A Preliminary Investigation of the Utility of Goal Model Construction

Naomi Cebula, Lily Diao, and Alicia M. Grubb

Department of Computer Science
Smith College, Northampton, MA, USA
{ncebula, ldiao, amgrubb}@smith.edu

Abstract. Goal models have long been used in the literature to model and reason about stakeholders' intentions. Prior work proposed several studies aimed at investigating what utility stakeholders derive from constructing and analyzing goal models. We designed and conducted an initial empirical study that explores the construction stage of goal modeling, asking whether stakeholders benefit from manually drawing their own model. We recruited eight qualified participants and asked each to create a goal model for a decision they were considering while talking out loud. Half of the participants in this study used BloomingLeaf, while the remaining participants drew goal models by hand. Using information gathered through an online pre-study questionnaire, we constructed a goal model for each participant. Participants' compared our generated models with their manually created ones. We used open coding to find themes and categories for qualitative responses. Analysis was mixed on whether participants preferred their own model to the researcher generated one. The results of which have implications for future studies, as well as goal model adoption and automation.

1 Introduction

In early project planning, goal-oriented requirements engineering (GORE) approaches have been advocated to help stakeholders make trade-off decisions. While there have been many GORE languages [1], they all encapsulate the representation of functional and non-functional requirements into a central artifact, called a *goal model*. These techniques have been extended and applied to research projects downstream in software development [2], yet they have not achieved broad industrial adoption [3]. In prior work, Grubb described basic assumptions made as part of the GORE process, including the assumption that modelers can construct goal models [4]. Using these assumptions, Grubb proposed several studies aimed at investigating what utility stakeholders derive from constructing and analyzing goal models, where *utility* was defined as 'fitness for some desirable purpose or valuable end' [5].

Our aim is to investigate the construction stage of goal modeling. This study is a first attempt at exploring the questions proposed by Grubb [4]. We ask *to what extent do participants gain value in manually drawing goal models (in a tool or on paper)?* Utility can be more specifically seen as whether stakeholders

have gained any new insights, changed any decisions, or developed new values through modeling. Our *null hypothesis* in this work is that there is no utility in the goal model construction process (i.e., drawing models on paper or in a tool).

Contributions. In this paper, we present an exploratory study of the utility of goal model construction. We recruited, trained, and observed eight novice Tropos [6] modelers as they modeled a scenario of their choosing. Participants used either BloomingLeaf [7] or pencil and paper to construct their own model (called USER) and reviewed a second model of the same scenario (called AUTO). Adapting the OZ Paradigm [8], participants were told the AUTO model was automatically generated from their pre-study survey, but it was created by the authors of this paper. In this study, we consider four research questions: (RQ1) Were novice participants able to understand and use goal model constructs? If so, what difficulties were experienced by novices when being introduced to goal modeling and Tropos? (RQ2) To what extent were participants able to understand the AUTO model? (RQ3) Did participants choose to extend the USER or AUTO model? (RQ4) What utility (if any) was described by participants or observed by researchers?

The remainder of this paper is organized as follows. Section 2 introduces the methodology of our study. Section 3 describes and interprets our study observations. We discuss implications and future work in Sect. 4.

2 Methodology

This study was reviewed by the Smith College Institutional Review Board (Protocol: 18-110), see supplemental information online[1] for the study protocol.

Study Design. Participants were asked to complete a pre-session questionnaire, a single in-person modeling session, and a brief post-study questionnaire. Throughout the study, participants were asked to explore one of three motivating scenarios that was currently relevant to their life: choosing between majors, planing for after graduation, and deciding between study abroad options.

In the *pre-study survey*, participants were asked who, what, and why questions to elicit scenario information, which allowed us to create goal models of their trade-off decisions, known as the AUTO model. Four participants each were randomly assigned to either the *Paper* or *Tool* group.

In the *in-person modeling session*, participants were trained on the syntax and semantics of Tropos goal models (either on paper or in BloomingLeaf). To demonstrate Tropos language constructs, we used a modernized version of the Trusted Computing example [9]. After completing the training videos, we asked participants questions to ensure that they achieved an adequate level of understanding of Tropos. Participants were then given time to develop a goal model for their chosen scenario (i.e., the USER model). We asked the participants a list of questions to encourage them to expand and refine the USER model. Once participants felt they were finished with the USER model, we showed them the

[1] https://github.com/amgrubb/gore-study

AUTO model and asked them to extend it. After the participants finished, they were asked to compare the USER and AUTO model and choose between them, explaining their choice. Finally, the participants were given another opportunity to improve their chosen model and were asked followup questions.

In the optional *post-study questionnaire*, participants were asked if their decisions changed after the in-person session or if they had any further thoughts about the scenario they investigated.

Recruitment and Remuneration. We recruited eight participants in the summer of 2019. Participants were required to be proficient in English, without any formal training in goal modeling, and registered undergraduate students at Smith College in the 2019 school year. Each participant was paid $15.00 upon completion of the in-person session. The participants' information was kept anonymous and demographic information was not collected.

