A LifeLike and Groupe Bizness experience to train user requirements elicitation skills

Abstract: User Requirements Elicitation (URE) is a critical stage in the development of software systems. Since it is communicative and interdisciplinary in nature, the process of URE implies a combination of soft skills and an active listening attitude, which can be trained only with the practice. In fact, despite a variety of techniques and approaches to URE are employed, and different methodologies are suggested to choose the best technique for the specific situation, there is not at the moment a systematic training method. In the paper, a behavioural simulator reproducing a lifelike URE conversation is presented, which was developed exactly to train URE skills. The effectiveness of the simulator was verified through an experiment, whose design, implementation and results are described.

Keywords: User Requirements Elicitation, requirements engineering, online applications, behavioural simulator, soft skills training

1. Introduction

The first step when designing an online application is to define the information need it wants to fulfill and the services it is expected to provide. Contents and functionalities have to be implemented according to the typology of users the online application is mainly intended to reach. In order to use and enjoy it, in fact, users require the application to be able to guess and satisfy their requests. The website of an enterprise devoted to baby care, for instance, will probably be visited mostly by parents of young children, parents-to-be, and people somehow related to the previous two groups; therefore, to meet these users’ expectations, the website must provide contents related to infants’ needs like products available, location of stores, information about different stages of children’s growth, and functionalities like buying online or having the complete view of expensive products (e.g. car seats). The process of defining users’ requirements for online applications and, in general, for every type of software system, is called User Requirements Elicitation (URE). URE is one stage of the requirements engineering process, which comprises other activities such as requirements prioritization and operationalization, and it is, according to many, the most important stage for the project success (Hofmann and Lehner, 2001; Hickey and Davis, 2002). The term “elicitation” points out the delicate role of the analyst, who has to take an active listening attitude in the dialogue with system stakeholders and intended users, being able to seek, uncover and elaborate requirements. Literature about requirements engineering, however, does not provide a uniform presentation of the steps involved, not even a shared definition of the activity itself. It is recognized, though, that URE is about learning and understanding the needs of users and the intentions of clients in developing the system, with an important role being played by invention and creativity (Robertson and Robertson, 1999; Maiden, Gizikis and Robertson, 2004). The success of the process largely depends on the analyst’s communication skills and expertise, since URE is a) communicative, b) interdisciplinary and c) practical in nature, as explained below.

a) It can be seen as a dialogue among three actors: the client, whose intentions and expectations for the software have to be precisely clarified and made explicit; the intended users, whose characteristics have to be defined in order to understand their requirements; the analyst, who needs to collect the more elements as possible to realize the client’s desires and fulfill the users’ needs.

b) URE is performed in a variety of settings, from the development of websites and mobile applications to the design of complex piece of software, from the implementation of enterprise systems to the development of market product lines. Different techniques are employed and approaches adopted, depending on the specific context of the project; such techniques and approaches have been borrowed and adapted from different disciplines, such as the social sciences (e.g. communication sciences, marketing), organizational theory, knowledge engineering, group dynamics. Only a few of them, though, have been developed specifically for URE (Zowghi and Coulin, 2005).

c) URE is an early but critical stage in the development of software systems, since in many cases, apart from a set of shared fundamental goals, functions, and constraints for the system or an explanation of the problems to be solved, contents and functionalities have to be discussed, clarified, even identified. Stakeholders, then, might
have different positions, or might be blind towards real users’ needs or technical abilities. Most of the times, then, stakeholders and analysts come from widely different professional areas, and do not share enough common understanding of concepts and terms; it is also often the case that the analyst have not enough familiarity with the problem, or that stakeholders do not realize what is actually feasible or realistic. All these drawbacks only emerge during the practice, that is in the actual system engineering, and require practical solutions.

