Do-Form: Enabling Domain Experts to use Formalised Reasoning

Contents

1 Motivation

Do-Form is motivated by the long-term **vision** of making information systems dependable. In the past even mis-represented units of measurement caused fatal **engineering** disasters. In **economics**, the subtlety of issues involved in good auction design may have led to low revenues in auctions of public goods such as the 3G radio spectra. Similarly, banks' value-at-risk (VaR) models – the leading method of financial risk measurement – are too large and change too quickly to be thoroughly vetted by hand, the current state of the art; in the London Whale incident of 2012, JP Morgan claimed that its exposures were $67mn under one of its VaR models, and $129 under another one. Verifying a model's properties requires formally specifying them; for VaR models, any work would have to start with this most basic step, as regulators' current desiderata are subjective and ambiguous.

We believe that these problems can be addressed by representing the knowledge underlying such models and mechanisms in a formal, explicit, machine-verifiable way. Contemporary computer science offers a wide choice of knowledge representation languages well supported by verification tools. Such tools have been successfully applied, e.g., for verifying software that controls commuter rail or payment systems. Still, **domain experts** without a strong computer science background find it challenging to

choose the right tools and to use them. Do-Form aims at investigating ways to support them. Some problems can be addressed now, others will bring new challenges to computer science.

2 Topics

The call for papers listed the following topics of interest:

- for **domain experts**: what problems in application domains could benefit from better verification and knowledge management facilities? As possible fields we suggested:
  - Example 1 (economics): auctions, VaR, trading algorithms, market design
  - Example 2 (engineering): system interoperability, manufacturing processes, product classification
- for **computer scientists**: how to provide the right knowledge management and verification tools to domain experts without a computer science background?
  - wikis and blogs for informal, semantic, semiformal, and formal mathematical knowledge;
  - general techniques and tools for online collaborative mathematics;
  - tools for collaboratively producing, presenting, publishing, and interacting with online mathematics;
  - automation and computer-human interaction aspects of mathematical wikis;
  - ontologies and knowledge bases designed to support knowledge management and verification in application domains;
  - practical experiences, usability aspects, feasibility studies;
  - evaluation of existing tools and experiments;
  - requirements, user scenarios and goals.

3 Programme

The symposium is designed to bring domain experts and formalisers into close and fruitful contact with each other: domain experts will be able to present their fields and problems to formalisers; formalisers will be exposed to new and challenging problem areas. The programme combines talks, system demos, and tutorials to ensure close interaction among participants from both sides.

For the tutorials we are delighted to have three world-class economists, who will be presenting the following domains:

- **Matching markets** (**M. Utku Ünver**, Boston College): These include matching students to schools, interns to hospitals, and kidney donors to recipients. See the documentation for the 2012 Nobel Memorial Prize in Economic Sciences for more background information.[1]

- **Auctions** (**Peter Cramton**, University of Maryland): Peter has been working on auctions for Ofcom UK (4G spectrum auction), the UK Department of the Environment and Climate Change, and others – and most recently on the "applicant auctions" for the new top-level Internet domains issued by the ICANN.

- **Finance markets regulation** (**Neels Vosloo**, Financial Services Authority, UK): It is currently impossible for regulators to properly inspect risk management models. Test portfolios are a promising tool for identifying problems with risk management models. To what extent can techniques from mechanised reasoning automate some of the inspection process?

Further information about the speakers and their tutorials is available from the symposium homepage.

4 Submission and Review Process

We ran a novel two-stage submission process:

Stage 1: problem & tool ("nail & hammer") descriptions to be reviewed and matched with each other

Stage 2: regular extended abstracts, with a normal conference-like peer review

In stage 1 we solicited ...

- from **domain experts**: descriptions of canonical models and problems in their domain that might benefit from better verification and knowledge management facilities. Descriptions should focus on aspects of these models that domain users find particularly problematic, and suspect might be aided by formalisation tools

- from **computer scientists**: descriptions of formalisation, verification and knowledge management tools, with an emphasis on how they could be applied in a concrete real-world setting, or tailored to such application domains.

The symposium chairs, assisted by the PC members, reviewed and initially published commented versions of the stage 1 submissions on the symposium homepage (reproduced below), to provide orientation for stage 2. Where we identified matching problems and tools, we notified the respective authors. Excluding one withdrawal, we received seven stage 1 submissions, out of which two classified as nails, three as hammers, and two covered both aspects.

[1] `http://www.nobelprize.org/nobel_prizes/economics/laureates/2012/`

In stage 2, with a later deadline, we solicited regular submissions on any of the topics outlined initially. We encouraged submissions that specifically addressed topics identified in stage 1; for a tool description paper, this could, e.g., have been done by motivating the tool with a stage 1 problem, and sketching how the tool could, or will, be applied in this domain. We also encouraged the stage 1 authors to revise and expand their initial submissions. The referees were advised to judge submissions based on the PC's views of the likelihood of contributing to a better matching of hammers (formalisation and verification tools) to nails (domain problems).

