Efficient Visual Notations for Efficient Stakeholder Communication

Ralf Laue
University of Applied Sciences
of Zwickau, Germany
Department of Information Science
Email: ralf.laue@fh-zwickau.de

Frank Hogrebe
Hessische Hochschule für Polizei
und Verwaltung Wiesbaden, Germany
Department of Public Administration
frank.hogrebe@hfpv-hessen.de

Boris Böttcher, Markus Nüttgens
University of Hamburg, Germany
Email: boris.boettcher@uni-hamburg.de,
markus.nuettgens@wiso.uni-hamburg.de

***Abstract*—The visual syntax of modelling languages can support (or impede) the intuitive understandability of a model. We observed the process of problem solving with two notation variants of i^* diagrams by means of an eye-tracking device. The number of wrongly answered questions was significantly lower when the alternative i^* notation suggested by Moody et al. was used. For the eye-tracking metrics "time to solve a task" and "number of eye fixations", no such significant result can be given. Furthermore, we identified a deficiency for the "dependency" symbol in the alternative notation.**

I. Introduction

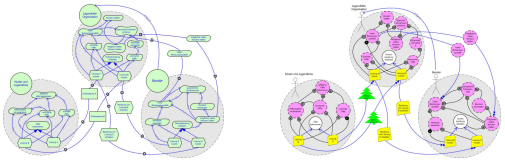
One of the key challenges in requirements engineering is to achieve efficient communication between requirements engineers and domain experts. Several graphical languages have been developed to support this communication. In the early phases of the RE process, the language i^* [1] is frequently used for modelling the social relationships between actors.

Based on the theoretical framework "Physics of Notations" [2], Moody et al. suggested improvements on the traditional visual syntax of i^* models in order to increase its understandability [3]. To compare the understandability of the alternative notation to the traditional one, we performed an experiment using eye-tracking device. This method can be used to measure how subjects interact with diagrams (see [4]).

II. Experimental Setup

For our experiment, we selected two i^* models. For both we created a version using the traditional i^* notation and a version using the visual symbols suggested by Moody et al. One model showed the classic meeting scheduler problem [1], the other model was an adaption of a model comparing possibilities for technical support to the youth counseling process [5]. Fig. 1 shows both versions of the latter model.

16 participants (8 for each notation variant) had to solve a set of six tasks while their eye movement was recorded. The participants have been recruited from students and the academic staff at the department of business information science at the University of Hamburg. Before the experiment, nobody of them was familiar with i^*. As a first step, each participant got an introduction to i^* by means of a self-learning computer course. Two different variants of the course have been developed (one using the alternative notation, the other using the alternative notation). This way, each participant



(a) traditional notation (b) alternative notation

Fig. 1. Two variants of an i^* model

was assigned to the "traditional notation" group or the "new notation" group. The participants had as much time as they wished to get familiar with the i^* concepts and the notation, and they had the possibility to ask additional questions if something remained unclear.

For collecting the data, we used the eye tracking device *Tobii T60 XL* with a 24-inch screen. Using the two diagrams, the participants had to answer six questions on the screen. They were instructed to press a key as quickly as possible if they have found the answer. Afterwards, they showed their answer on the screen. This way, we collected three kinds of data: whether the answer was correct, the time to complete the task and the record of the eye movement.

Our questions were as follows: What contributes to the softgoal "Low effort"? (Q1), How many dependencies from "Autom. Appointment Scheduler" are shown? (Q2), Which resources should "Appointment Participant" provide to other actors? (Q3), If "solution A" is chosen, which softgoals of "Kids and Youth" will be positively affected? (Q4), Are there softgoals of "Youth Counseling Organization" to which selection of "solution B" has a negative effect? (Q5), For which actors a dependency from "Counselor" is shown? (Q6).

III. Results

A. Correctness of the Answers

The mean number of errors in the group using the traditional notation was 2.625. For the group using the alternative notation, it was 1.125. We performed a one-tailed Mann-Whitney U-test on the data groups. The U value for the one-tailed Mann-Whitney test is 11 and the P-value is 0.014, i.e. the result is **significant** at the standard level of 95%.

All questions with exception of Q2 were answered better by the group using the alternative notation. Note that Q2 required

978-1-4799-3032-6/14/$31.00 © 2014 IEEE

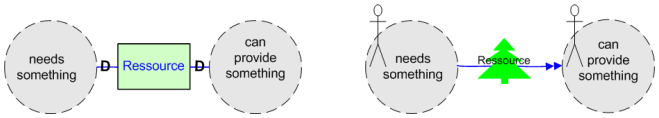


Fig. 2. Notations for Expressing a Dependency (Left: Traditional Notation, Right: Alternative Notation)

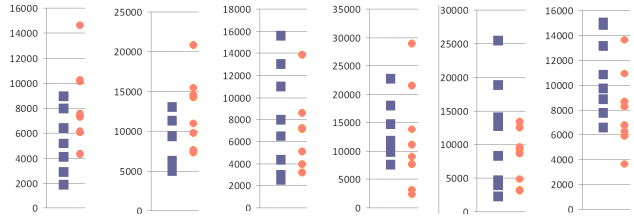


Fig. 3. Time Needed (in ms), Result for each of the Six Questions (Blue Squares - New Notation, Red Circles - Traditional Notation)

to reason about the *direction* of a dependency. As shown in Fig. 2, it is easy to misinterpret the dependency symbol ⟶ in the new notation in a way that a resource moves from the left actor to the right one. However, the intended meaning is that the left actor *depends on* the right actor who has to provide the resource. If we exclude the "dependency counting" question Q2, we get a **highly significant** result (P-value 0.00148).