Data Analysis. With participants' consent, in-person modeling sessions were recorded and transcribed. Two of the authors independently used open-coding on the transcripts to explore themes [10]. In a group meeting, all authors discussed and converged on relevant themes. A second pass was then performed on the transcripts to ensure consistency in coding.

3 Results

In this section, we describe and interpret our study observations.

Understanding Tropos (and BloomingLeaf). We begin by answering RQ1: Were novice participants able to understand and use goal model constructs? If so, what difficulties were experienced by novices when being introduced to goal modeling and Tropos? After watching the tutorial videos, participants were able to identify actors, decomposition links, and contribution links. Some participants were confused about the difference between the specific types of contribution links. For example, one participant asked, "What's the difference between ++ and ++S?", while another asked "Why do some [links] have two negatives and others have one negative?". We were able to clarify participants' questions about model semantics prior to continuing; thus, we believe we can compare observations between participants.

Participants were encouraged to continue asking questions throughout the session for additional clarification. We observed two additional patterns over the remainder of the modeling. First, participants asked clarifying questions about model elements, including intentions. For example, a participant asked if "soft goals are things that don't have cut-offs?". These questions were most likely to occur at the beginning of the modeling process when participants had not yet fully transferred concepts from the Trusted Computing example. Second, participants asked how factors in their scenario could be added to the model. This style of question was most common once participants identified the main goals of their scenario. For example, participants asked, "Is my interest a resource or a task?" and "Should this link be and?". Overall, we found that novice participants were able to understand and use goal modeling constructs with assistance.

Model Understanding. Next, we consider RQ2: To what extent were participants able to understand the AUTO model? After each participant constructed the USER model, the researcher showed them the AUTO model and asked questions to understand how well it fit the participant's own scenario. None of the participants removed any nodes or connections from the AUTO model. Four participants proposed new connections between existing intentions and new intentions to be added. Two participants proposed how an intention could be further decomposed. After reviewing the AUTO model, all of the participants were able to present a clear cut answer about the major trade-offs and decisions in their own scenario, which were not clearly indicated in their pre-study questionnaire; thus, all participants were able to understand the AUTO model. This provides additional evidence that participants understood the goal model constructs.

Comparing Auto Model with Self-created Model. Following from the previous question, we look at RQ3: Did participants choose to extend the USER or AUTO model? Five participants chose the AUTO model (3 Paper & 2 Tool), while three chose the USER model (1 Paper & 2 Tool). For participants who chose the AUTO model, their reasoning was that the AUTO model was more organized or had more information, sample responses included, "I think *auto* is more clear and the arrows make more sense" and "...there are more direct connections between all of the components that go into my goals...". For participants who chose the USER model, they found it contains more detailed information, for example, "I'll extend the one I just made, it has a lot of links I don't want to recreate on *auto*". In reviewing the participants' rationale, we became concerned that participants may have chosen the AUTO model because it was more visually appealing than their own model, which is addressed in our future work.

Utility Evaluation. Finally and most importantly, we looked at RQ4: What utility (if any) was described by participants or observed by researchers? To evaluate what utility participants gain from constructing the model, we coded the responses to the post-study questionnaire and the transcripts of the in-person sessions, including the results to the question, "what did you learn from each model?". We observed four aspects of utility in the model construction activities in this study.

(1) Elicit Underlying Assumptions & Motivations. When talking out loud, participants discussed their motivation behind achieving a certain goal while putting that goal on the canvas. Thinking in the context of goal modeling helped participants to explore tacit information, though most of this information does not appear in the final model. Two comments included in this category were, "probably it is also a goal [to become pre-dental], 'cause after coming to Smith, I find that I really want to take a lot of chemistry courses and want to eventually make it a major", and "one of the goals of neuroscience is to achieve a deep understanding in one of the hard sciences, a task would be taking a bunch of upper-level classes".

(2) Explicitly Consider Structural Relationships. Study participants responded that they learned how to break down a bigger goal into smaller tasks and considered how to weigh different factors in the process. This lead to new insights

about how they could structurally think about the ultimate goal of their scenario. One participant said that she learned "...a different way to weigh things, compare things that are required vs possibilities. A new way to put thoughts down". Another participant said that this made her "...look at things as smaller steps, which is really valuable. It very much so set into stark terms the decisions. I'll have to make in the future, which I've been putting off".

(3) Learned Something New. Participants responded that after modeling their goals, some of the factors or goals were presented and linked within the model which they had not seriously considered before. This forced them to reconsider a perspective or factor that they had previously ignored. Two sample responses were, "I wasn't thinking a lot about how different perspectives would contribute to my decision-making, especially [my] pre-health advisor...but in this model it represents that.", and "I think one of the biggest conclusions is that industry job would satisfy less of my personal goals than teaching or research".