We invited research and position papers, as well as tool and system descriptions, and finally formalised knowledge representations with human-readable annotations.

In addition to the revised and expanded versions of the stage 1 submissions, we accepted four new submissions in stage 2. These included two tool and system descriptions, one description of a formalised knowledge representation and one research paper.

The submission and review process was supported by the EasyChair system.

5 Matchmaking

To provide orientation for submission stage 2, we published for each stage 1 submission a summary of their hammer/nail aspects and gave advice on how they could be matched.[2] Several authors followed this advice and made explicit references to stage 1 submissions in their papers, using the submission numbers given below.

Submission 1 (nail)

SAsSy – Scrutable Autonomous Systems (Nava Tintarev, Nir Oren, Roman Kutlak, Matt Green, Judith Masthoff, Kees van Deemter and Wamberto Vasconcelos)

Failure to understand complex (e.g. distributed) autonomous systems may lead to unrealistic expectation or lack of trust (e.g. in dangerous environments), and thus limits their adoption. The authors discuss the importance of making them scrutable, i.e. adding to them a mechanism that can explain the alternatives for actions and its decisions, so that it can communicate with humans in an understandable way on a level that allows users to cope with potentially large amounts of data. Making autonomous systems scrutable rests on four foundations:

1. a representation for planning
2. human-understandable reasoning mechanisms
3. translating logical statements into natural language
4. techniques that enable the user to cope with a deluge of information

Each of them requires formalisation and has its own unique challenges, as does the integration of all of these into a single system.

[2] Most stage 1 submissions have subsequently been revised for stage 2.

The SAsSy project aims at developing a hammer for this nail but is in an early stage and could therefore (re)use existing hammers. Respondents to this work can be found on two different ends: non-trivial applications as well as contributions to the different parts (1–4). The authors have so far identified the following possible hammers, and studied relevant literature:

- 1. knowledge acquisition
- 1. and 2. argumentation theory
- 2. and 3. natural language generation
- 4. diagrams, user modelling, user evaluation

Submission 2 (nail and hammer)

Transparency of Environmental Computer Models (Martine de Vos, Jan Top, Willem Robert van Hage and Guus Schreiber)

The environment is complex to model: long-term processes, complex mechanisms, lots of variability, not everything well understood, lack of empirical data. Current computer models of this, as well as the process of their development, lack transparency: What is incorporated in the model and what not, how trustable are the results? This restricts their applicability and reusability by outsiders/non-developers. The following is needed to make them more transparent:

1. make their underlying conceptual model explicit
2. encourage model developers (i.e. domain experts) to do (1), using formalisation
3. encourage them by incentives

The following previously existing solutions are not yet completely sufficient:

- teaching best practice to modellers (when they don't see an urgent need for transparency, while society has this need)
- high-level annotation of scientific models and workflows, and provenance annotation of scientific data

The authors present the approach of explicitly describing the conceptual models themselves (concepts and relations) as ontologies, and using the latter to

- facilitate reviews of the model
- communicate the model and its underlying scientific knowledge

Respondents to this work can help to address requirements (1–3), or apply ontologies as a means of review and communication in other application settings.

Submission 3 (hammer)

A Vision of Collaborative Verification-Driven Engineering of Hybrid Systems (Stefan Mitsch, Grant Olney Passmore and Andre Platzer)

Hybrid systems with both discrete and continuous dynamics are an important model for real-world physical systems. The key challenge is how to ensure their correct functioning w.r.t. safety requirements. Promising techniques to ensure safety seem to be model-driven engineering to develop hybrid systems in a well-defined and traceable manner and formal verification to prove their correctness, forming the vision of verification-driven engineering. Despite the remarkable progress in automating formal verification of hybrid systems, the construction of proofs of complex systems often requires significant human guidance, since hybrid systems verification tools work over an undecidable theory and cover different fields of mathematics. It is thus not uncommon for verification teams to consist of many players with diverse expertise (e.g. on several fields of mathematics, verification, machine-supported theorem proving, etc.) This paper presents a verification-driven toolset for collaboratively engineering large-scale hybrid systems, which supports

- modeling hybrid systems
- exchanging and comparing models and proofs, and
- managing verification tasks

Particular strenghts include

- decomposition of hybrid systems (so that they can be verified by verifying properties of their subsystems),
- efficiently solving multivariate polynomial inequalities (by integrating existing provers; work in progress),
- collaborative development of models and proofs (in a generic language and in domain-specific ones)

Respondents to this work can offer new application domains.

Submission 4 (hammer)

Interacting with Ontologies and Linked Data through Controlled Natural Languages and Dialogues (Ronald Denaux, Vania Dimitrova and Anthony Cohn)

Ontologies play an important role in many different appication areas. Although the languages in which they can be speficied look superficially simple, domain expert have problems to apply them appropriately. This paper presents three steps to enable domain experts to build ontologies:

1. Rabbit, a controlled natural language frontend for OWL, which has been successfully tested with domain experts in cartography

2. ROO, an editor that helps domain experts to avoid modelling mistakes by guiding them through the ontology authoring process

3. a facility that enriches this editor with interactive feedback on logical consequences of new facts to be added to an ontology, i.e.:

   (a) the fact already is in the ontology

   (b) can be derived from it,

   (c) contradicts it,

   (d) is new (and if so whether anything follows from it)

   So far this makes authors aware of problems, but does not really yet help to resolve them.