As a communicative activity, URE is influenced by the multifaceted, unexpected, unforeseeable variables occurring in human interaction, and it is, thus, subject to a large degree of error. The establishment of a collaborative and positive attitude among the participants is essential to reduce incorrect, incomplete, inconsistent collection of information, and to overtake misunderstandings and misalignments between the parties involved. Despite the many techniques and approaches developed to perform URE, as well as classifications to organize such plethora of techniques and help analysts to select the most suitable one for their case, the training of novices still largely relies on practice (learning by doing) and experts’ emulation (learning by observing). The need to work towards reducing the gap between experts and novices has been widely denounced, but “requirements elicitation still remains more of an art than a science” (Zowghi and Coulin, 2005).

In this paper, an experiment is reported, which was designed to validate the effectiveness in training URE skills of a behavioural simulator, exactly developed for this task. The simulator derived from a joined idea between an academic stakeholder (i.e. the authors of this paper), who teaches and makes research in the area of online communication, and an industrial stakeholder, who is expert in the design, construction and validation of behavioural simulators to train soft skills in different working settings.

In what follows, we first set the state of the art of current techniques and approaches to URE, then we present the simulator, we describe the experiment and discuss the results.

2. A framework to understand and analyze online communication

2.1 Techniques, methodologies and models of URE

A wide variety of techniques, approaches and tools have been used for requirements elicitation. Zowghi and Coulin (2005) selected a core group of eight of them, which they claimed are representative of those that are both state of the art and state of practice. These eight ‘families’ of techniques are, briefly, the following: interviews, which are best applied when there is a limited understanding of the domain from the part of the analyst; domain analysis, where related documentation and applications/competitor systems are examined; group work, implying a direct commitment of the stakeholders and cooperation among them; ethnography, which is particularly effective to investigate intended uses of the system and the addressed public; prototyping, which may be expensive but is extremely useful when entirely new applications have to be developed; goal based approaches, in which high-level goals of the system are decomposed and elaborated into sub-goals; scenarios, that are narratives describing expected interactions between users and the system; viewpoints, where the domain is modeled from different perspectives (such as from its operation, its interface, its competitors, its users) in order to have a complete description of the system.

Usually, more than one technique or approach is adopted, because of the multifaceted and iterative very nature of the URE process (see previous section). Methodologies are reported in the literature, which propose combinations of approaches and techniques to achieve the best possible results in specific situations and environments (Checkland and Scholes, 1990; Goguen and Linde, 1993; Martin, 2003).

Also process models have been proposed over the years to describe the different stages of the URE and, thus, to guide the selection of the techniques to be used (Sommerville and Sawyer, 1997; Hickey and Davis, 2002). However, these models only sketch generic roadmaps of the process; their inability “to provide definitive guidelines is a result of the wide range of tasks that may be performed during requirement elicitation, and the sequence of those activities being dependent on specific project circumstances” (Zowghi and Coulin, 2005: 23).

When committed to develop online communication projects, our research team also employs a mix of techniques to perform URE, including interviews, meetings and focus groups with stakeholders, user scenarios. An original methodology developed by the team exactly for URE, is called User Requirements with Lego (URL) (Cantoni et al. 2009a; Cantoni et al. 2009b). URL is based on Lego Serious Play, “an experiential process
designed to enhance innovation and business performance” (Lego, n.d.), by ‘giving your brain a hand’; in fact, the core idea is doing while thinking, in order to stimulate and enhance creativity. URL is a sort of extension of Lego Serious Play, designed to support the definition of Information Architecture and content strategies in online communication. In particular, URL helps in finding tacit, difficult to grasp communicative requirements that usually do not emerge with other techniques. For this reason, URL has to be intended as an additional methodology, used besides formal and structured strategies (such as interviews and focus groups) to uncover and define user requirements.

The development of the URL methodology as well as the construction of the behavioural simulator presented in this paper, were driven by the Online Communication Model (OCM), a model which ideally represents all the components of online communication artifacts – hereafter referred to with the generic term ‘online applications’ – like websites or mobile applications, and constitutes the framework of our understanding of URE. Differently from other models used to map URE, which are mostly based on processes and domain knowledge, OCM looks at computer-based systems from the point of view of communication, adopting a holistic approach.

