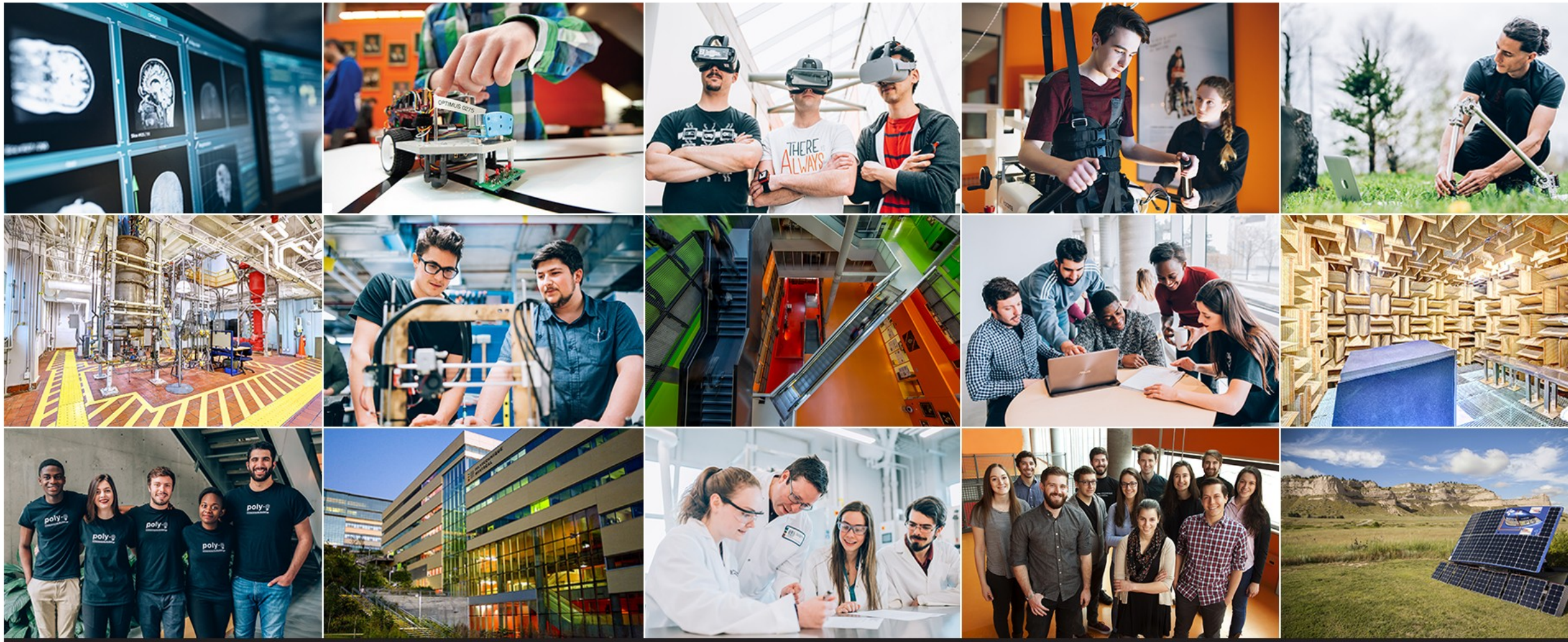


POLYTECHNIQUE MONTRÉAL

TECHNOLOGICAL UNIVERSITY

INF[67]900E Lecture 3 – Reviewing

Dr. Bentley James Oakes

bentleyjoakes.github.io

Core Ranking Improvement

Source: CORE2021
Rank: A
Field Of Research: 4612 - Software engineering
Asked for A, uprank to A (Data 1) (Decision)

Source: CORE2020
Rank: C
Field Of Research: 4612 - Software engineering

http://portal.core.edu.au/conf-ranks/1181/

This Lecture

1. Deliverables
2. Review process
3. Review structure
4. Review principles

Deliverables

February 8th

- One/two page critical review of a paper
- Template on Moodle
- Evaluated on being **constructive**, **specific, professional, structured**
  - 0 to 3 points for each category

February 15thth

- One page evaluation of another student's review
- Template on Moodle
- Evaluated out of 5 marks on having **lessons present, lessons missing, general feedback**

Review Process

Why do we want reviewing?

Papers in literature and presented should be (among other qualities):

1) Sound
2) Significant
3) Novel
4) Verifiable/transparent
5) Easy-to-understand/appealing

Reviewing helps:

- Decide which articles **worth reading and presenting**
- Articles become **higher quality**
- **Authors/reviewers** become **better researchers**

Why should we as a researcher review?

