Evaluating the Software Problem Representation on the Basis of Rationale Trees and User Story Maps: Premises of an Experiment

Konstantinos Tsilionis¹(✉) , Joris Maene¹, Samedi Heng² ,
Yves Wautelet¹ , and Stephan Poelmans¹

¹ KU Leuven, Leuven, Belgium

{konstantinos.tsilionis,yves.wautelet,stephan.poelmans}@kuleuven.be

² HEC Liège, Université de Liège, Liège, Belgium
samedi.heng@uliege.be

Abstract. User stories describe system requirements from the users' point of view using structured natural language. These artifacts mostly assist the development team in agile projects where sets of them express the expected features of the to-be software. Nevertheless, when faced with large user stories' sets, further structuring needs to be introduced to make the software problem more understandable/clear. Indeed, the latter can hardly be understood on the basis of huge flat lists of sentences. The present paper describes the premises of an experiment destined to evaluate whether novice modelers can understand the requirements problem better through a visual representation (called Rationale Tree) built out of a user stories' set. It does so by comparing, on the same objective, the Rationale Tree method to the industry-adopted User Story Mapping technique with a modeling exercise. This was performed with the participation of two student groups where one used the former technique while the other one employed the latter. The ultimate goal of this exercise is to identify the actual performance of the Rational Tree in delivering understandability to the requirements problem (when input is provided in the form of a complex user stories' set) in order to identify possible bottlenecks and improvement opportunities within the method itself.

Keywords: User stories · User story · Rationale Tree · User Story Mapping · Agile requirements engineering

1 Introduction

User stories are requirements expressions written in natural language, generally structured around three dimensions, i.e. the WHO, the WHAT and (possibly) the WHY [9]. They are usually manifested in the format 'As a <type of user>, I want <some goal> so that <some reason>' [1]. Even though the user

stories' simple structure is beneficial in describing globally the features of the to-be system especially when developed in an agile manner, still delineating thoroughly every aspect of the software requirements specification problem from a list of user stories remains challenging.

The *Rationale Tree* [8,10] is one way of dealing with this challenge; it represents a conceptual modeling-based method to visually represent elements extracted from a user stories' set and to link these elements. The method uses parts of the i* strategic rationale diagram constructs [12] and visual notation to build various trees of relating user story elements in a single project with the purpose of identifying depending user stories, identifying Epic ones and group them around common Themes. In essence, this visually-aided method aims to recognize and reduce any occurring modeling redundancies during the stages of requirements' analysis and design at the start of a software development project. The Rationale Tree was introduced theoretically in [10]; it was applied on real life cases in [8] and its ability as a stand-alone method to improve the understanding of user stories' set has been tested on novice modelers in [6,11]. On the other hand, *User Story Mapping* [3] is an industry-adopted user stories' structuring method based on listing, under the scope of an epic user story, all of the related lower-level functions described in ordinary user stories.

The study aims to overview whether a conceptual modeling-based approach such as the Rationale Tree offers a better resolution to the elicitation of user requirements when compared to a simpler user stories' structuring approach (User Story Mapping). Under this scope, an experiment – whose premises are presented in this paper – was organized with the participation of master-level students taking on the role of novice modelers. Half of these students were asked to perform a modeling exercise (see Sect. 2) based upon a case study's user story set by using the Rationale Tree method. At the same time, the other half of the group had to undergo the structuring of the same user story set but using the User Story Mapping technique. This paper overviews the ability of the first group to produce a rationale tree of acceptable quality and the one of the second group to produce a user story map of acceptable quality. Further analysis will be the subject of a following communication.