2.2 Online Communication Model

OCM (Cantoni and Tardini, 2006; 2010) goes beyond a naive dichotomy that sees online applications either as mere technological artefacts, to be handled by engineers or, on the opposite side, as advertising tools, to be managed by visual communication experts. The one or the other interpretation is only partially true, but both suggest that online applications are static objects. The model (Figure 1), instead, presents them as dynamic entities with a proper life and typical activities, like a shop or a press agency. OCM ideally considers all the elements and the actors involved in the communicative activity taking place online, and groups them in four dimensions or pillars:

1. contents and services/functionality: the more or less structured ensemble of information pieces and services provided in the application, such as information provision, news reporting, buying, chatting, product or service reviewing;
2. accessibility tools: the collection of technical instruments, which make the contents and services accessible, like hardware, software, and interface;
3. people who manage, who are the group of people who design, implement, maintain and promote the application;
4. users/clients, who are the group of people who access and use the application.

The first two dimensions are related to ‘things’, while the other two are related to ‘people’. There is, then, a fifth dimension that completes the framework:

5. ecological context or relevant info-market, which gives to every element of the application its precise meaning, value and place within the broader context of the web.

Figure 1: Online Communication Model (adapted from Cantoni & Tardini, 2006).

3. Behavioural simulators to train soft skills
The concept

Simulators are well known in education. Usually they are employed to train hard skills, that are specific teachable abilities which can be defined and measured, like typing, accounting, using a software. Simulators are also very popular to train procedural knowledge, that is how to do things, like driving a car, flying a plane, operating a patient (see, for instance, the wide literature on surgical simulators or machine drive simulators). On the contrary, as far as we know, there aren’t simulators for training soft skills, which are personality-driven abilities, related to the emotive and communicative sphere, like patience, interpersonal relation, teamwork, communicative attitude. Soft skills can make a difference in the way a person performs in certain activities and at work, but are difficult to be defined, taught and, thus, measured.

The simulator validated in our experiment was designed to train URE skills, which comprise a number of soft skills, many of them related to people’s communicative attitude, like negotiation style, self-control, empathy, complaints management. It was realized thanks to a joined idea of eLearning Lab, the laboratory of Università della Svizzera italiana in charge of improving the quality of teaching through the use of ICTs, and LifeLike Interaction, an international enterprise which designs and realizes behavioural simulators to improve job outcomes that rely in the interaction among people (LifeLike, n.d.a). LifeLike simulators are based on the idea that “our brain does not accumulate data, rather, it memorizes experiences in the form of stories” (LifeLike, n.d.b); different types of behavioural simulators have been developed by LifeLike for specific tasks, like to enhance customers’ needs understanding, team management, communication effectiveness, sales closing. In the next section, the simulator developed to train URE skills is presented.

How it works

The didactical idea behind the simulator is an interaction between user and game, based on a narrative and relational model developed by LifeLike (the LifeLike Interaction®). The user plays an interactive game-interview, in which s/he plays the role of an online communication consultant, who meets a client willing to create a website for her enterprise. The game experience is highly realistic and based on emotion. In order for the dialogue to be as much realistic as possible, the situation has to be credible and the characters must have a professional as well as a psychological and private profile. The client, in our case, is Mrs Manuela Cristicchi, the manager of a family-driven hotel on the Adriatic coast (Italy), who wants a website for her hotel, because she knows that her competitors also have a website and understands the potentiality of ICT as new promotion channels. The character of Manuela is played by a professional actor and the video sequences are built with advanced cinematographic techniques. All LifeLike simulators are built following the same technique. Actors are required a fine performance, since reading the script is not enough: they have to interpret a number of different moods and attitudes, which are communicated mostly non-verbally, with kinesics, facial expression, tone of voice (see Figures 1 and 2). Real-time is mandatory, given that time has a real impact on the evolution of the interview and the reaction of the client; in fact, there is no pause button. Starting from a ‘dream dialogue’, that is the perfect dialogue between an analyst and a client, every dialogical turn was modified so to include a potential problem or an error in the URE process. The interview the user plays is the result of a unique combination of modified dialogical turns, according to his/her choices during the meeting with the client. About 16 millions combinations are possible. The dream dialogue was elaborated by eLab, who has expertise in the field of online communication, URE and usability in particular; the dream dialogue was then modified by LifeLike according to its script rules.