- Assist community
- Build reputation
- Win *best reviewer awards*
- Discover interesting work
- Read about new topics/approaches/styles

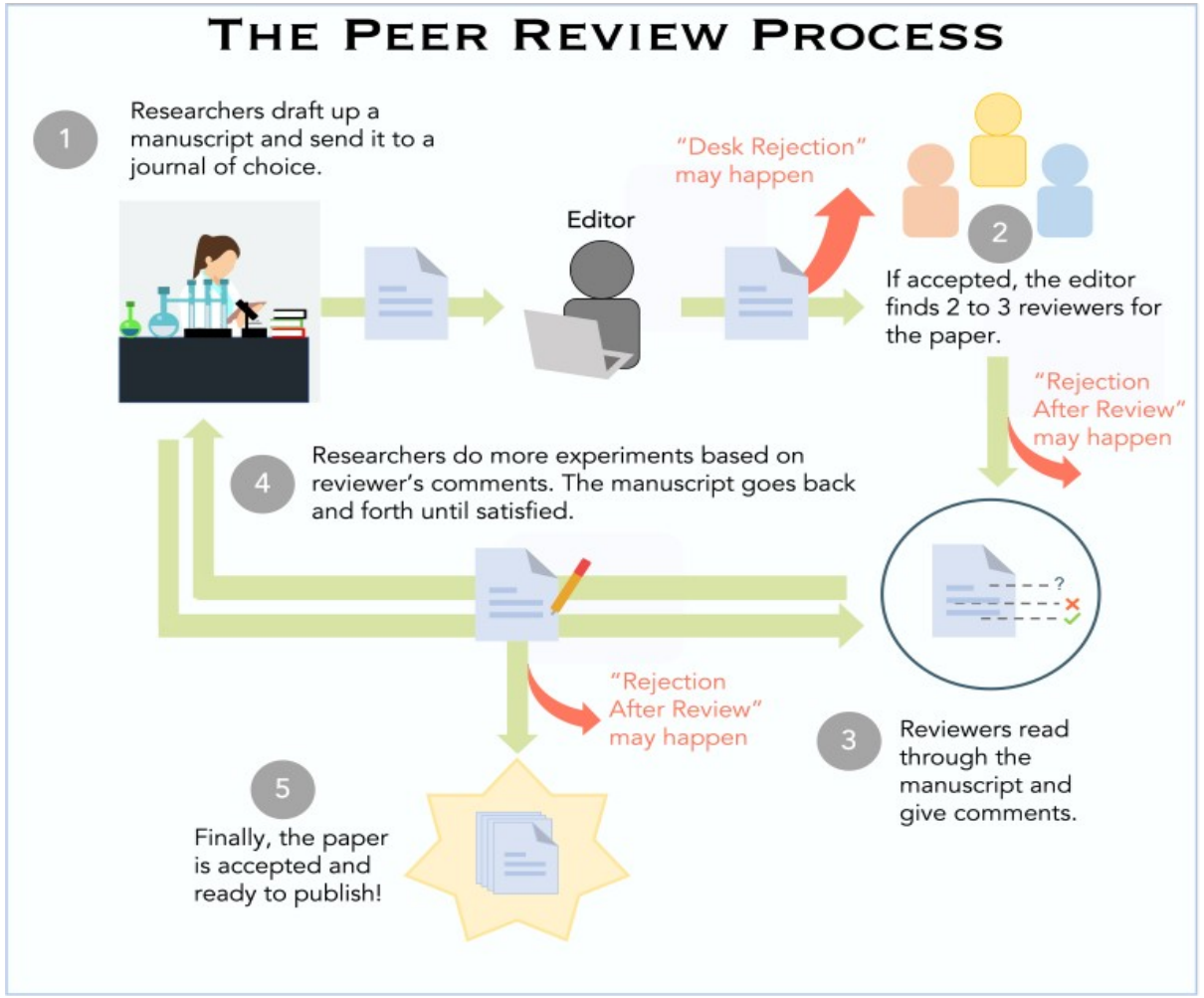
Best way to learn writing papers
is to critically review papers

Journal Submission Process

1) Submission
2) Sent to reviewers
3) Iterations:
    1) Reviewers review and recommend
    2) Reject, major revision, minor revision, accept
    3) Authors revise and resubmit, with response letter
4) If accepted, published

