

Automatic and Adaptive Assessments by Bayesian Networks

A Case Study from SUPSI

Francesca Mangili*

Alessandro Antonucci†

September 14, 2022.

Istituto Dalle Molle di Studi di Intelligenza Artificiale (IDSIA)
Department of Innovative Technologies (DTI)
University of Applied Sciences and Arts of Southern Switzerland (SUPSI)
Polo universitario Lugano, Via la Santa 1 CH-6962 Lugano-Viganello (TI)

Introduction

In the recent times a considerable attention has been paid to the development of new technologies for education, with the pandemics further accelerating such process. *Assessment* represents a key element of any educational action, but, as a matter of fact, less attention has been paid to the creation of new technologies in this context. This sounds unfortunate, as the development of techniques for automatic assessment has been a classical field of investigation for Artificial Intelligence (AI) and the developments induced by recent advances in the area of deep learning considerably increased the potential of these techniques. In this short document we summarise the research carried out in the last years at the University of Applied Arts and Sciences of Southern Switzerland (SUPSI) for the development of innovative technologies for automatic assessments based on AI techniques [5]. This includes a case study about language proficiency assessment, describing an adaptive tool able to reduce the duration of a test for entry level assessment [6]. This kind of approaches might be a natural option for various assessment tasks in the Swiss universities.

Motivation

In general mastering a competence not only requires to hold individual basic components of such competence, but also integrating those skills and applying them in different contexts. Current assessment tools often focus on individual skill components only. Developing novel tools able to model skill interactions, thus making possible the assessment of complex skills, is an important challenge for modern academic institutions. Having as a goal the use of such assessment models for designing interventions (either human-driven or based on intelligent tutoring systems), the models should be also interpretable and, ideally, describe causal relations in the model.

For those reasons, our research has especially focused on the use of probabilistic graphical models. Following a consolidated research direction in the field, we adopted Bayesian networks (BNs) for the modelling of the relations between multiple skills of the student and his/her answers to given questions. Such networks are a natural formalism for causal relation modelling and, due to their simple graphical representation, are understandable also in the absence of a strong mathematical background.

A limit to the widespread adoption of this type of models is the effort required for the elicitation of both: (i) the structure of the relations between basic skills, and (ii) the parameters specifying them. A possible solution is learning the model from data. However, in educational domains, large amounts of homogeneous student data are rarely available and other directions should be considered instead. Therefore, some research was devoted to the analysis and implementation of different elicitation strategies and simplifying assumptions.

*Francesca Mangili (francesca.mangili@supsi.ch). Researcher at IDSIA-DTI-SUPSI.

†Alessandro Antonucci (alessandro.antonucci@supsi.ch). Senior Lecturer-Researcher at IDSIA-DTI-SUPSI.

Impact

Our research in the context of expert-elicited BN models, has identified two strategies for reducing the elicitation issues discussed in the previous section, encouraging their more widespread use. The first strategy consists in a robust, set-valued, elicitation of the parameters intended to give higher realism to the probabilistic estimates to be provided by domain experts [1, 5]. The second is instead based on parametric probabilistic models involving logical gates to be embedded in the BN structure and leading to simpler elicitation efforts [2]. The latter technique has been successfully adopted in an European project having SUPSI as a partner and intended to provide scope-tailored and industry-ready hardware to SMEs.¹ Open-source libraries have been also released² to simplify the implementation of such tutoring systems [3]. As a last step in this simplification process, we have proposed an approach for deriving the model structure directly from more commonly used assessment rubrics [4]. Such a flexible setup makes especially simple the elicitation process for any type of task, given the relative assessment rubric, making it possible to develop a learner model without a too big modelling effort.

Probabilistic learner models can thus more easily be adopted to represent skills interactions and, as a consequence, to robustly assess complex skills. They can, thus, be the base for the development of adaptive assessment or intelligent tutoring system delivering a customised user experience. This is for instance the framework adopted by SUPSI to assess English and German proficiency of freshman students, hence assigning them to courses of appropriate levels [5, 6]. The original (standard) quiz was delivered online and included a large number of questions of varied difficulty, due to need of assessing with a single evaluation tool the very different language skill levels of freshman students. Based on the learner model, the AI tool developed was able to select the questions to be administered to the individual student based on the skill level inferred from her/his previous answers. Such a tailored selection of questions was proven able to guarantee the same assessment accuracy of the complete test with a reduced number of questions.

Conclusions and Outlooks

Assessment is a fundamental part of the learning process. Accurate assessment of complex and interacting skill levels may be used to offer personalised learning paths to higher education students coming from different backgrounds, thus potentially reducing dropouts and improving the quality of the educational process. Moreover, while *summative* assessments, having the value of certification, may require some degree of standardisation, continuous assessment has a formative role which could be better achieved if tailored on single individuals. The BN learner modelling framework proposed in [4] represents an important step in this direction, as it strongly reduces the effort required for the model definition. An obvious development is the integration with deep learning technologies for natural language understanding, which would add more flexibility to the current framework.

References

- [1] A. Antonucci, F. Mangili, C. Bonesana, and G. Adorni. A new score for adaptive tests in Bayesian and credal networks. In J. Vejnárová and N. Wilson, editors, *Symbolic and Quantitative Approaches to Reasoning With Uncertainty*, pages 399–412, Cham, 2021. Springer International Publishing.
- [2] Alessandro Antonucci, Francesca Mangili, Claudio Bonesana, and Giorgia Adorni. Intelligent tutoring systems by Bayesian nets with noisy gates. In *The International FLAIRS Conference Proceedings*, volume 35, 2022.
- [3] C. Bonesana, F. Mangili, and A. Antonucci. ADAPQUEST: a software for web-based adaptive questionnaires based on Bayesian networks. In *AI4EDU: Artificial Intelligence for Education (IJCAI-2021 Workshop)*, Virtual Event, 2021.
- [4] F. Mangili, G. Adorni, A. Piatti, C. Bonesana, and A. Antonucci. Modelling assessment rubrics through Bayesian networks: a pragmatic approach. *Proceedings of 2022 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, 2022.
- [5] F. Mangili, C. Bonesana, and A. Antonucci. Reliable knowledge-based adaptive tests by credal networks. In A. Antonucci, L. Cholvy, and O. Papini, editors, *Symbolic and Quantitative Approaches to Reasoning with Uncertainty. ECSQARU 2017*, volume 10369 of *Lecture Notes in Computer Science*, pages 282–291. Springer, Cham, 2017.
- [6] F. Mangili, C. Bonesana, A. Antonucci, M. Zaffalon, E. Rubegni, and L. Addimando. Adaptive testing by Bayesian networks with application to language assessment. In Alessandro Micarelli, John Stamper, and Kitty Panourgia, editors, *Intelligent Tutoring Systems: 13th International Conference, ITS 2016, Zagreb, Croatia, June 7-10, 2016. Proceedings*, Lecture Notes in Computer Science, pages 471–472, 2016.

¹<https://kitt4sme.eu>.

²<https://github.com/IDSIA/adapquest>.