B. Time Needed and Number of Eye Fixations

Before creating dotplot diagrams for visualising the distribution of the answering times, we removed outliers according to a commonly used method: We calculated the 25%-quartile q_{25} and the 75%-quartile q_{75} and disregarded measurements outside the interval $[q_{25} - 1.5 \cdot (q_{75} - q_{25}), q_{75} + 1.5 \cdot (q_{75} - q_{25})]$. Fig. 3 shows the distribution of the answering times to each of the 6 questions for both groups. For Q1 and Q5 (first two diagrams from the left), we observe that users using the new notation performed better. For Q3 and Q4, no differences can be observed, for Q2 and Q6 (the two rightmost diagrams) users performed a little better when using the traditional notation. While all these observations are **not statistically significant**, we observed once again that the questions for which the performance was better when using the traditional notation were the two questions that required to understand the direction of a dependency (Q2 and Q6).

From the recordings of the eye-tracker, we extracted the number of eye-fixations on the diagram while working out the answer to a question. The heatmaps in Fig. 5 visualise the durations of eye fixations when working with question Q1. A high number of fixations indicates less efficient search in the diagram. After removing the outliers in the same way as described for the variable "time needed" above, we obtained the dotplots shown in Fig. 4 where the questions are shown in the same order as in Fig. 3. It can be observed that the patterns in Fig. 4 and Fig. 3 show some similarities. Once again, **no statistically significant** differences between the groups could be found using the data shown in Fig. 4.

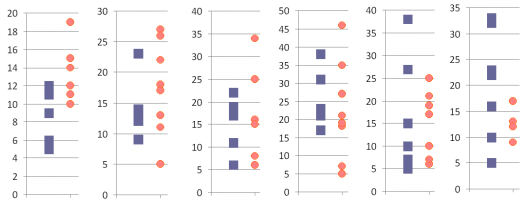


Fig. 4. Number of Eye Fixations, Result for Each of the Six Questions (Blue Squares - New Notation, Red Circles - Traditional Notation)

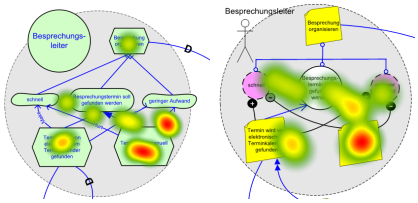


Fig. 5. Eye-tracking Heatmaps (Same Question, Different Notations)

IV. Conclusion

In our experiment, the number of correct answers significantly depend on the notation. This supports the hypothesis that i^* can be used more effectively when the visual notation would be improved. Answering times and number of eye fixations did not give a conclusive result. In general, we can state that the improvements suggested by Moody et al. are not always superior to the original notation. In particular, we have reason to assume that the arrow which symbolises a dependency can lead to a misinterpretation of the dependency direction. Our results are limited due to the fact that the number of participants in the experiment was rather low, therefore additional experiments will be necessary to recheck the results. The most important threats to validity refer to external validity: The subjects of our experiments (although having a background in business information science) did not have industrial experience, so the results may be different for other groups of i^* users. Second, answering questions on a screen as quickly as possible is not the typical use case for i^* models, therefore additional experiments with a more typical scenario are necessary. We are currently performing additional paper-and-pen experiments with a much larger number of participants and plan to publish a paper with full results.

References

[1] E. S. K. Yu, "Towards modeling and reasoning support for early-phase requirements engineering," in *3rd IEEE International Symposium on Requirements Engineering*, 1997, pp. 226–235.

[2] D. L. Moody, "The "physics" of notations: a scientific approach to designing visual notations in software engineering," in *32nd ACM/IEEE International Conference on Software Engineering*, 2010, pp. 485–486.

[3] D. L. Moody, P. Heymans, and R. Matulevicius, "Visual syntax does matter: improving the cognitive effectiveness of the *i** visual notation," *Requir. Eng.*, vol. 15, no. 2, pp. 141–175, 2010.

[4] A. I. Molina, M. A. Redondo, M. Ortega, and C. L. Rodero, "Evaluating a graphical notation for modeling collaborative learning activities: A family of experiments," *Sci. Comput. Program.*, vol. 88, pp. 54–81, 2014.

[5] J. Horkoff and E. S. K. Yu, "Evaluating goal achievement in enterprise modeling - an interactive procedure and experiences." in *The Practice of Enterprise Modeling*, ser. LNBIP, vol. 39. Springer, 2009, pp. 145–160.