The game is articulated in the following steps:
1. the user logs into the system and is introduced into the simulator with a short video; s/he also receives a short description of Manuela Cristicchi’s professional and personal profile;
2. the user plays the game: at each dialogical turn, s/he has to choose among three to five alternative statements, which might be either questions for or answers to Mrs Cristicchi (see Figure 3);
3. at the end of the game, the user gets: a score in percentage representing his/her performance in the URE dialogue according to the four stages of the game (opening, requirements analysis, solution proposed, closing, see (Figure 4), an analytical feed-back on different stages of his/her dialogue with the possibility to check the game sequence (Figure 5), a comment by the Manuela, in the form of a telephone call she makes to her son and during which she reports her impression of the meeting and her decision;
4. the user is invited to play again to improve his/her performance.

The feed-back is based on the following indicators, weighted by the system to give the overall game score (see Figure 6):

- **negotiation style**: refers to the user’s spontaneous approach to the task. It can be assertive, aggressive or appeasing;
- **process of communication**: indicates the extent to which the user was able to follow the ideal dialogical path, managing difficulties and unexpected situations;
- **quality of relation**: refers to the atmosphere created by the user during the dialogue, that is the extent to which s/he was able to put the client in a comfortable state;
- **focus on the other part**: is about the ability of the user at taking the right decision and accomplish the task without wasting time and disappointing the interlocutor;
- **self control**: refers to the perceived consistency of the user’s behaviour.

At the end of the game, before the simulator gives its feedback, the user is asked to evaluate him/herself, through a set of questions related to the above indicators, to which s/he has to assign a score in percentage. Picture 6 reports the comparison between the scores assigned by the user to him/herself and the scores assigned by the system.

Like a muscle, soft skills need continuous and systematic training. As one single training session is not enough to strengthen a muscle, in the same way, one single game is not enough to enhance the desired skills. LifeLike Interaction fine tuned a training protocol for behavioural simulators to be effective, which requires the trainees to play a certain number of games over one year, and their performances being monitored thanks to periodical...
check-ups.

Figure 6: Game performance according to the system indicators and comparison between the user’s self-evaluation and the system evaluation.

4. Validation of the LifeLike simulator to train URE skills

Research questions

The experiment was driven by two research questions:

a) Do users playing systematically with the LifeLike simulator improve their performance with the simulator?
b) Do users playing systematically with the LifeLike simulator enhance their user requirements elicitation skills?

Question a) refers to the ability of players to reach better results practicing with the simulator, that is if they are able to understand the underlying rationale of the simulator and learn from it to improve their scores. If a selected group of trainees do not improve its performance with a tool/game by systematically training with it, this means either that trainees have some kind of learning difficulty – due, for instance, to a wrong estimate of the knowledge level required – or that the tool/game presents design failures – like unclear rules or inconsistent rationale.

Question b), on the other side, represents the core of the experiment, as well as the success or failure of the simulator to accomplish the task it was developed for.

The experiment was conducted with a class of 69 students, attending the second year of a Bachelor program in Communication Sciences, along the second semester of the academic year (17th February 2014 – 30th May 2014). The class was divided into two groups, the experimental and the control group, which trained with the simulator according to the schedule in Figure 7. Students were randomly assigned to one or the other group, paying attention to have a balance between men and women. The experimental protocol followed five main steps described below. Figure 7 is a representation of the experimental procedure for the two groups.
Experimental protocol

1) 1st check-up: both groups played 4 games over a 2 hours class, after an overview of the experiment and a brief introduction to the simulator. Students were assumed to have the same level of specific knowledge (i.e. related to URE and online communication) and to have no experience with the practice of URE. Goal of the first check-up was to measure the individual as well as the group starting performance.