2 Research Approach and Sampling Technique

For the purposes of this study, we performed a *modeling exercise* (referred also as the *drawing exercise*) with the participation of a student group. Our goal was to capture the modeling capabilities¹ of novice modelers when employing the Rationale Tree method in order to discover possible benefits and/or drawbacks derived from its application. Since a generic class of novice modelers could be instantiated by many possible roles and task-specifications within the software development industry there was no available comprehensive sampling frame to rely on. Consequently we decided to use non-stochastic purposive sampling in

¹ Modeling capability will denote the capacity of novice modelers to produce qualitative artifacts while employing a specific user stories' structuring method each time.

the form of typical cases [4] representing our participating students. In particular, our sample consisted of the Master program of Business Administration with a specialization track in Business Information Management from the ‘Katholieke Universiteit Leuven’ (Brussels Campus). This group totaled seventy-two students previously educated in general business and economics subjects other than the Information Technology field. Their limited experience in the domain of software modeling and requirements engineering was suitable in order to test the level of general comprehension and learning transferability of the Rationale Tree (compared to the User Story Mapping) as represented by the quality of the artifacts produced by the students when exercised in a specific case study with a defined user stories’ set.

3 Design of the Modeling Exercise

The exercise began by introducing a mix of open-ended and closed questions intended to: a) collect background information concerning the participating students, and b) survey their previous knowledge corresponding to the domain of software modeling. For this introductory part, the students had to write down their primary occupation at the time the exercise was taking place and their previous experience with software modeling. This last question asked the students to elaborate on the specific modeling languages they were accustomed to and how they acquired this kind of knowledge. Following, the format of the questions changed from open-ended to a Likert style as they became more specific to the structure of user stories, the Rationale Tree method and the User Story Mapping technique. In particular, students were asked to what extent they were aware of these techniques (in general) and to what extent they were accustomed to specific features incorporated in these techniques (i.e., ‘are you familiar with the i* modeling language?’). Due to size limitations, all the information collected from these questions is presented in Appendix 1².

Next, the entire group was split randomly into two equally-sized smaller groups. The first group³, amounting to thirty-six students, had to employ the Rationale Tree method throughout the entire endeavor. The exercise began by introducing a detailed theoretical explanation of the method accompanied by a solved example to instigate further the students’ understanding. Next, each student was given the same case description (‘Company X’ case, see Fig. 1) and had to identify/correspond the user stories’ dimensions to the modeling elements of the Rationale Tree. Following, each student individually had to draw a *RT diagram*⁴ based on the information provided from the previously determined user stories’ set. The second student group⁵, comprised also of thirty-six students, had to follow exactly the same procedure to perform the modeling exercise; they were, however, instructed to employ the User Story Mapping technique throughout

² All Appendices are available at: <http://dx.doi.org/10.17632/b4zyddbrbc.1>.

³ The first student group will also be referred to as the ‘Rationale Tree group’.

⁴ ‘RT diagram’ denotes the artifact produced by employing the Rationale Tree method.

⁵ The second student group will also be referred to as the ‘User Story Mapping group’.

their entire endeavor. In this aspect, this particular group was provided with a theoretical explanation about the User Story Mapping technique accompanied by a solved example with the ultimate goal of each student producing a *USM*⁶. A full description of the tasks and expectations involved in this drawing exercise is given in Appendix 2. Both groups were given the same time to perform this drawing exercise (one hour). An experimental design in the aforementioned form is justified in [5] as it is purposed to analyze the students' ability to respond to complex software engineering cases given that their viewpoint provides an alternative conceptual framework, a wider scope and a rich set of data.

The case study used for this modeling exercise was retrieved from [6] and referred to the case provided by a 'Company X' remained unnamed for privacy reasons. This particular case study (presented in Fig. 1) was reviewed and optimized using the Quality User Story framework [2,6]. The latter represents a linguistic approach to engage in a first-level evaluation of the quality in a set of user stories. Overall, the framework establishes a set of fourteen criteria addressing the syntactic, the semantic, and the pragmatic content of the text that should define high-quality user stories [2]. Therefore, a quality-enhanced user story set, after the implementation of the aforementioned framework, was used in the present exercise in order to build a 'golden standard' for the RT diagram. The 'golden standard' method refers to the construction of an 'ideal' artifact created by our research team. Every correct element and link was awarded with points. The RT diagrams produced by the Rationale Tree group approaching this 'ideal'

Company X has developed a web application where users can analyze DNA samples. If End Users want to make use of the analysis's platform, they need to sign into the web application. Before signing in, they need to complete a registration process via a personal activation code, generated by the partners. Because the information used and shared on the platform is of high confidentiality, the End User has the choice to enable a two-factor authentication during the registration. If enabled, every time an End User wants to log in to the platform, he/she receives a two-factor authentication.

