Insights from Using Goal Models for Teaching Data Structures

Marcela Ruiz[1,*], Xavier Franch[2]

[1]*Zürich University of Applied Sciences, Winterthur, Switzerland*
[2]*Universitat Politècnica de Catalunya, Barcelona, Spain*

Abstract

Previous research has shown the feasibility of using goal models as a notation to describe data structures. Such results motivated further empirical research to elucidate the extent to which teaching data structures using goal models helps students to better understand (i.e., describe and compare) data structures, and be more efficient when solving their assignments. This paper reports on the experience of introducing goal models for teaching data structures in an undergraduate computer science program during the Autumn semester 2022. We report on the used teaching materials, partial results regarding students' performance, and our vision for building a platform for teaching data structures with goal models.

Keywords

iStar 2.0, data structures, teaching experience, experimental insights

1. Introduction and Motivation

Previous research has shown the feasibility of using goal models based on the i* framework and its corresponding language iStar 2.0 [1] as a notation to describe data structures to be used in classical imperative programs written in e.g., Java or C# [2]. In this line of work new constructs were introduced in iStar2.0 and adding them to the iStar2.0 metamodel to allow the description of well-known abstract data types [3] like sequences, functions, and graphs, and the data structures implementing them [4]. Such results motivated empirical research to elucidate the extent to which teaching data structures using goal models help students to better understand (i.e., describe and compare) data structures, and be more efficient when solving their assignments [5]. This paper reports on the experience of introducing goal models using the iStar 2.0 language for teaching data structures in an undergraduate computer science program during the Autumn semester 2022. We built knowledge capsules containing the description and implementation of data structures, and in contrast to previous empirical evaluations, we expanded the analysis by evaluating subjects' performance in terms of effort. We report on the used teaching materials, partial results from the quasi-experiment, and our vision for building a platform for teaching data structures with goal models. The rest of the paper is structured as follows: Section 2 presents the design of the quasi-experiment; Section 3 presents partial

The 16th International iStar Workshop, September 03–04, 2023, Hannover, Germany
*Corresponding author.
marcela.ruiz@zhaw.ch (M. Ruiz); xavier.franch@upc.edu (X. Franch)
0000-0002-0592-1779 (M. Ruiz); 0000-0001-9733-8830 (X. Franch)

results about subjects' performance in terms of validity of described data structures and effort, and we conclude the paper with main insights and outlook about the future work.

2. Teaching Data Structures with Goal Models: Design of a Quasi-Experiment

During the Autumn semester 2022 (from September 2022 to January 2023) we conducted a quasi-experiment [6] (subjects and objects of study were not randomly selected from the population) to investigate students' performance' effects when describing and comparing data structures (DS) after being exposed to DS described using iStar 2.0. This quasi-experiment has been designed according to Wohlin et al. [6], and it is reported according to Jedlitschka and Pfahl [7]. This report differentiates from a previous experiment we conducted in 2020, in which we did not measure effort and we clearly did not have questions regarding the main differences between DS belonging to two different hierarchies [3] [5].

The subjects were formed by a group of third-semester computer science students enrolled in the course of Algorithms and DS (ADS) offered at the Zürich University of Applied Sciences. The total number of subjects was 54 belonging to two distinct groups with the same schedules. Both groups have students with experience in industry. In general, some students are currently working in industry (part-time students), and none of them had been in contact with the i* framework or used iStar2.0 for describing DS. The course planning was not updated to incorporate the experimental set-up, still maintaining the original course objectives. However, the content presented to one of the two groups was updated adding iStar2.0 to some of the DS descriptions and examples taught. The group who received materials of DS with iStar 2.0 was considered as experimental group. The subjects executed the experimental task as part of the course's end-of-semester (closed book) oral exam. The main research goal (MRG) and research question (RQ) of our study are presented below:

MRG: The main research goal, according to the Goal/Question/Metric template [8] is to **gain** empirical knowledge on the use of iStar 2.0 to describe data structures **for the purpose** of *understanding whether a high-level, goal-oriented approach as provided by iStar 2.0 could help to describe data structures in a more structured form* [2], **with respect to** subjects' accuracy and effort in selecting and comparing data structures **from the point of view of** *computer science students and the researchers (authors of this paper)* **in the context of** an algorithms and data structures course at the Zürich University of Applied Sciences in Switzerland.

