Teaching of Formal Methods for Software Engineering

Keywords: Formal Modelling, Collaboration, Teaching

Abstract: The use of Formal Methods (FMs) offers rigour and precision, while reducing ambiguity and inconsistency. The major barriers hindering the adoption of FMs in industry are the problems of understandability, comprehensibility, and scalability. To solve the understandability problem, from one side, the readability of the method have to be increased, but from another side, an appropriate teaching and learning approach have to be introduced. This paper presents an overview of existing approaches on teaching of FMs and Logic, also discusses the common issues in teaching of this subjects.

1 INTRODUCTION

For the development of embedded real-time systems in most cases experts of different disciplines have to cooperate, and for such a cooperation a specification of the developing system, i.e., precise and detailed description of its behaviour and/or structure, is important. Embedded systems are real, but their behaviours are modelled by mathematical objects, about which one can argue formally. One aim of Formal Methods (FMs) is to prove or to automatically evaluate behaviour properties of a system in a systematic way, based on a clear mathematical theory.

When dealing with formal methods, we are mainly concerned with the methods’ soundness and correctness, and sometimes also its mathematical elegance, but usually do not take into account such aspects as readability, usability, or tool support. This leads to the fact that FMs are perceived by most engineers as “something that is theoretically important but practically too hard to understand and to use”, (Crocker, 2006) even suggests to replace the somewhat “unattractive” name ‘formal methods’ with ‘verified software development’. We might have here a phenomena similar to the research field of Artificial Intelligence (AI): deeply appreciated at the beginning of AI discipline, it was seen as almost useless in 70s, cf. (Crevier, 1993), but was “reborn” with the paradigms of ‘expert systems’ and ‘intelligent agents’, and ‘data mining’.

Even small changes of a formal method can make it more understandable and usable for an average engineer. Moreover, human factors engineering need to be incorporated into the software development process, cf. (Spichkova et al., 2015), but the starting point for an adaptation of FMs in practice is the education in these methods. In the last decade, a number of teaching programs has been initiated to solve this problem. For example, the Top SE program in Japan was introduced with the aim to produce “superarchitects” who can promote practical use of advanced, scientific methods and tools, including formal methods, for tackling problems in software engineering, cf. (Ishikawa et al., 2009)

In this paper, we discuss the common issues in teaching of FMs and logic, as well as review the various approaches for teaching Formal Methods for Software Engineering that have been proposed, and discuss how they address the above mentioned challenges. The focus of our analysis here is on collaborative and communication aspects of software development using formal methods and logical modelling.

2 TEACHING FORMAL METHODS: CHALLENGES

The discourse on what to teach in Formal Methods and how to teach it has been going on for decades. It seems to be widely agreed that Formal Methods educators face the following challenges:

- There is a great diversity in the students’ background and cognitive skills due to the globalisation of higher education, which requires constant adaptation, cf. (Hoare, 2013; Feast and Bretag, 2005).
- Students have decreasing mathematical background as the curricula become more practice-oriented, leaving less room for theoretical courses, cf. (Bjørner and Havelund, 2014; Crocker, 2006; Zamansky and Farchi, 2015a).
- Students have less motivation as they are strongly focused on the direct relevance of what they study
to their daily practice, often failing to see the link of Formal Methods to the real world, cf. (Tavolato and Vogt, 2012; Jeanette, 2000).

2.1 Cultural Background

As per UNESCO statistics\(^1\), the number of students who have crossed a national border to study, or are enrolled in a distance learning programme abroad, grows around the world. To denote this group of students, UNESCO introduced a new term – *internationally mobile students* (IMSs). In 2012, at least 4 million students went abroad to study, up from 2 million in 2000, representing 1.8% of all tertiary enrolments or 2 in 100 students globally. According to this statistics, 5 destination countries hosted nearly one-half of total IMSs: the United States (hosting 18%), United Kingdom (11%), France (7%), Australia (6%), and Germany (5%).

Internationalization of the higher education has created the so-called *borderless university*, which provides better opportunities for learning and increases the human and social sustainability, cf. (Woodcraft, 2012; Vallance et al., 2011; Penzenstadler et al., 2012). An obstacle to successful transnational teaching and learning could be the diversity in cultural and technical/educational backgrounds of teachers and students (as well as among the students). For example, the learning style and perception of the plagiarism problems is different by students coming from Asian and European countries, cf. (Zobel and Hamilton, 2002). This diversity has to be taken into account while teaching and assessing the students.

A partial solution to this problem might be introduction of *active and inductive* learning in Software Engineering education process (Sedelmaier and Landes, 2015). In the traditional *deductive* teaching, the lecturer introduces general theoretical principles and mathematical models, proceeds with examples on the applications of these principles and models, and concludes with practical exercises. An alternative to the deductive teaching is an inductive teaching approach that includes a range of instructional methods, such as inquiry learning, problem-based learning, project-based learning, case-based teaching, discovery learning, and just-in-time teaching, cf. (Prince and Felder, 2006) for more details.

