Corrections to *Predicting SMT solver performance for software verification* based on examiners' comments

Andrew Healy

January 7, 2017

0 General comments

0.1 While the work focuses largely on machine learning, this is not reflected well in the overall research hypotheses, which feels both too general and difficult to evaluate (e.g. efficient with respect to what?). The overall hypothesis should be changed to better reflect the machine learning focus.

0.2 The use of a pre-solver makes it hard to judge how much contribution the actual machine learning makes. A discussion on the impact of the pre-solver should be provided.

0.3 Further commentary on the selection and extraction of features is needed. Motivation for the approach taken should be clearly outlined and justified. Examples of alternative features that could have been used should be provided and the reasons for exclusion should be described.

0.4 Small numbers should be spelt in words. If the number is at the start of a sentence then spell as a word. Fix overhanging words throughout.

Fixed throughout.

1 Chapter 1

1.1 Sec. 1.1.3: last paragraph: rephrase (2 nd sentence needs to be broken up).

1.2 Sec. 1.2, I am not convinced by the hypothesis, as it is hard to measure. Can you say beyond a single prover or random prover maybe? I would also bring in machine learning as this is in my part the main contribution of this work.

1.3 Sec. 1.2, para 2: A PO my → A PO may

Typo fixed.

2 Chapter 2

2.1 Sec. 2.1, Remove we before stated (second last line)

Removed clause. Sentence now begins: *"It is Why3's driver-based..."*.

2.2 Sec. 2.1, Q1 (First sentence): starting with "integrating..": rephrase as sentence makes no sense. Fix overhang. Change Why3 is clearly, to Why3 is arguably

Broke first sentence into two sentences. Beginning with the clause *"For users of SV tools,.."*.

2.3 Sec. 2.1.1, You could also mention TIP: Tons of Inductive Problems (which is used within the automated induction community). There was a benchmark paper at CICM 2015 that you may want to cite.

2.4 Sec. 2.1.1, change a SMT-LIB to an SMT-LIB. Insert has before a wider scope.

Fixed.

2.5 Sec. 2.1.1, You need to explain "function point" and not just cite it. It gets a bit hard to read as a standalone document if you only give citations.

2.6 Sec. 2.1.1. (first paragraph P13): You can also write specification in proof script: the procedural tactic applications are just part of it, so this needs rephrasing. Remove the word being line 2.

2.7 Sec. 2.2: It wasn't clear why you are discussing software metrics to this details, so a bit more motivation would be beneficial

2.8 Sec. 2.2.1 P14: change journals to journal.

Fixed.

2.9 Sec. 2.3: Although not as close to your work as others, I believe that a literature review should be broad and shallow, so I would also expect other approaches to learning. One example is tactic learning, e.g. (i) Automatic Learning of Proof Methods in Proof Planning, (ii) Hazel Duncan's PhD thesis (Edinburgh Uni), (iii) Typed meta-interpretive learning for proof strategies (which I was involved in). You could also mention the Aris project in Maynooth, which uses case-based learning.

2.10 P15: fix overhang

Included in Comment 0.4 of this document.

2.11 Sec. 2.3.2: Here you discuss/compare concepts that haven't been introduced, with forward references to the place they are introduced (e.g. section 4.2.5). I'd suggest that you delay comparison to when you actually introduce the concept. Another example is that you talk about k-nearest neighbour without having introduced it. Use lowercase k for kNN.

2.12 Sec. 2.4 P19: Rephrase second last sentence.

3 Chapter 3

3.1 P20: Sentence 1 needs closure or remove it and make the point as part of the next sentence.

3.2 Sec. 3.1.1 P22: I didn't understand why you don't see POs from Atelier-B to be software verification related? It is after all used to develop software. Do you only consider code-level verification? Please clarify/expand

3.3 Sec. 3.1.2: Correct spelling of university.

Typo fixed.

3.4 Sec. 3.1.2 P23: Insert a dash between non commercial.

Fixed.

3.5 Sec. 3.2 P24: I didn't understand why you used POs and lemmas. Why lemmas? Won't a PO be generated from it? It may just be a matter of clarifying terminology.

3.6 Sec. 3.2: Change predicated to predicates

Typo fixed.

3.7 Sec. 3.2.1 P25: In my experience feature selection is the key for successful learning so I was hoping for a bit more details of why you chose these features. Did you for example consider semantic properties (such as associativity of operators, inductively defined types, ...)?

3.8 Sec. 3.2.1: P25: You need to explain how to calculate McCabe's cyclomatic complexity, or at least reference the section where this is explained.

3.9 Sec. 3.2.1: P25: Note that a lemma asserts and does not check.

Terminology changed: *"...lemma checks that..."* → *"...lemma asserts that..."*.