RQ1: When the subjects analyse different DS, to what extent do they identify the intrinsic characteristics of each data structure in the right way? To answer this research question, we compare completeness, validity, and effort when data structures are described and characterized by subjects that have been exposed to DS taught using iStar 2.0; to a control group that has been exposed to DS taught in the traditional way (without iStar 2.0).

Variables. We consider one independent variable: DS specification, i.e., the way DS are described:

- DS specified with iStar 2.0, as defined in [2]
- DS specified in the traditional way, serving the purpose of a control group.

For this study, we consider the following dependent variables, which are expected to be influenced to some extent by the independent variable:

Completeness. The degree to which all the intrinsic characteristics of a certain data structure have been properly described. To facilitate this calculation, the researchers take into account a reference solution containing the minimum indispensable description.

Validity. The degree to which the characteristics of a certain data structure are described in the right way. Acting as reviewers, the researchers identify properly or wrongly described characteristics based on a reference model, and then discuss them until they agree on the verdict.

Efficiency. The degree to which the subjects' efficiency is affected when describing DS. Acting as reviewers, the researchers take note of the time each subject took when describing DS.

Hypotheses. We define null hypotheses (represented by a 0 in the subscript) that correspond to the absence of an impact of the independent variables on the dependent variables. Alternative hypotheses (represented by a 1 in the subscript, e.g., $H1_1$ is the alternative hypothesis to $H1_0$) suppose the existence of such an impact. Alternative hypotheses correspond to our expectations: DS specified with iStar 2.0 will have a positive impact on the dependent variables.

$H1_0$: DS specified with iStar 2.0 do not influence the completeness of described DS.

$H1_1$: DS specified with iStar 2.0 do influence the completeness of described DS.

$H2_0$: DS specified with iStar 2.0 do not influence the validity of described DS according to incorrect characteristics.

$H2_1$: DS specified with iStar 2.0 do influence the validity of described DS according to incorrect characteristics.

$H3_0$: DS specified with iStar 2.0 do not influence the subjects' effort when describing DS measured in time unit (minutes and seconds).

$H3_1$: DS specified with iStar 2.0 do influence the subjects' effort when describing DS measured in time unit (minutes and seconds).

The subjects were provided with four different types of input questions to increase the external validity of the experiment based on the questions' objective (see Table 1). A summary of the experimental procedure is presented in Figure 1, and a sample description of the experimental objects is presented in Table 2.

Both control and experimental groups were provided with the original materials for teaching DS in the traditional way. In addition, both groups were exposed to a set of self-assessment tests that would allow them to monitor their learning progress during the semester. For the experiment, the experimental group was exposed to DS specified using iStar 2.0 by means of descriptions and examples called knowledge capsules. The knowledge capsules were built for three hierarchies of DS: sequences, graphs, and maps. The knowledge capsules are available in the online appendix (https://shorturl.at/oxT38).

Table 1
Distribution of experimental objects per group.

		Input DS question type (i.e., object)	
	Data Structures Specification	Group 1	Group 2
Control Group (24 subjects)	Subjects exposed to data structures specified in the traditional way	Question A1 Question B1 Question C1 Question D1	Question A2 Question B2 Question C2 Question D2
Experimental group (30 subjects)	Subjects exposed to data structures specified with iStar 2.0	Question A3 Question B3 Question C3 Question D3	Question A4 Question B4 Question C4 Question D4