A recent work (Alharthi et al., 2015; Spichkova and Schmidt, 2015) presents an approach on requirements specification and analysis for eLearning systems and for the geographically distributed software and systems. The eLearning systems are usually developed to use within different organisations or even different countries. The organisation/country-specific requirements can differ in each particular case because of technical, cultural, or legal diversity. The challenge is to deal with this diversity in a systematic way, avoiding contradictions and non-compliance.

2.2 Mathematical Background

The fundamental role of Logic and FMs is recognised in the ACM CS (Sahami et al., 2011) and SE (LeBlanc et al., 2006) undergraduate curriculum guidelines. The pioneer of SE, David Parnas, believes that a solid understanding of logic is essential for a software engineer: “*Professional engineers can often be distinguished from other designers by the engineers ability to use mathematical models to describe and analyze their products*.”, cf. (Parnas, 1993). In his works, Parnas has noted the problem of understandability of formal specifications more than 20 years ago. To make the formal specifications more attractive for practitioners, Parnas suggested to make the formal expressions and system specifications shorter, changing their size and perceived complexity.

A decade later, a Symposium on Teaching Formal Methods was to explore the failures and successes of formal methods education, cf. (Tea, 2004). Now another decade is gone, but we are facing very similar problems: understandability and readability of FMs. Moreover, there is increasing concern that computing curricula are drifting away from fundamental commitment to theoretical and mathematical ideas (Tucker et al., 2001). As (Bjørner and Havelund, 2014) points out, “*Today’s masters in computing science and software engineering are not as well educated as were those of 30 years ago*”. Mandrioli (2015) further comments on a “*general tendency towards soften the teaching of engineering principles; this requires a certain amount of heroism to convince students that not everything can be obtained without effort*”, cf. (Mandrioli, 2015).

The problems are even more acute in universities of applied sciences (Tavolato and Vogt, 2012) and in the Information Systems discipline (Zamansky and Farchi, 2015a). The lack of empirical evidence that mathematical background is directly relevant for practitioners makes it even more difficult to convince decision makers that this situation must be handled. A notable example of such evidence is the Beseme project (Page, 2003): in a three-year study, empirical data on the achievements of two student populations was collected: those who studied discrete mathematics (including logic) through examples focused

\(^1\)http://www.uis.unesco.org/Education/Pages/international-student-flow-viz.aspx
on reasoning about software, and those who studied the same subject illustrated with more traditional examples. An analysis of the data revealed significant differences in the programming effectiveness of these two populations in favour of the former.

2.3 Lack of Motivation

Currently, FMs have very limited use in industrial software development process, which is a significant hurdle in making the based on FM courses attractive to the students. Woodcock et al. present survey of industrial use, comparing the situation in 2009 with the most significant previous surveys, and discuss the issues surrounding the industrial adoption of formal methods, cf. (Woodcock et al., 2009). Thus, we have a vicious cycle: To embed FMs into the software development lifecycle on industrial level, the FMs and the corresponding mathematical background have to be a part of the university curriculum. But the students are not motivated to learn FMs until they are not largely adopted by industry.

As FMs require a mathematical background and abstract thinking skills, many students have negative perceptions and even fear of courses that require dealing with complex mathematical notations. This is strongly related to the phenomenon of mathematical anxiety, cf. (Wang et al., 2014; Sherman and Wither, 2003). The term mathematical anxiety was introduced in 1972 by Richardson and Suinn as “feelings of tension and anxiety that interfere with the manipulation of numbers and the solving of mathematical problems in a wide variety of ordinary life and academic situations,” cf. (Richardson and Suinn, 1972). As stressed by Wang et al., mathematical anxiety has attracted recent attention because of its damaging psychological effects and potential associations with mathematical problem solving and achievement.

3 APPROACHES FOR TEACHING OF FORMAL METHODS

Teaching Software Engineering courses is a difficult task, as it requires imparting abstract reasoning skills necessary for problem solving (Sprankle and Hubbard, 2011). One of the solutions to this problem would be embedding in the Software Engineering curriculum such subjects as logic and formal specification, because these subjects can provide a good starting point for exploring concrete ways in which we abstract thinking can be taught (Zamansky and Farchi, 2015a).

The approaches we discuss in Section 3.2 mostly focus on overcoming issues that comes from the diversity in mathematical background (or even lack of a solid mathematical background) as well as from the lack of motivation to study formal concepts. We do not attempt to cover the issues that comes from the diversity in the learning style and perception of the plagiarism problems, because these issues are not FMs specific.

3.1 Teaching Strategies

As pointed out by Ferreira et al., the success of teaching process depends on the amount of self-learning and self-discovery that is left for the students: “If the teacher discloses all the information needed to solve a problem, students act only as spectators and become discouraged; if the teacher leaves all the work to the students, they may find the problem too difficult and become discouraged too. It is thus important to find a balance between these two extremes.”, cf. (Ferreira et al., 2009). MathIs project, presented by Ferreira et al., aimed to reinvigorate secondary-school mathematics by exploiting insights of the dynamics of algorithmic problem solving.