2) 2nd check-up: both groups played 4 games over one week (the game location being not relevant), after attending classes on URE and receiving details about the design, features and use of the simulator. Students were taught about different techniques and approaches to URE, case studies were discussed, and the OCM was presented as theoretical framework which drove the design of the simulator and constitutes the point of reference for the URE practice. Goal of the second check-up was to measure the individual as well as the group improvement in playing with the simulator, after receiving field-specific and tool-related knowledge. After the second check-up, students were split in groups. The experimental group was required to freely train with the simulator over 4 weeks (from 14th April to 5th May), playing a minimum of 20 games. After that period, accounts of students belonging to the experimental group were blocked.

3) 1st Listening test: after the training period, on May 6th both groups underwent a ‘listening test’, that is a test specifically designed to measure if students’ URE skills benefited from the training. The name recalls the core of the elicitation activity, which implies the analyst to actively listen to his/her client, in order to uncover and clarify the requirements of the online application at stake. The test consisted in a video (see picture 1) showing a lifelike dialogue between an analyst and a client willing to re-design the website for her luxury cruises enterprise (see Figure 8). Students had to carefully look – and listen! – at the video and annotate on a pre-set scheme three categories of elements:
a) what they learnt from the dialogue in terms of the 5 elements of the Online Communication Model (e.g. contents and functionalities of the website, people internal or external to the enterprise devoted to manage and produce it); this part counted 3 out of 10 points;
b) what else they need to know to successfully complete the URE, but that did not emerge from the dialogue, again in terms of the 5 elements of the OCM; this part counted 3 out of ten points;
c) communication problems between the client and the analyst, listed according to three typologies, that are problems regarding the speakers, problems regarding the world and its representation (i.e. different perceptions and priorities), and the time factor (i.e. how the time was managed) (see Cantoni and Piccini, 2004); this part counted 4 out of ten points.

4) 3rd check-up: the day after the listening test (May 7th), a final check-up was made, in order to test if the
experimental group improved its scores thanks to the systematic training with the simulator. From May 8th until
the course exam held on June 11th, the control group was given the possibility to freely train with the simulator,
in order for those students to reach the same level of preparation for the exam as the students of the other group.
Playing with the simulator did not directly grant students a higher score in the exam, but was supposed (as the
experiment in fact later showed) to support their preparation to that part of the exam devoted to URE.

5) 2nd listening test: one part of the exam, held on 11th June 2014, was constituted by a second listening test,
aimed at testing if students improved in understanding URE, thanks to the knowledge acquired during the
course and to the activities proposed. An ad hoc lifelike video (see Figure 9) was shot for the second listening
test, showing a URE conversation between an analyst and the communication manager of a cultural association
devoted to restoration of pieces of art in Venice (Italy), who wished to develop a website to let interested people
know about the association activities and to fund-raising.

It has to be noted that students had access to the feedback part of the game only during the systematic training,
while the check-up games did not give them any feedback. This was to avoid that the control group could learn
from its mistakes by listening to the feedback.

Figure 8: a frame of the first listening test. Figure 9: a frame of the second listening test.

5. Results

Training with the LifeLike simulator and improvement in game performance

At the first check-up, the two groups performed almost the same (median = 51% for the c.g.; median = 49% for
the exp.g.), meaning that they had the same starting level of competence. After in-class lectures on URE and a
familiarization with behavioural simulators, they both improved a bit their performance: the median of the
control group increased of 2,5% and that of the experimental group increased of 2%. This shows that
knowledge of the domain as well as knowledge of the tool have an influence in the game performance. At the
third check-up, the experimental group which had played 3 weeks with the simulator, dramatically improved its
performance, moving from 51% to 59,5%, that is gaining 8,5%. The control group did not play and had a gain
of only 3% (the median moved from 53,5% to 56,5%).

