Methods in AI Research 16

Assignment part 2

Chris Janssen & Rosalie Iemhoff

Criteria

- **Graded on 5 point scale:**
  1. **Insufficient.** This would lead to a fail.
  2. **Barely sufficient**
  3. **Sufficient.** We would expect the typical student to achieve this level.
  4. **Above standard**
  5. **Outstanding.** Meant for work that exceeds expectations by far. We expect to rarely award this.

(On next slides: C or R indicates who will mark)

Marking

A. General writing (C & R)
   - minimum number of typos and grammatical errors
   - Is writing concise?
   - Is argumentation flow correct in general?

B. Title & abstract (C & R)
   - Is the title an actual title that reflects the content/topic/research question (i.e., not “Final assignment of MAIR”)?
   - Does the abstract follow conventions?
   - Is it a concise and accurate representation of all aspects of the report (and not just an introduction)?

Marking

C. Introduction – coherence (R)

- Does the introduction form a coherent argument…
- …that leads to a clear, theoretically and practically supported motivation for the research approach?

D. Introduction – embedding in context (R)

- is the work embodied appropriately in wider set of literature on relevant topics e.g. as discussed in class (& potentially beyond that).
- For example, is it embedded in relevant work on reasoning systems?

Marking

E. System description (-)

- This will not be graded, but will be negatively graded if missing.
- Provide some description of your system in a way that is relevant to the experiment: we need to understand what your system does when used by the user
- So: keep this concise

Marking

F. Experimental methods - conventions (C)

- Are appropriate conventions followed for the methods section (see Appendixes)?
- For example, is each subsection (participants, materials, design, procedure, measurements) written down *explicitly*, *concisely* and *correctly…*
- … up to a level that readers can understand the method easily and are able to replicate the experiment.

G. Experimental methods – set-up of study (C).

- Overall, was an appropriate method chosen to address the research question?
- Was care taken in execution of the project?
- This grading criterion will also take effort & (appropriate forms of) creativity (in the design) into account.

Marking

H. Results - descriptive statistics (C)
   - Are appropriate descriptive statistics provided (e.g., mean, standard deviation, 95% CIs or other relevant and appropriate metrics)

I. Results - inferential statistics (C)
   - Are appropriate inferential statistics provided?
   - Are these reported according to conventions (no "dump of software output")
   - Are these interpreted correctly?

Marking

J. Results – visualization (C)
   - Is an appropriate visualization of results used?
   - Is this clear?
   - (for example: a bar plot with error bars, clear use of labels, appropriate sized fonts)

K. General discussion - coherence, theory, implications (R)
   - is the general discussion coherent with the introduction?
   - Are results related back to the literature in an appropriate and correct way?

Marking

L. General discussion – limitations (C)

- are limitations of the experimental set-up reported clearly?
- Are these indeed the most important limitations?

M. Conclusion (C)

- Is there a concise conclusion that answers the research question?
- Note: in the context of this course it can also be that you can not answer the research question yet (e.g., due to null results or an inconsistent pattern of results) in which case your conclusion should reflect that no definitive answer is there yet

Marking: final remarks

- **We don't expect Noble prize winning work**
  - **Whether result is "significantly different" is not a criterium**
- **We do expect well-written documents of well executed projects (within constraints of the course)**
  - **Follow conventions and requirements**
  - **Be clear and concise in your explanation**
  - **Make use of each other for proof-reading**
  - **Be critical**
- **We assume that everyone contributed (more or less) equally**
  - **In appendix of assignment: *briefly* comment on distribution of work (does not count to page limit)**

Questions?

Chris Janssen & Rosalie Iemhoff