested and reviewed, we found some weaknesses after the experiment. We believe statement #11 "a vulnerability is the same as a risk" failed to express what we really intended to check. We wanted to investigate whether a vulnerability automatically is considered as a risk, or do the subjects recognize the need for identification of unwanted incidents also. Unfortunately it was probably quite easy to guess that the two terms were not identical and mark "Incorrect". The same goes also for statement #19 " a risk and an unwanted incident is the same" where the intention was to investigate whether the subjects were aware of that unwanted incident is a part of a risk, not if the subjects see the difference between two words.
- Data interpretation: since one of the alternative answers for each statement was "I don't know", we choose to interpret missing answers as "I don't know" This gave us some data points that are not directly filled in by subjects, but since there were no good reasons for not answering, neither short time nor fear of giving wrong answers to the test we decided that this was legitimate.

For the icon experiment:

- Domain knowledge and questionnaire design: the students were not tested in their understanding of the specific domain modelled and the questions dealt with the CORAS language notation only. The risk scenarios were not too technical detailed, and taken from a real security analysis which means that they are representative for its type.
- Familiarity with material: the students had seen the CORAS icons before and were more used to them than the other icons. This is a problem, even if we used identical risk scenarios, kept the stereotype name in both and replaced the CORAS icons with two standard UML icons the students with CORAS models had a small advantage. Still we believe this was too small to make a real difference, in most cases the group with UML icons had similar mean score to the other group.
- Time pressure: it was important to have some time pressure to investigate fast diagram navigation. The questionnaire also contained extra tasks in the end to make sure that no subjects finished before time and thereby could disturb the other subjects.

In general:

- Norwegian experiment material: the findings reported in this paper are based on experiments with Norwegian students. Even though the experiment material was mainly in Norwegian and used Norwegian subjects we assume that our findings are not limited to yield only for Norwegians. The conceptual model is developed in an international project and all the concepts have precise translations into Norwegian.
- Student as subjects: Using students as subjects will always be criticized, but we believe that students are well suited for this explorative study. There are no other requirements than being familiar with a part of the system assessed in order to participate in or understand the documentation from a security analysis. These experiments did not require knowledge about a particular system; we only tested terminology and modelling notation. In our opinion, unless a company already conducts security analyses using this type of terminology, the concepts tested here will be just as unfamiliar for professionals as for students. Many of the students were 6-9 months from graduating with a master degree in informatics, and many will become employees in companies that conduct this type of security analyses. Both the experiments were completely anonymous so the students did not have to worry about score and grades, which could potentially influence their behaviour.

6. Lessons learned

The results from the terminology experiment have taught us that some parts of the conceptual model are not as intuitive as desired. On the basis of these results we propose a set of changes to the original model (see Figure 10):

- To avoid confusion about frequency, probability and consequence we have included probability as a

specialization of frequency.

- In order to emphasize that a risk consists of a frequency value, consequence value and an unwanted incident, these terms have been grouped together in a logical manner that illustrates how they are components of a risk. The black diamond symbolizes that if a risk is eliminated this will also remove the frequency and consequence values (composition). The white diamond means that the unwanted incident is a part of the risk, but if the risk is eliminated the unwanted incident may still exist in other risks (aggregation).