1. As an End User I want to change my personal information So that I can correct errors in my personal information.
2. As an End User I want to receive an activation code So that only known customers can register.
3. As an End User I want to register myself So that I can sign in with a user account.
4. As an End User I want to register an account with a valid activation code of a customer, my personal information and a secret password.
5. As an End User that is unauthenticated, I want to read and accept the terms and conditions upon registration of an account So that I will act in compliance with the regulatory requirements.
6. As an End User I want to be able to log in with a two-factor authentication so that my log in is more secure.
7. As an End User I want the application to ask for my authentication So that only authenticated End Users can sign into the platform.

Fig. 1. The 'Company X' case study for the drawing exercise.

⁶ The 'USM' abbreviation denotes the artifact of the User Story Mapping technique.

solution the most, were awarded with the most points (see Sect. 5). This golden standard was then ‘converted’ into a USM to compare and evaluate optimally the artifacts received by the User Story Mapping group also.

4 Hypothesis

As justified before, the comparison between the different features of the Rationale Tree and the User Story Mapping started by providing a thorough theoretical explanation of each approach before the actual drawing exercise takes place. Whether the notions were sufficiently understood was measured by the quality of the artifacts extrapolated from the answers of the students. Conversely to the User Story Mapping technique, the Rationale Tree contains links and decompositions making it more difficult to be comprehended by novice modelers. The latter incorporates also a broader choice of elements with semantics open to interpretation, adding thusly to its overall complexity. Hence, *our main hypothesis*:

‘The artifacts resulting from the drawing exercise are of lower quality in the Rationale Tree group than the User Story Mapping group.’

The quality in the artifacts produced by these two student groups represents the level of comprehension of the requirements problem, as presented in the case study and its accompanying user stories’ set, when assisted by the employment of the two aforementioned user story structuring techniques.

5 Evaluation Methods for the Students’ Artifacts and Validation of the Hypothesis

This section is devoted to the evaluation and comparison of the students’ drawn artifacts. We used two evaluation methods to assess the **RT diagrams**: The first one (the *Three Criteria* evaluation method [11]) consisted of the evaluation of three main criteria namely completeness, conformity and accuracy. A score was allocated to each model element and link identified in each of the RT diagrams reproduced by the respondents. *Completeness* checked whether the respondents have modeled all WHAT and WHY dimensions in a complete way, receiving a point for each modeled element. *Conformity* referred to the respondents ability to identify which modeling construct would represent each user story dimension. For each conformed element half a point was given. *Accuracy* referred to the correct identification of a number of fundamental links as identified from the golden standard artifact. For each correctly identified link, four points were given. If the link was present but the wrong type of link was used, only one point was given. The second evaluation method (the *Golden Standard* [6]) was based on an ‘ideal’ artifact, created by the research team, whose every correct element and link was awarded with the maximum of points. The RT diagrams provided by the students approaching this ‘ideal’ solution the most gathered the most points. The **USM**

artifacts were evaluated based on five identified criteria namely completeness, consistency, accuracy, correctness and accuracy. Appendix 3 provides a detailed description of each evaluation method as well as the grading system per criterion.

5.1 How Well Can the Rationale Tree Group Identify Correctly the WHO-, WHAT-, WHY-Dimension of the User Stories' Set?

Before commencing with the actual evaluation of the produced artifacts for the two groups, we analyzed whether the identification of the WHO-, WHAT-, and WHY-dimensions within the user stories' set was easily feasible for the participating students. We started with that because a proper identification of the user stories' dimensions is not an easy task for novice modelers.

