UAI-2014
Conference on Uncertainty in Artificial Intelligence
July 23-27, 2014, Quebec City, Canada

Reviews For Paper

Paper ID	198
Title	Interactive Learning from Unlabelled Instructions

Masked Reviewer ID: Assigned_Reviewer_10

Review:

Question	
Novelty: This is arguably the single most important criterion for selecting papers for the conference. Reviewers should reward papers that propose genuinely new ideas or novel adaptations/applications of existing methods. ... (For the rest of this question, see http://auai.org/uai2014/reviewCriteria.shtml)	Standard interactive learning assumes that the system (e.g., robot) knows how to interpret the feedbacks given by humans either because they are either pre-defined or learned during a calibration phase. In this paper, the authors propose an approach to learn both the task to be performed and the meaning of the feedbacks. As underlined by the authors, this general problem has already been introduced by Grizou et al. As I am not an expert in the domain, I would have liked a longer discussion about how the approach proposed in this paper relates/compares to that of Grizou et al. Finally, the authors validate their approach on artificial data and pre-recorded EEG data.
Technical Quality: Are the results technically sound? Are there obvious flaws in the conceptual approach? Are claims well-supported by theoretical analysis or experimental results? Did the authors ignore (or appear unaware of) highly relevant prior work? ... (For the rest of this question, see http://auai.org/uai2014/reviewCriteria.shtml)	I find the formalization not easy to follow because it is not always rigorous and there were problems in the notations: - p.2, par.1, l.-1: why the order in the triplet of a history is changed? - I would have appreciated a few words on the sets of states, actions and signals. Are they finite, infinite, continuous spaces? - It would have been good to say that a task is completed by a sequence of actions, even if it seems obvious. - In (3), why D_i is "upperscripted" by \xi_t? - In the second line of (3), wouldn't it be better to write p_\hat \theta(\xi_t|D_{i-1}^\xi_t) as p(\xi_t|D_{i-1}^\xi_t, \hat\theta)? - is p(\xi_t | D_0^\xi_t) equal to p(\xi_t) introduced in (1)? - what does \xi_{1:i} mean? - In (4), what is N? Should it be n instead? Or is it M from p.2, col.2, l.2? - the first line of (4) should be explained more. Does it use (3)? - I fail to understand l_c. Isn't label affected to signal? - In (6), argmin should be min? - page 5, there should not be no underscript for weights W^\xi. - From the definition of (7)-(9), only one term of the pseudo-likelihood is used? - In Sec.4.2, \theta is first defined as a set of parameters and then after (10), it is a probability distribution.

Potential Impact and Significance: Is this really a significant advance in the state of the art? Is this a paper that people are likely to read and cite in later years? Does the paper address an important problem (e.g., one that people outside UAI are aware of)? .... (For the rest of this question, see http://auai.org /uai2014 /reviewCriteria.shtml)	The problem tackled in the paper is of great importance. However the explanation of the paper has to be improved.
Quality of Writing: Please make full use of the range of scores for this category so that we can identify poorly-written papers early in the process.	Quality of writing is marginal to ok - it needs significant additional editing
Overall Numeric Score for this Paper:	Decent paper, but may be below the UAI threshold. I tend to vote for rejecting it, although would not be upset if it were accepted.
(Optional) Additional Comments to the Authors: please add any additional feedback you wish to provide to the authors here. For example, if the quality of writing in the paper is not excellent, please provide some feedback to the authors on how the writing could be improved.	Minor comments: replace "the the" by "the" (several occurrences) add a comma after "i.e." and "e.g." - page 2, column 2, last sentence of paragraph 1: We will denote D_i to the history [We will denote D_i the history] - Section 2.2: release this assumption [-> relax this assumption] - Equation 6: the braces are missing under argmin - next line: it exists [-> there exists] - page 5, first line of paragraph 3: see eq.4 [-> see Eq. 4] - two lines below Equation 9: there is an high [-> there is a high] - page 5, last sentence of column 1: when former [-> when the former] - page 5, equation of column 2: a is both a free and bound variable

Masked Reviewer ID: Assigned_Reviewer_4

Review:

Question

Novelty: This is arguably the single most important criterion for selecting papers for the conference. Reviewers should reward papers that propose genuinely new ideas or novel adaptations/applications of existing methods. ... (For the rest of this question, see http://auai.org/uai2014/reviewCriteria.shtml)	The authors present a method for planning a MDP given only a black box simulator that can be called to generate a reward or transition. They provide high probability finite-sample bounds on the number of calls needed to compute a near-optimal action from the current state. These bounds depend on problem-specific characteristics, and can be tighter in some cases than those previously derived. In particular they depend on kappa which defines the set of states that contribute in a significant way to near-optimal policies from the initial state. Kappa will generally be unknown in advance, but their results suggest that if the actual problem has properties that make kappa small, then the resulting sample complexity may be much better than some prior results. The proposed problem seems like a natural small step from related settings. To achieve their results I think the key insight was to be able to reuse many samples to evaluate many policies. In this way the paper reminded me slightly of Pegasus (Ng and Jordan) which fixes a set of random trajectories to be used to do policy evaluation over many policies. The settings are somewhat different but it would be interesting to have this approach referenced in related literature.
Technical Quality: Are the results technically sound? Are there obvious flaws in the conceptual approach? Are claims well-supported by theoretical analysis or experimental results? Did the authors ignore (or appear unaware of) highly relevant prior work? ... (For the rest of this question, see http://auai.org/uai2014/reviewCriteria.shtml)	The results seem sound, though I have not done a detailed pass over all the proof steps. The primary contribution of this paper is theoretical. No empirical results are given and it was unclear, even given the appendix on making the algorithme efficient, how significant an advance this approach is empirically over existing approaches. Still the theoretical results are interesting. It would be nice if the authors could continue to expand on the intuition and insights that allowed them to prove their results. Some of the paper is very clearly and nicely articulated (especially the first few pages), but other parts seem too slow or fast. For example, Figure 1 and section 2.3 could be eliminated, and this would give more space to provide more intuition and details in section 3.

Potential Impact and Significance: Is this really a significant advance in the state of the art? Is this a paper that people are likely to read and cite in later years? Does the paper address an important problem (e.g., one that people outside UAI are aware of)? .... (For the rest of this question, see http://auai.org/uai2014/reviewCriteria.shtml)	Section 1.3 tries to outline the significance and impact of the proved result, relative to prior work. The primary benefit over other approaches seems to be for these interesting cases where only a subset of the states are important. But it's not clear how many domains have this structure. Some empirical results would significantly strengthen these.
Quality of Writing: Please make full use of the range of scores for this category so that we can identify poorly-written papers early in the process.	Quality of writing is good, but could be improved with some editing
Overall Numeric Score for this Paper:	A good paper overall, accept if possible. I vote for acceptance, although would not be upset if it were rejected because of the low acceptance rate.

Masked Reviewer ID: Assigned_Reviewer_6

Review:

Question	
Novelty: This is arguably the single most important criterion for selecting papers for the conference. Reviewers should reward papers that propose genuinely new ideas or novel adaptations/applications of existing methods. ... (For the rest of this question, see http://auai.org/uai2014/reviewCriteria.shtml)	The paper presents an approach for learning a task taking advantage of user feedback but considering that the meaning of the user signal is not known to the system; so both, the feedback signal and the task have to been learn simultaneously. There are several works on taking advantage of user feedback for learning a task, such as learning by demonstration, learning by instruction and reward shaping in the context of reinforcement learning; there is not much work on learning the meaning of the user signals at the same time. Thus the paper proposes a way to this using a probabilistic model and a planning paradigm based on reducing the uncertainty (variance) in both task and signal.
Technical Quality: Are the results technically sound? Are there obvious flaws in the conceptual approach?	The proposed approach is evaluated in a synthetic experiment and in a real experiment that uses EEG data. The results show that (i) using variance to guide the search is a good option (ii) self calibration works better the traditional calibration procedure used in EEG experiments.

Are claims well-supported by theoretical analysis or experimental results? Did the authors ignore (or appear unaware of) highly relevant prior work? ... (For the rest of this question, see http://auai.org/uai2014/reviewCriteria.shtml)	The method assumes a fixed number of tasks which is known a priori and a relatively small model in terms of states and actions. It also assumes that the user signal has only two values (correct/incorrect). Thus a natural question is if the proposed approach will work when these assumptions are relaxed, in particular with respect to the number of tasks and size of the problem? The EEG taks solved is not explained in the paper, so it is difficult to asses the significance of the results in this experiment.
Potential Impact and Significance: Is this really a significant advance in the state of the art? Is this a paper that people are likely to read and cite in later years? Does the paper address an important problem (e.g., one that people outside UAI are aware of)? .... (For the rest of this question, see http://auai.org/uai2014/reviewCriteria.shtml)	The proposed approach is useful for certain scenarios, in particular for brain-computer interfaces where the user signal is difficult to decode. In this context it could be a significant contribution. However, it seems unnecessary for certain domains in which it will be relatively easy to predefined the meaning of the signals used for feedback.
Quality of Writing: Please make full use of the range of scores for this category so that we can identify poorly-written papers early in the process.	Quality of writing is excellent
Overall Numeric Score for this Paper:	A very good paper, should be accepted. I vote and argue for acceptance, clearly belongs in the conference
(Optional) Additional Comments to the Authors: please add any additional feedback you wish to provide to the authors here. For example, if the quality of writing in the paper is not excellent, please provide some feedback to the authors on how the writing could be improved.	Explain the EEG task solved in the paper.