4. a dialogue-based user interface for giving more differentiated feedback while the author is building an ontology (but useful beyond ontology authoring)

Respondents to this work could evaluate whether the tools presented here are really independent of the application domain, or try to scale the techniques to more expressive logics.

Submission 5 (hammer)

Explaining the Outcome of Knowledge-Based Systems; a discussion-based approach (Martin Caminada, Mikołaj Podlaszewski and Matt Green)

Nonmonotonic reasoning is a framework for analysing the construction of arguments based on rules of thumb which are subject to exceptions. This submission proposes such a discussion-based approach for explaining the outcome of knowledge-based systems that use nonmonotonic (defeasible) inference. An interactive Python-based frontend is available, which aids in the process of constructing arguments. A key feature is the ability not only to establish an argument, but also to give a (one hopes) explanation of the justification of the argument suitable for human users to digest.

Respondents could apply this technique in their respective domains; it seems particularly promising in the social sciences.

Submission 6 (hammer and nail)

Developing an Auction Theory Toolbox (Christoph Lange, Colin Rowat, Wolfgang Windsteiger and Manfred Kerber)

Auctions allocate trillions of dollars in goods and services every year. Auction design can have significant consequences, but its practice outstrips theory. The authors aim at advancing auction theory with help from mechanised reasoning. To that end they are developing a toolbox of formalised representations of key facts of auction theory, which will allow auction designers to have relevant properties of their auctions machine-checked. In the starting phase of this effort, they are investigating the suitability of different mechanised reasoning systems (Isabelle, Theorema, and TPTP) for

reproducing a key result of auction theory: Vickrey's celebrated 1961 theorem on the properties of second price auctions – the foundational result in modern auction theory. Based on their formalisation experience, they give tentative recommendations on what system to use for what purpose in auction theory, and outline further steps towards a complete auction theory toolbox.

Respondents could apply a similar toolbox building methodology in other domains, or could contribute additional mechanised reasoning know-how to the authors' auction theory toolbox building effort.

Submission 7 (nail)

Model Validation and Test Portfolios in Financial Regulation (Neels Vosloo)

Modern finance relies on software, whether to price assets ('valuation models') or to compute portfolios' riskiness ('capital models'). Given the quantities of money involved, errors in financial software can lead to tens or hundreds of millions of dollars of losses. At the same time, model validation and checking is typically performed using traditional, manual methods. This constellation of observations has led to high-level concern about whether the modern financial system is overly reliant on computational models, whose possibilities for failure are not well understood or controlled (q.v. the 2012 Beddington/Foresight report).

In the midst of this clearly very large domain area, Vosloo outlines a tractable starting point for the consideration of formal methods/mechanised reasoning: how can regulators develop minimal 'test portfolios', capable of efficiently ensuring that capital models incorporate relevant risk factors, benchmarking those capital models against each other, and stress testing them under a variety of market conditions.

Respondents could outline how to apply verification techniques to this problem.

6 Organisation

Do-Form had the following programme committee:

1. Bill Andersen, Highfleet, Baltimore, US
2. Rob Arthan, Lemma 1, Reading, UK
3. Christoph Benzmüller, Mathematics and Computer Science, Free University of Berlin, Germany
4. Peter Cramton, Economics, University of Maryland, US
5. James Davenport, Computer Science and Mathematical Sciences, University of Bath, UK
6. Michael Grüninger, Mechanical and Industrial Engineering, University of Toronto, Canada
7. Manfred Kerber, Computer Science, University of Birmingham, UK (**co-chair**)
8. Michael Kohlhase, Computer Science, Jacobs University Bremen, Germany
9. Christoph Lange, Computer Science, University of Birmingham, UK (**co-chair**)
10. Till Mossakowski, DFKI and University of Bremen, Germany
11. Colin Rowat, Economics, University of Birmingham, UK (**co-chair**)
12. Todd Schneider, Raytheon, Sterling, VA, US

13. Richard Steinberg, London School of Economics, UK
14. Geoff Sutcliffe, Computer Science, University of Miami, US
15. Theodore L. Turocy, Centre for Behavioural and Experimental Social Science, University of East Anglia, UK
16. Makarius Wenzel, Formal Testing and System Exploration, University of Paris Sud, France
17. Wolfgang Windsteiger, RISC / JKU Linz, Austria

The three chairs are running ForMaRE, an effort to apply Formal Mathematical Reasoning in Economics[3]. ForMaRE is supported by EPSRC grant EP/J007498/1 ("Formal Representation and Proof for Cooperative Games: A Foundation for Complex Social Behaviour").

Do-Form web page: `http://www.cs.bham.ac.uk/research/projects/formare/events/aisb2013/`

[3] `http://www.cs.bham.ac.uk/research/projects/formare/`