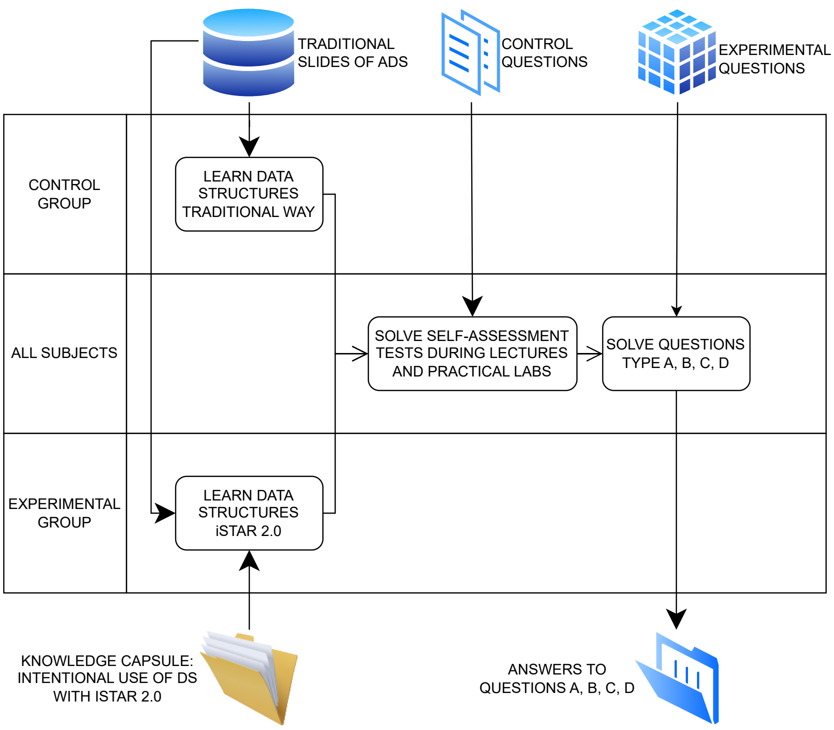

Figure 1: Experimental procedure in a nutshell.

Table 2
Description of the experimental objects: sample with questions' type.

A. Description of the main differences and similarities between two DS.
Question A1: Which are the differences and similarities between stacks and lists in ADS? Answer: Differences: Stacks return the top element in the stack; lists return a specific element from the list. In addition, lists can be traversed. Similarities: both data structures have elements arranged linearly, operations for element addition and removal; and a get operation that is specialised for stacks and lists.
B. Description of the main requirements (i.e., operations) of a given DS.
Question B1: What is an AVL tree, and which are its main operations? Answer: Greater/Smaller than function, Update, Insert, Delete, Get... Individual access to elements greater or smaller than a given key.
C. Description of the main differences between two DS from two different hierarchies
Question C1: We need to create a data structure that we can control its size, provides fast lookup to elements by providing a key. Having the elements ordered is not a requirement. What is the most appropriate data structure and why? Answer: Hash tables are the most appropriate data structure because they demand a known approximate size. This size can be extended in the future if needed. A hash function needs to be provided to store and find elements. There is not a specific order to store elements.
D. Control question about DS without associated iStar 2.0 language description.
Question D1: What is the main difference between a singly and doubly linked list? Answer: Doubly linked lists have a pointer to the previous element.

Table 3
Partial results from measuring the variable effort from the comparative quasi-experiment.

	Time average taken per group	
Questions Type	Control group	Experimental group
A	0.56	1.55
B	0.74	1.31
C	0.96	1.44
D	0.61	0.65

3. Students' Performance: Partial Results Regarding Answers' Validity and Subjects' Effort

After collecting all the results from the exams, the descriptive result for the variable validity is presented in Figure 2. Regarding effort, the average time taken by the groups on each question is summarized in Table 3.

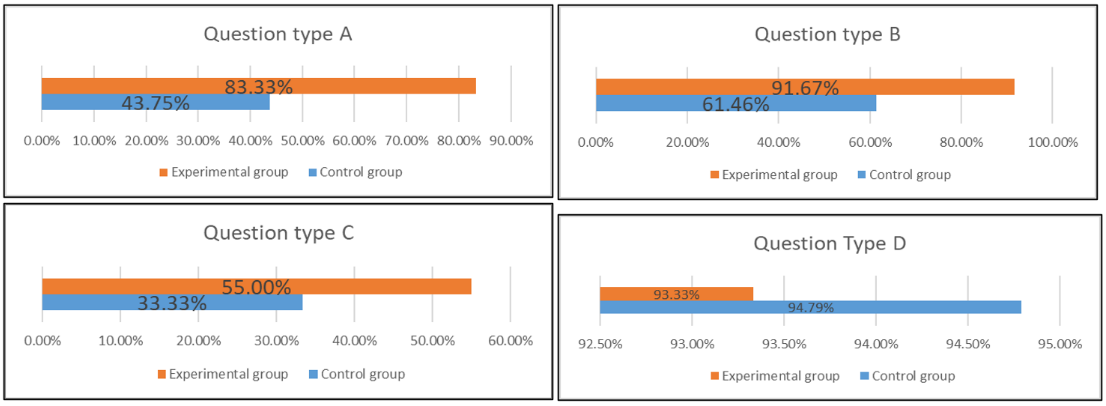

Figure 2: Partial results from measuring the variable validity from the comparative quasi-experiment.

4. Analysis of Threats to Validity of the Results

The execution of the quasi-experiment and partial results presented in this paper have some threats to validity that we want to highlight.