As suggested in (Wang and Yilmaz, 2006), the approaches in integrating FMs into software engineering curriculum can be divided into three main categories:

1. to avoid FMs,
2. to devote a specific course with emphasis on formal verification of source code;
3. to redesign the entire program so that formal methods are integrated throughout the curriculum.

However, the work of Wang and Yilmaz does not mention another category, which can be very promising for integrating FMs into software engineering curriculum: to introduce a specific course that

- covers basics of logic and FMs,
- does not require a deep knowledge in mathematics, as only the core aspects of the FMs will be introduced,
- uses visualisation and ramification strategies to make the material more understandable and less “boring”.

One examples of courses from the above category is the course Applied Logic in Engineering, or a “logic for everybody” course (Spichkova, 2016). Other examples include the Logic and FM course designed for Information Systems students (Zamansky and Farchi, 2015b), and a series of courses specifically adapted to the needs of university of applied sciences are described in (Tavolato and Vogt, 2012). Recently
courses in the spirit of “computational thinking for everybody” envisioned by J. Wing in (Wing, 2006a) have begun to be offered at various departments, e.g., IS103 Computational Thinking course at the Singapore Management University and the COMP101 Computational Thinking and Design course at the University of Maryland.

Another way to attract students while teaching FMs was presented by Curzon and McOwan: Within the engagement project cs4fn, Computer Science for Fun, they taught logic and computing concepts using magic tricks, cf. (Curzon and McOwan, 2013).

There are also recent approaches on embedding the e-learning and blended learning strategies in teaching of mathematics and logic, cf. (Pokorny, 2012).

3.2 Visualisation and Tool Support

Visualisation tools using a notional machine have been used since the early 1970s to promote understanding of programming constructs (Mayer, 1975; Mayer, 1981).

The cognitive load can be reduced through visualisation of the learning tasks (Pane and Myers, 1996; Powers et al., 2007). Visualisation can help correct misconceptions as they commonly occur when outcomes are not readily visible, cf. (Sirkkä and Sorva, 2012). Advanced tasks could be designed to facilitate critical thinking by rewarding optimal or near optimal solutions as they cause students to reflect on their strategies. Moreover, publishing optimal benchmark figures for different configuration may promote greater interaction about possible strategies thus leading to a form of social constructivism that promotes higher levels of learning (Kozulin et al., 2003).

Vosinakis et al. introduced the MeLoiSE platform (Meaningful Logical Interpretations of Simulated Environments) for teaching Logic Programming, cf. (Vosinakis et al., 2014). The platform developers focused on visualisation aspects, to allow the students experience a collaborative visual interface to the Prolog programming language.

AutoFocus tool, cf. (Hölzl and Feilkas, 2010; Spichkova et al., 2012), was developed as a scientific prototype for formal modelling of distributed, timed, reactive systems. The implemented modelling language was based on a graphical notation that can be used to teach basic modelling constraints and the usability aspects (Spichkova et al., 2013).

Korecko et al. developed a toolset for support of teaching formal aspects of software development, cf. (Korecko et al., 2014). The toolset focuses on Petri nets and B-Method and visual representation of a train schedule example. This approach can stimulate abstract reasoning skills, but besides abstract reasoning skills, students require a so-called computational thinking (Wing, 2006b), a mental activity in formulating a problem to select a computational solution. The key attributes of computational thinking, as introduced by Wing, are (i) reformulating a complicated problem into a problem (or set of problems) which is already known and which we are able to solve; (ii) using abstraction and decomposition when analysing and decomposing a large complicated task or designing a complex system; and (iii) choosing an appropriate representation for the problem and/or modelling the relevant aspects of a problem to make it analysable.

The KeY-Hoare tool, cf. (Bubel and Hähnle, 2008), was also developed to teach Hoare Logic. KeY-Hoare is based on a variant of Hoare logic with explicit state updates which allows one to reason about correctness of a program by means of symbolic forward execution.

Sznuk and Schubert developed a tool for teaching Hoare Logic, HAHA (Hoare Advanced Homework Assistant), cf. (Sznuk and Schubert, 2014). To estimate the impact that introduction of a tool has on the educational process, they used statistical methods of quantitative psychology (Trierweiler and Stricker, 1998). In contrast to KeY-Hoare, HAHA is based on classical Hoare logic. Classical Hoare logic requires backwards reasoning, which can be argued to be less natural and harder to learn.

Another tools used in educations were Why3 (Filliâtre and Paskevich, 2013), Dafny (Leino, 2010), as well as proof assistants Isabelle(Nipkow et al., 2002), and Coq proof (Henz and Hobor, 2011). Why3 is a tool for deductive program verification based on the WhyML language, which is an intermediate language in verifiers for Java, C, and Ada programming languages. Dafny can be used to verify functional correctness and termination of sequential, imperative programs.

4 CONCLUSIONS

Despite the impressive volume of work on teaching FM, this field still lacks systematization; it is easy for educators to get lost in the “jungle” of the proposed methods and tools. This position paper presents our ongoing work in providing a “jungle map” to teaching FM, i.e., systematically reviewing the variety of approaches to teaching FM, taking into account the aims of the course, its target audience, its respective mathematical background and motivation.
REFERENCES