Long process: 6m, 1 or 2 years

But, paper will get stronger



https://sitn.hms.harvard.edu/flash/2022/peer-review-in-science-the-pains-and-problems/

Conference submission differences

- Has *program committee* (PC) = reviewers
- Shorter time-frame
- Abstract submission one week before, for PC *bidding*
- Usually one round, can be two rounds, can have rebuttal period
- Decisions are different:

  - 5. **Strong accept, award quality -** this paper should be accepted and it is a good candidate for a distinguished paper award
  - 4. **Accept** – this paper should be accepted
  - 3. **Weak accept** – this paper may be accepted, but I will not fight for it
  - 2. **Weak reject** – this paper may be rejected, but I will not fight against it
  - 1. **Reject** – this paper should be rejected

Single versus double-blind reviewing

As stated in the call for papers, submissions are supposed to be **sufficiently anonymous** that a reader cannot determine the identity or affiliation of the authors.

The main purpose of the doubly-anonymous reviewing process is to **reduce the influence of potential biases on reviewers' assessments.** You should be able to review the work without knowing the authors or their affiliations.

Single-blind: The papers contains the details of the authors.
Pros: Easier to do as nothing needs to be hidden
Cons: Bias by reviewers against authors (personally/professionally), their institutions/countries

Double-blind: Paper has no details about the authors
Pros: Removes the bias
Cons: Difficult to do, can lead to awkwardness (how to mention prior work?)

Review Principles

Reviewing Principles

Edit your review, making it as constructive and clear as possible. Even a very negative review should be respectful to the author(s), helping to educate them. - ICSE guidelines

My perspective:

Pretend a senior colleague asked you for feedback and help on a paper, what would you say?
Be:

- **Respectful** – Address the content, don’t be harsh.
- **Helpful and constructive** – Mention what you don’t like and offer suggestions, be detailed
- **Quick** in responding – Return the review quickly

Ethics in reviewing

- Don’t share papers (!)
- Don’t try to publish your work first
- Don’t recommend your papers to cite (unless extremely relevant)

- Mention *conflicts of interest* to editor
  - Friends, family, co-authors from past 5 years, same institution

- Check for (self-)plagiarism from past/other works
- Check for unethical research

How to review

- Read the paper to get a sense of it (1 hr.)
- Read it in detail and write major/minor comments (2 hr.)
- Write the summary/pros/cons/recommendation (1 hr.)
- The next day, read the review again (0.5 hr.)

Don’t rush a review!

Review Structure

Overall merit *

Present on reviews matching "round:Reviews OR round:Reviews-artifacts-check"

Please rate the overall quality of the manuscript. If you use (1) Reject or (4) Accept, then you are ready to argue against or in favor. Use (5) Strong accept if you want to nominate this as a distinguished paper

- 1. Reject
- 2. Weak reject
- 3. Weak accept
- 4. Accept
- 5. Strong accept (award quality)

Reviewer expertise *

Present on reviews matching "round:Reviews OR round:Reviews-artifacts-check"

- X. I am an expert on this topic (know the related work well)
- Y. I am knowledgeable on this topic
- Z. I am an informed outsider

Confidential Comments (authors will not see these comments)

Public Comments (these will be made available to the author)

Overview
req Which category describes this submission?
✓ Practice / Case Study / Experience Report
Technology / Tool
Research
Survey
Other
req Please rate the submission
Excellent
Good
Fair
✓ Poor

Please address the following questions/concerns in your review:
1. Are the title, abstract, and keywords appropriate?
2. Does the introduction state the objectives of the submission in terms that encourage the reader to read on?
3. How relevant is this submission to the readers of this journal? The target audience of the journal are practitioners and researchers from industry and academia with a vested interest in high quality modeling practices and research. Indicate the extent that the paper will be relevant to this target audience.
4. How does this submission advance the field of software and system modeling research and practice? Comment on any novel contributions or significant insights gained. The journal aims to publish papers that deepen understanding of modeling practices and techniques or contribute significant new ideas that revolutionize or incrementally advance the field.
5. Is the submission technically sound? For example, comment on (1) adherence to standards if standard notations/techniques/methods are used, (2) soundness of mathematical expressions, and (3) soundness of conclusions drawn from objective premises.
6. Does the submission contain sufficient and most appropriate references? Journal versions of work are preferred over conference versions. Indicate important missing references, if any.
7. Comment on the organization of the submission. Is it focused? Is the length appropriate for the topic?
8. Please comment on the readability of this submission. Please comment on the degree of effort required to read and understand this paper.