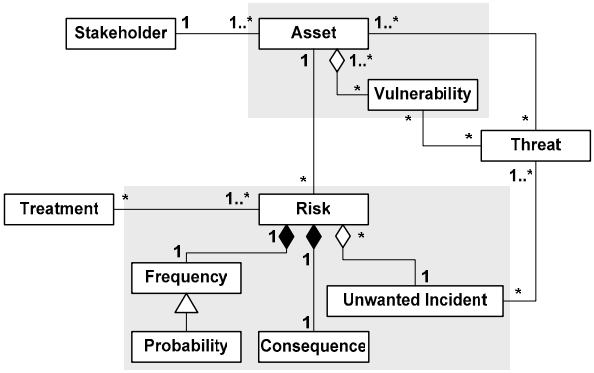

Figure 10 The revised conceptual model

- The direct relation between treatment and unwanted incident has been removed and instead we have connected treatment to risk. By doing this we specify that a treatment is directed towards a risk, but not whether it targets a vulnerability, a threat, an unwanted incident or a combination of these.
- The association between asset and risk was a major source for misunderstandings. Nevertheless we have decided to keep the relation because we believe it is important. The subjects seem to think of an unwanted incident in the way we use the term risk, and by removing the direct association between asset and unwanted incident we hope that this misunderstanding will be less common.
- We chose to remove "entity" from the model. The term was often misunderstood (48.9% incorrect or uncertain answers). In addition it is not natural to speak of general entities in a structured brainstorming, only entities assigned value which then are called assets.
- The association between asset and vulnerability has been changed from a regular relation to become a part of the asset in the sense that an asset can have a vulnerability.
- In the original conceptual model it is not explicitly modeled that "a threat exploits a vulnerability to initiate an unwanted incident". This is an intuitive relation that achieved a high percentage of correct answers and we feel it is correct to make it clearer than it was in the original model.
- The new model tries to emphasize that an unwanted incident is a part of a risk and therefore one or more threats are always associated with the risk through its unwanted incident.
- We have also chosen to highlight concepts grouped together in the form of compositions and aggregations, i.e. vulnerability is tightly connected to asset and risk is a concept that includes three other concepts.

7. Conclusion

A structured brainstorming within security analysis gathers system experts with the aim to identify risks and treatments for the system. The participants need to have a common understanding of concepts like asset stakeholder and threat, in addition to more traditional security analysis terms. If the analysis leader asks them to identify unwanted incidents for the system, it is important that the participants understand that this is different from risks or threats. They need to understand a security risk scenario quickly and interpret it in the same way. The CORAS language for structured brainstorming is specially developed to support structured brainstorming and document the results in security risk scenarios. The language is based on a conceptual model, consisting of security analysis specific terms. The suitability of its notation and its underlying conceptual model is very important for the usability of the language. We aimed to investigate this in two experiments, one focusing on the terminology in the underlying conceptual model and the other on the use of graphical icons. We used students as subjects and the results from this study will be used as input in further experiments with professionals to see whether we obtain similar results.

The results from the terminology experiment showed that few subjects had problems with relations like stakeholder and asset or threat and vulnerability. These terms are part of the daily language and most people have an intuitive understanding of them. Concepts like frequency, consequence and probability created more confusion and a large majority did not see probability as a kind of frequency. These results show the importance of using terms that are intuitive and do not conflict with the daily language.

The CORAS language is a UML profile that uses special graphical icons to symbolise selected security analysis terms. The results from the icon experiment showed that stereotyping with CORAS icons vs. only UML icons improves the speed of navigation, i.e. identification of specific risk scenario elements. The results showed a statistical significant difference in the

number of questions completed in favour of the group using CORAS icons. The icons did on the other hand not significantly affect the correctness of interpretation of risk scenarios.

The conclusion from our icon experiment is that using CORAS' graphical icons helps the participants in a structured brainstorming session to identify the same risk scenario elements faster than if one does not use CORAS icons. When it comes to correctly understand the scenarios we believe it is necessary to look at other aspects as well as graphical icons.

Future work

The future goal of our work is to develop a modelling language that supports and documents structured brainstorming, whose *usability* and *suitability* is thoroughly supported by empirical and analytical evidence. The assessment of the CORAS language has provided us with information that will be used as input in designing the new language. In addition to more student experiments, we have planned an analytical assessment of the CORAS language, based on the quality framework of [18]. Through industrial field studies in the SECURIS project we aim to test our hypotheses and gather experience from real users. The CORAS language will also be tested in SINTEF ICT's usability laboratory using subjects that are representative for real users. The results from these investigations will help us identify the requirements for a language that in the best possible ways supports and documents brainstorming sessions in security analysis.

Acknowledgements

The research on which this paper reports has partly been funded by the Research Council of Norway project SECURIS (152839/220). The authors thank Jan Heim and Fredrik Seehusen (SINTEF ICT) for valuable input.

References

[1] Australian/New Zealand Standard for Risk Management (AS/NZS 4360): Strathfield Standards Australia, 1999.