(4) Update Conclusions. Finally, participants responded that their personal preference was further reinforced or changed after the modeling experience. Some said that the idea of "soft goals" helped them to see the broader picture of the scenario, which influenced their decisions. For example, one participant said, "I think my preference is towards the french major now. It shifted more in that direction.", while another reinforced their concerns, saying, "I guess I was thinking a lot about money issues, but when you draw [out the model] it feels real and pressing."

4 Discussion

In this section, we discuss our observations and propose future studies.

Threats to Validity. The observations made in this study may be erroneous due to threats to validity, specifically construct validity. Subjects may have gained utility from discussing their scenarios. We did not separate the utility in talking through a problem from the utility of modeling the problem. Future studies will isolate this factor and compare them with the results of Kwan and Yu [11]. Since our participants asked clarifying questions while modeling, the researchers' answers may have biased the participants in some way. Our training videos can be improved by contrasting contribution links, providing more examples, and considering the work of Liaskos et al. [12]. Future studies should recruit participants already familiar with goal modeling and Tropos. Participants may have been *apprehensive towards evaluation* and wanted to perform well for our study. They may also have been biased by their own assessments due to *hypothesis guessing*. Finally, bias may exist in our analysis as a result of *experimenter expectancies*.

Discussion of Utility. In Sect. 3 we found evidence of utility; thus, we reject the null hypothesis as stated, but this is insufficient to answer our study question: To what extent do participants gain value in manually drawing goal models (in a tool or on paper)? Since this was an exploratory study, none of the observations were conclusive but instead inform our future work.

Future Work. We expect to refine our protocol and make the changes described above to address our threats to validity. We will repeat this study as a series of controlled experiments with larger numbers of participants over longer periods of observation, isolating each variable in this study. In future work, we intend to explore participants rationale for choosing the AUTO or USER model.

To investigate the effects of 'talking out loud', we will compare the use of goal modeling to other methods of choosing between trade-offs (e.g., pro and con lists). For each of these studies we should consider novice users (representing stakeholders), trained modelers, and stakeholders not involved in the modeling process. Finally, we hope to partner with software organizations to study goal model construction in an industrial context.

Summary In this paper, we presented an exploratory study of the utility of goal model construction. We found that participants understood and were able to use model constructs and found utility in both the USER and AUTO models. Participants gained insights into their assumptions, goals, and the structural nature of their scenarios with some participants changing their decision. Future work will repeat our study with trained modelers and address the issues noted in our discussion above.

References

1. J. Horkoff and E. Yu, "Comparison and Evaluation of Goal-Oriented Satisfaction Analysis Techniques," *Requirements Engineering*, vol. 18, no. 3, pp. 199–222, 2013.
2. J. Horkoff, T. Li, F.-L. Li, M. Salnitri, E. Cardoso, P. Giorgini, J. Mylopoulos, and J. Pimentel, "Taking goal models downstream: A systematic roadmap," in *Proc. of RCIS'14*, 2014, pp. 1–12.
3. A. Mavin, P. Wilkinson, S. Teufl, H. Femmer, J. Eckhardt, and J. Mund, "Does Goal-Oriented Requirements Engineering Achieve Its Goal?" in *Proc. of RE'17*, September 2017, pp. 174–183.
4. A. M. Grubb, "Reflection on Evolutionary Decision Making with Goal Modeling via Empirical Studies," in *Proc. of RE'18*, 2018, pp. 376–381.
5. OED Online, *"utility, n."*, Oxford University Press, 2018, www.oed.com/view/Entry/220771. Accessed: 2018-03-13.
6. A. Fuxman, L. Liu, J. Mylopoulos, M. Pistore, M. Roveri, and P. Traverso, "Specifying and Analyzing Early Requirements in Tropos," *Requirements Engineering*, vol. 9, no. 2, pp. 132–150, May 2004.
7. A. M. Grubb and M. Chechik, "BloomingLeaf: A Formal Tool for Requirements Evolution over Time," in *Proc. of RE'18: Tool Demos*, 2018, pp. 490–491.
8. J. F. Kelley, "An Empirical Methodology for Writing User-Friendly Natural Language Computer Applications," in *Proc. of CHI'83*, 1983, pp. 193–196.
9. J. Horkoff and E. Yu, "A Qualitative, Interactive Evaluation Procedure for Goal- and Agent-oriented Models," in *Proc. of CAiSE'09*, 2009, pp. 19–24.
10. A. Strauss, J. Corbin, and J. Corbin, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory.* SAGE Publications, 1998.
11. A. Kwan and E. Yu, "Goal Modeling without Stress: An Empirical Study of User Engagement," in *Proc. of iStar'17*, 2017, pp. 85–90.
12. S. Liaskos, N. Alothman, A. Ronse, and W. Tambosi, "On the Meaning and Use of Contribution Links," in *Proc. of iStar'19*, 2019.