Judging from the students' choices, the **WHO-dimension** was identified correctly by 73% of the students assigned to the Rationale Tree method. However, this particular group had difficulty in identifying elements of the **WHAT-** and **WHY-dimension**. Indeed, Table 1 shows the participants' augmented difficulty in distinguishing between *Task* and *Capability* on the one hand, and *Soft-* and *Hard-Goal* on the other. We therefore consider that the differentiation between the former mentioned elements can be ambiguous for novice modelers when using the Rationale Tree method.

Table 1. WHAT-, WHY-dimension analysis for Rationale Tree group.

MBA student	Task	Capability	Soft-goal	Hard-goal	Not present
US2WHAT	45.9	48.6	5.4		
US2WHY	2.7	5.4	51.4	40.5	
US3WHAT	94.6	5.4			
US3WHY	21.6	16.2	56.8	5.4	
US4WHAT	29.7	70.3			
US4WHY					100
US5WHAT	59.5	40.5			
US5WHY	10.8		43.2	45.9	
US6WHAT	16.2	81.1	2.7		
US6WHY			13.5	86.5	
US7WHAT	62.2	35.1	2.7		
US7WHY	2.7	5.4	54.1	37.8	

***Bold elements** indicate the most occurring element within that sample.

5.2 How Well Can the User Story Mapping Group Identify Correctly the WHO-, WHAT-, WHY-Dimension of the User Stories' Set?

Judging from the students' choices, the **WHO-dimension** was identified correctly by 62.9% of the students assigned to the User Story Mapping technique. Table 2 reveals that for this particular group there was also a lot of variability in the identification of elements which points towards the existence of ambiguity in the artifacts' semantics. Students tend to identify the WHY-dimension solely as an *Activity*, which is the highest level used element in the USM. Goal-level tasks are identified solely as activities, although this depends on the context of each case. Our used case (Company X) did not have much details, and due to the hierarchical structure of the User Story Mapping technique, an activity must be followed by a task and finally a detail. This unlike the Rationale Tree where goals, tasks and capabilities can all be used interchangeably.

Table 2. WHAT-, WHY-dimension analysis for User Story Mapping group.

MBA student	Activity	Task	Detail	Not present
US2WHAT	5.7	57.1	37.1	
US2WHY	51.4	11.4	34.3	2.9
US3WHAT	20	77.1	2.9	
US3WHY	71.4	25.7		2.9
US4WHAT		37.1	62.9	
US4WHY	2.9	2.9		94.3
US5WHAT	2.9	74.3	22.9	
US5WHY	62.9	8.6	25.7	2.9
US6WHAT	8.6	74.3	17.1	
US6WHY	71.4	5.7	20	
US7WHAT	5.7	74.3	17.1	2.9
US7WHY	51.4	14.3	25.7	8.6

***Bold elements** indicate the most occurring element within that sample.

5.3 Comparison of the Resulting RT Diagrams and USM Artifacts for the ‘Company X’ Case

Table 3 compares various descriptive elements related to the acquired points for each drawn artifact, based on their corresponded evaluation method. We observe that the means are similar for the artifacts produced by both groups. Standard deviation is lower for the resulting USM. Hence, the artifacts resulting from the modeling exercise were similar in quality regardless to which group the students were assigned to, which is an element that leads us towards the rejection of our *main hypothesis*.

Table 3. Descriptive statistics of three evaluation methods.