[2] British Standard BS4778: Quality vocabulary: Availability, reliability and maintainability terms. Glossary of international terms: British Standards Institute, 1991.

[3] CORAS, http://coras.sourceforge.net

[4] IEC 61508: Functional Safety of Electrical/Electronic/Programmable Electronic Safety-Related (E/E/PE) Systems, 1998-2000.

[5] IEEE-1471: "Recommended Practice for Architectural Description of Software-intensive Systems" (IEEE1471-2000 Std): IEEE, 2000.

[6] IEEE Std 610.12: IEEE Standard Glossary of Software Engineering Terminology, 1990.

[7] ISO/IEC 13335: Information technology - Guidelines for management of IT Security, 1996-2000.

[8] ITU-T X.800: Security architecture for open system interconnection for CCITT applications. (Technically aligned with ISO 7498-2). Geneva, 1991.

[9] Object Management Group (OMG), http://www.omg.org

[10] J. Ø. Aagedal, F. den Braber, T. Dimitrakos, B. A. Gran, D. Raptis, and K. Stølen, "Model-based risk assessment to improve enterprise security," *Proc. Enterprise Distributed Object Communication (EDOC'2002)*, pp. 51-62, 2002.

[11] C. Atkinson, T. Kühne, and B. Henderson-Sellers, "Stereotypical Encounters of the Third Kind," *Proc. 5th International Conference on The Unified Modeling Language*, pp. 100-114, Dresden, 2002.

[12] F. den Braber, T. Dimitrakos, B. A. Gran, M. Soldal Lund, K. Stølen, and J. Ø. Aagedal, *UML and the Unified Process*: IRM Press, 2003.

[13] F. den Braber, M. S. Lund, and K. Stølen, "Using the CORAS Threat Modelling Language to Document Threat Scenarios for several Microsoft relevant Technologies," SINTEF ICT, Technical Report (STF90 A04057), 2004.

[14] F. den Braber, M. S. Lund, K. Stølen, and F. Vraalsen, "Integrating Security in the Development Process with UML," in *Encyclopedia of Information Science and Technology*, M. Khosrow-Pour, Ed.: Idea Group Reference, 2005, pp. 1560-1566.

[15] F. den Braber, A.-B. Mildal, J. Nes, K. Stølen, and F. Vraalsen, Experiences from Using the CORAS Methodology to Analyze a Web Application, in *To appear in Journal of Cases on Information Technology*, 2005.

[16] T. Dimitrakos, B. Ritchie, D. Raptis, J. Ø. Aagedal, F. den Braber, K. Stølen, and S. H. Houmb, " Integrating model-based security risk management into eBusiness systems development - the CORAS approach.," *Proc. 2nd IFIP Conference on E-Commerce, E-Business, E-Government (I3E'2003)*, pp.159-175, 2002.

[17] I. Hogganvik and K. Stølen, "Empirical Investigations of the CORAS Language for Structured Brainstorming," SINTEF ICT (STF90 A05041), 2005.

[18] J. Krogstie and S. d. F. Arnesen, "Assessing Enterprise Modeling Languages Using a Generic Quality Framework," in *Information Modeling Methods and Methodologies*: Idea Group, 2005, pp. 63-79.

[19] L. Kuzniarz, M. Staron, and C. Wohlin, "An Empirical Study on Using Stereotypes to Improve Understanding of UML Models," *Proc. 12th IEEE International Workshop on Program Comprehension (IWPC'04)*, pp.14-23, 2004.

[20] M. S. Lund, I. Hogganvik, F. Seehusen, and K. Stølen, "UML profile for security assessment," SINTEF ICT, Technical report (STF40 A03066), 2003.

[21] F. Redmill, M. Chudleigh, and J. Catmur, *Hazop and software Hazop*: Wiley, 1999.

[22] J. Rumbaugh, I. Jacobson, and G. Booch, *The Unified Modeling Language Reference Manual*: Addison Wesley Longman, Inc., 1998.

[23] R. E. Walpole, R. H. Myers, and S. L. Myers, *Probability and Statistics for Engineers and Scientists*, 6 ed: Prentice Hall International, Inc., 1998.