Paper Review Template (for deliverable)

Overview
Summarize the paper in a few sentences. Lets a) the editor know what the paper is about, and b) the author know that the reviewer read it properly

Pros
For editor, why to accept. For authors, the strong points

Cons
For editor, why to reject. For authors, the weak points

Comments/Recommendation
Overall, why you chose your recommendation

Major Comments
For each section of the paper:

- Your thoughts, what can be improved, your suggestions
- Unclear text, missing related work

- **Minor Comments**
- Typos/grammar issues
- Basically, anything that can be fixed in five seconds

Summary: “This work proposed a [...] with [...] for extracting both the structural and functional connectivities from fmri data, it is very interesting work since a few works has been working on both the structural and functional connectivities patterns on this field. However, I would like to see the discussion of this work on how to expand to dynamic brain network on both the structural and functional patterns.”

Strength: “as above”

Weaknesses: “as above”

Recommendation: “accept”

AC cannot use the review and make any decision without reading the paper

Summary: “This paper proposes a [...] to combine generic keypoint and CNN information into a single, highly efficient memory-based model for indexing and classifying generic 3D medical image data.”

Strength: “none”

Weaknesses: “- no novelty according to a conference as MICCAI - no well written, so many English errors - only 1 expert on each dataset”

Recommendation: “reject”

Judgements are not supported by any arguments

Reviewer: 1

Public Comments (these will be made available to the author)
Summary

The paper presents a method for the formal verification of pre/post-condition contracts in model transformations. Transformations must be defined using the declarative subset of ATL. Given that verification uses symbolic execution, results are valid for any potential input model. When the contract is not fulfilled a sample input model is provided as a witness. The approach is assessed quantitatively and qualitively in several case studies, where some optimizations (i.e. automatic transformation slicing) are applied to speed up verification.

Comments

The paper title is adequate, and the abstract and keywords capture the important aspects of the submission. The introduction motivates the problem and summarizes the contribution. It also identifies that this paper is an extended version of a MODELS 2015 paper and highlights the new contributions in this submission (which in my opinion justify this extension).

The topic addressed by this paper (quality of model transformations) fits within the scope of SoSyM. Furthermore, it consider one of the most widespread languages (ATL) and it covers a sufficient subset of the language to support realistic transformations. Therefore, I think the paper will be relevant to any reader working with ATL or studying the quality of model transformations.

STRENGTHS:
1- The major strength of this submission is the ability to support most features in declarative ATL (except for using blocks, which can be rewritten anyway). Thus, this method seems to be applicable in a wide variety of transformations.
2- Moreover, experimental results show reasonable execution time and memory usage.
3- Thanks to symbolic execution, verification results hold for any input model.

WEAKNESSES:
1- The definition of the approach is not very formal. There is no formal proof of the equivalence of the original ATL transformation and the DSLtrans notation used for verification.
2- Furthermore, sections describing the mapping from ATL to DSLtrans (4.2.1 and 4.2.2) could be improved (see below).
3- Finally, a section of paragraph aggregating the shortcomings of the current prover and tool infrastructure would be welcome.

Hence, my recommendation is acceptance of the paper, provided that the issues regarding weaknesses (2) and (3) can be addressed. I believe that a formal proof of correctness is out of the scope of this venue.

Major comments

- Section 4.2.1: Rather than describing the semantics and the mapping, it mixes ATL semantics and DSLtrans semantics in the description and focuses too much on the specific Families2Persons case study. Hence, it is complex to understand (a) what is the semantics of ATL, (b) what is the semantics of DSLtrans and (c) an intuition on how ATL semantics is preserved in the HOT. This section could use some reorganization to better cover and separate these three aspects (a), (b) and (c).

- Section 4.2.2: Again, rather than describing how OCL constructs are translated, I feel that this section focuses too much on specific cases in the Families2Persons case study. A more general explanation on a systematic strategy to translate complex OCL expressions would be welcome.

- Shortcomings of the prover: Some limitations of the current implementation are mentioned throughout the paper, e.g. rule subsumption in Section 8.2, lack of support for using blocks in Table 1. A paragraph or section aggregating these limitations would be most welcome by practitioners.

Minor comments

- Section 6: There is no information on the metamodel size in the case studies, while Section 6.3.1 claims they are sufficient to cover a variety of metamodel sizes. Please include this information somewhere.

- Section 7.2: This Section mentions that the UML-to-Kiltera transformation has 17 DSLtrans rules, while Table 2 states there are 14 DSLtrans rules. I understand they are slightly different versions of the same case study, but please mention the differences among them.