Mean values followed the same evolution of the median values with very little difference, meaning that the
samples were normally distributed.
Training with the LifeLike simulator and improvement in URE

Table 1 reports both the overall results (mean and median) of the two groups’ performances in the first and the second listening test, and the values they obtained in each of the three parts of the test.

Table 1: Results of the experimental and control group at the 1st and the 2nd listening test.

<table>
<thead>
<tr>
<th></th>
<th>I Listening Test</th>
<th>II Listening Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Experimental Group</td>
<td>Control Group</td>
</tr>
<tr>
<td>Mean</td>
<td>5,98</td>
<td>5,04</td>
</tr>
<tr>
<td>Median</td>
<td>6,60</td>
<td>4,90</td>
</tr>
<tr>
<td>Mean part I</td>
<td>3,11</td>
<td>2,64</td>
</tr>
<tr>
<td>Median part I</td>
<td>3,40</td>
<td>2,50</td>
</tr>
<tr>
<td>Mean part II</td>
<td>2,86</td>
<td>2,36</td>
</tr>
<tr>
<td>Median part II</td>
<td>3,10</td>
<td>2,40</td>
</tr>
<tr>
<td>Mean part Ia</td>
<td>1,67</td>
<td>1,57</td>
</tr>
<tr>
<td>Mean part Ib</td>
<td>1,44</td>
<td>1,08</td>
</tr>
<tr>
<td>Median part Ia</td>
<td>1,80</td>
<td>1,60</td>
</tr>
<tr>
<td>Median part Ib</td>
<td>1,70</td>
<td>0,80</td>
</tr>
</tbody>
</table>

The experimental group performed better than the control group in the first listening test, with a mean advantage of nearly one out of ten points (5,98 for the exp. g. against 5,04 for the c.g.). If the median is considered, the performance is even higher: the value separating the higher half of the experimental group is 6,6 against 4,9 of the control group.

The experimental group performed better in each of the three parts of the test, in particular, the median values of the second part (what else they need to know) show a relevant difference of nearly one point out of four (exp. g. 1,79 against c.g. 0,80); probably, training with the simulator, users became more sensitive to the need of collecting all the relevant information to develop a successful application and, conversely, more alert in detecting the missing information.
Results of the second listening test show a low improvement for the experimental group – both overall and in the single parts – but a dramatic improvement of the control group performances. After the systematic training with the simulator, the mean scores of the control group raised 1.5 points out of ten; the median reaches 6.47 against the previous 5.04. If the single parts are taken into account, users’ performances improved especially in the first two parts (what they learnt and what else they need to know) than in the third part (communication problems), confirming the above observation.

6. Conclusion

Behavioural simulators are a new type of systems developed to train soft skills, that are personality-driven abilities related to the emotive and communicative sphere. In the paper, the design process, the development and the first results of an experiment aimed at validating a behaviour simulator to train user requirements elicitation skills were presented. The experiment involved students of a second Bachelor year attending a course on Online Communication. Both their performance in playing with the simulator and their ability in the practice of URE were measured, in order to verify the effectiveness of the simulator. Results showed that the median values of the experimental group – after a period of systematic training – improved of 8.5% against an improvement of 3% of the control group. Results of the listening tests, which were designed to test students’ URE skills out of the simulator environment, also showed an improvement in the performance, arguably given by the systematic training with the simulator. In particular, users performed better in the parts of the test related to the requirements identification (i.e. what they learnt from the client’s words) and to the unsatisfied requirements need (i.e. what else they need to know that did not come out of the dialogue with the client).

Further analysis should be performed to test the correlation between performance and users’ gender. A second experimental round, then, should be designed, in order to measure the correlation between users’ performance and number of games played.

7. References