		% of points on USM	% of points on RT diagram based on Golden Standard evaluation	% of points on RT diagram based on Three Criteria evaluation
N	Valid	35	37	37
	Missing	37	35	35
Mean		.5281	.5536	.5522
Median		.5690	.5806	.5965
Std. Deviation		.13698	.20109	.17229
Minimum		.14	.03	.11
Maximum		.71	1.29	.83

6 Conclusion

The present study illustrates a first-level comparison between a relatively-new and conceptual modeling-based method of structuring user stories' sets to the industry-adopted technique. In total, the elements/links presented in the Rationale Tree method were perceived as somehow difficult to identify and differentiate one from another. Our analysis explicitly revealed a prolonged confusion between the elements used to describe the WHAT- and WHY-dimension within a user stories' set such as Tasks/Capabilities and Soft-Goals/Hard-Goals. This is an indication that more thought should be placed in the semantics and rules to facilitate the differentiation between such functional elements. However, despite our initial expectations, the Rationale Tree group produced artifacts similar in quality compared to the ones of the User Story Mapping technique. A factor that might have influenced these results is the use of the particular case study which proved to be difficult in representing the requirements problem no matter the method used by the students. However, the use of this case provided us with the necessary consistency in studying and comparing our results with the ones received from previous studies. Moreover, the complexity of the case adds value to the goal of this exercise which was to be considered as a first-level simulation of the elaborate tasks that modelers have to perform when taking on the task of requirements elicitation and analysis. An even more in-depth comparison of the ability of understanding the software problem and, more particularly, specific key aspects like missing requirements, Epics and Themes on the basis of both techniques will be the subject of a broader communication. Future work also includes reproducing the same kind of experiment but opposing the rationale tree and the use case diagram [7].

References

1. Cohn, M.: User Stories Applied: For Agile Software Development. Addison-Wesley Professional, Boston (2004)

2. Lucassen, G., Dalpiaz, F., van der Werf, J.M.E.M., Brinkkemper, S.: Improving agile requirements: the quality user story framework and tool. *Requirements Eng.* **21**(3), 383–403 (2016). <https://doi.org/10.1007/s00766-016-0250-x>
3. Patton, J., Economy, P.: *User Story Mapping: Discover the Whole Story, Build the Right Product*. O'Reilly Media, Inc., Sebastopol (2014)
4. Saunders, M., Lewis, P., Thornhill, A.: *Research Methods for Business Students*. Pearson Education, London (2016)
5. Teddlie, C., Tashakkori, A.: *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*. Sage, Thousand Oaks (2009)
6. Wautelet, Y., Gielis, D., Poelmans, S., Heng, S.: Evaluating the impact of user stories quality on the ability to understand and structure requirements. In: Gordijn, J., Guédria, W., Proper, H.A. (eds.) *PoEM 2019. LNBIP*, vol. 369, pp. 3–19. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-35151-9_1
7. Wautelet, Y., Heng, S., Hintea, D., Kolp, M., Poelmans, S.: Bridging user story sets with the use case model. In: Link, S., Trujillo, J.C. (eds.) *ER 2016. LNCS*, vol. 9975, pp. 127–138. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-47717-6_11
8. Wautelet, Y., Heng, S., Kiv, S., Kolp, M.: User-story driven development of multi-agent systems: a process fragment for agile methods. *Comput. Lang. Syst. Struct.* **50**, 159–176 (2017)
9. Wautelet, Y., Heng, S., Kolp, M., Mirbel, I.: Unifying and extending user story models. In: Jarke, M., et al. (eds.) *CAiSE 2014. LNCS*, vol. 8484, pp. 211–225. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-07881-6_15
10. Wautelet, Y., Heng, S., Kolp, M., Mirbel, I., Poelmans, S.: Building a rationale diagram for evaluating user story sets. In: *2016 IEEE Tenth International Conference on Research Challenges in Information Science (RCIS)*, pp. 1–12. IEEE (2016)
11. Wautelet, Y., Velghe, M., Heng, S., Poelmans, S., Kolp, M.: On modelers ability to build a visual diagram from a user story set: a goal-oriented approach. In: Kamsties, E., Horkoff, J., Dalpiaz, F. (eds.) *REFSQ 2018. LNCS*, vol. 10753, pp. 209–226. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-77243-1_13
12. Yu, E., Giorgini, P., Maiden, N., Mylopoulos, J.: *Social Modeling for Requirements Engineering*. MIT Press, Cambridge (2011)