Conclusion validity: The random heterogeneity of the subjects is a threat to the validity of the results of this quasi-experiment. We have selected two distinct groups that are enrolled in the DS courses. However, we did not run any randomization analysis or techniques to select the subjects and groups. This threat affects the potential generalization of the results.

Internal validity: In this comparative quasi-experiment the control and experimental groups received dissimilar materials for learning data structures with and without using iStar 2.0. There is a threat regarding the interaction of groups and treatment. Despite the groups not being mixed in with other courses or activities in their study program, there is the possibility that subjects have exchanged materials and experiences. To mitigate this threat, we created control questions to evaluate how different both groups are regarding experience and knowledge regarding data structures without iStar 2.0. Since the experiment takes place during the whole academic semester, it is difficult to control the potential interaction between groups and treatment. Regarding instrumentation, the evaluators used chronometers to take note of the time each subject took to answer each question. There is a threat regarding the precision in which the chronometers were started or stopped. To mitigate this threat, both evaluators took measures separately and the values were checked for consistency and error analysis.

Construct validity: It is important to highlight that the experimenter was also involved in the evaluation of the results from the subjects, which can generate regarding experimenter expectancies. To mitigate this threat, one evaluator who is not involved in the experiment evaluated the results from the subjects.

5. Conclusions and Future Work

This paper summarizes our ongoing effort on investigating the effects of introducing iStar 2.0 for teaching data structures in undergraduate programs in Computer Science. For this purpose, we have conducted a quasi-experiment in which we introduce a set of knowledge capsules for teaching and exemplifying data structures using iStar 2.0. In contrast to previous empirical evaluations, we extend our variable set and introduce the analysis of subjects' effort when describing data structures. The partial results presented in this paper show that subjects exposed to data structures described with iStar 2.0 outperform subjects in a control group following the regular teaching material. On the other hand, the average of effort taken by the subjects when solving an exercise shows that subjects in the control group were faster. This can probably be explained by the fact that iStar 2.0 notation motivates students to further think about the applications and requirements of each data structure. These results further motivate research in the line of exploring the extent students' efficiency can increase with the constant use of data structures specified with iStar 2.0.

We plan to further analyze the partial results obtained from this experiment and conduct subsequent replications to collect more empirical data in the Autumn semester 2023-2024. In addition, we plan to build a platform to make available the knowledge capsules with examples. The first prototype of a tool to allow the modelling and simulation of data structures is currently under construction and we envision the platform will facilitate the massive adoption of intentional modelling for describing and comparing data structures. We want to explore if these knowledge capsules can be connected with the module concept [9] [10] that we are using as model encapsulation mechanism for the DS themselves [3].

References

[1] F. Dalpiaz, X. Franch, J. Horkoff, istar 2.0 language guide, 2016. arXiv:1605.07767.

[2] X. Franch, Using i* to describe data structures, in: Thirteenth International iStar Workshop (iStar 2020), CEUR-WS.org., 2020, pp. 49–54.

[3] X. Franch, Extending istar2.0 metamodel to define data structures, in: Proceedings of the 14th International iStar Workshop (iStar 2021), CEUR-WS.org., 2021, pp. 28–34.

[4] X. Franch, Estructuras de Datos: Especificación, Diseño e Implementación, Ed. UPC, 1993.

[5] X. Franch, M. Ruiz, Goal-oriented models for teaching and understanding data structures, in: Conceptual Modeling (ER 2021), Springer International Publishing, 2021, pp. 227–241.

[6] C. Wohlin, P. Runeson, M. Hst, M. C. Ohlsson, B. Regnell, A. Wessln, Experimentation in Software Engineering, Springer Publishing Company, Incorporated, 2012.

[7] A. Jedlitschka, D. Pfahl, Reporting guidelines for controlled experiments in software engineering, in: 2005 International Symposium on Empirical Software Engineering, 2005., 2005, pp. 10 pp.–. doi:10.1109/ISESE.2005.1541818.

[8] R. Solingen, E. Berghout, The goal/question/metric method: A practical guide for quality improvement of software development (1999).

[9] X. Franch, Incorporating modules into the i* framework, in: B. Pernici (Ed.), International Conference on Advanced Information Systems Engineering (CAiSE), Springer Berlin Heidelberg, 2010, pp. 439–454.

[10] A. Maté, J. Trujillo, X. Franch, Adding semantic modules to improve goal-oriented analysis of data warehouses using i-star, Journal of Systems and Software 88 (2014) 102–111. doi:https://doi.org/10.1016/j.jss.2013.10.011.