Typos

- Page 3, line 22: "hold when any" = "hold for any"
- Page 7, line 24: "internal traces" = "internal trace"
- Page 9, line 33: "which receive" = "which receives"
- Page 16, line 57: "no distinguishing" = "no distinction"
- Page 20, line 33: "two different path condition" = "two different path conditions"
- Page 27, line 52: "rule that" = "rules that"
- Page 30 and 31: Problems with upper-case in several reference titles, e.g. "Dsltrans", "Syvolt", "uml-rt", "atl", "turing".

Summary: "Authors propose X, a new semantic and fully-convolutional segmentation architecture. X essentially is a U-Net with bi-directional recurrent skip connections. Compared to other recurrent U-Net architectures with gated RNN blocks, X uses existing layers and concat blocks and does not require any extra parameters. Authors validate the method on two segmentation tasks and one super-resolution task, outperforming baseline methods from literature and simpler architectures."

Strength: "- Simplicity: X's main strength is that no extra parameters are required, since the recurrence is realized directly on the layers - Extendability: The method can be applied to already existing U-Net segmentation problems with minor changes to the model architecture. Even though this is not investigated in this work, an extension to 3D segmentation should be straightforward, as no extra parameters are required. The high number of network parameters in 3D makes the incorporation of additional gated RNN architectures (GRU LSTM) particulary "costly", while X would keep the model complexity constant."

Weaknesses: "- Limited novelty: the proposed network appears like a special case of the previously proposed R-U-Net (Wang et al.), with l=0, without gates, and with a concat merging of the hidden layers/states. - Limited discussion of recurrence: in principle, authors realize a vanilla RNN directly on the hidden representations in the U-Net. Hence, training requires an unrolling of the X and backpropagation-through-time (BPTT) on the recurrence time steps, which may cause vanishing gradients (as in vanilla RNNs). Authors use very few timesteps (in this work, t=1/2/3). Larger temporal context, in combination with gating of units (as in GRU/LSTM) could further improve results, but to what degree this could be necessary/helpful, is not discussed here. - Limited comparison to state-of-the-art: Authors compare to R2-U-Net, but not to Wang et al. (R-U-Net) - No statistical evaluation of results: paired tests would give statistical weight to the argument of "superiority" of the proposed method. "

Comments:

" Lack of clarity: - Better explanation of the training stage: it would help to have a clear separation of the training and test stage. The training stage should explain unrolling of the network architecture through time (ideally accompanied by a figure), and how training is performed. - #params: Authors claim that no extra parameters are required compared to a vanilla U-Net, however, the concatenation of decode features from the previous iteration with the current iteration's encode features (i.e. the reverse direction) causes larger feature maps, which require deeper convolutional filters (i.e. more channels in each filter) and hence more parameters. This increase may be negligible in a network with 15.0M parameters, but a brief clarification would be helpful (maybe I am still misunderstanding sth). [...]

For future work, I would recommend:
- Extension to 3D: the simplicity and compactness makes this approach particularly attractive for 3D segmentation.
- Explore performance on many more problems: X could be universally applicable, but here it is used on only a few tasks. I would strongly recommend to apply X to the medical image segmentation decathlon (http://medicaldecathlon.com/). I would not expect X to end up leading the board, but it would be interesting to see whether X can actually scale to a wide variety of tasks, and especially in higher dimension (i.e. perform at least as good as an equivalent 3D U-Net on all tasks). If so, this could become an attractive alternative architecture next to U-Net in future. [...] "

Recommendation: "accept"

From ICSE 2023 guidelines – **Section 6 is required reading**

At ICSE, we evaluate papers against *five criteria*, as independently as possible.

- **Soundness** address research questions and supported by rigorous application of appropriate research methods
- **Significance** contributions beyond prior work, future implications
- **Novelty** sufficiently original with respect to state-of-the-art
- **Verifiability and Transparency** how paper supports independent verification or replication
- **Presentation**

  quality of writing, clearly readable figures/tables, clear and concise

  We recognize that not all authors are fluent English writers. But if the language issues make the paper not comprehensible, it is not yet ready for publication.

These criteria should be mentioned in your review

THANK YOU!

Topics covered:

- Reviewing process
- Reviewing principles
- Review structure

POLYTECHNIQUE MONTRÉAL

TECHNOLOGICAL UNIVERSITY

Bentley.Oakes@polymtl.ca

https://bentleyjoakes.github.io/