

Reviewer 6: Accept

The paper describes an algorithm, simulation and real-robot experiments in which a robot co-learns instruction-meaning association and a new task from a pre-defined possible task-space. The instructions are given via spoken words and the task space is given from a set of a finite reward distribution. The authors also describe how to integrate a-priori knowledge into the system as well as known instructions, e.g. buttons. The algorithm is very interesting and the analysis of its performance is thorough. The authors show that indeed the robot learns a new task as well as the association between the instructor's spoken words and their meaning. They also show that its efficiency can be increased by a-priori knowledge and fully known feedback. Applying learned associations to a new task is also shown to be beneficial.

However, the overall readability of the algorithm section is problematic. I was a little confused regarding n , f , z : which is the instruction, which is the meaning and what is the third one? A really short example in this stage can increase the reader's understanding.

Also, in Section II.A: (i) after equation (1), \hat{x}_i is not defined. (ii) In equation (2), the transition from second to third line is not explained [where did the dependence on s and a gone?]. (iii) Some verbal meaning to variables could help, such as $z_{ij}^{\hat{x}_i}$ and w_{ij} .

In Section II.B: G is defined but not used in equation (5), why? Where is the usage of the optimal action given?

In contrast, Section III.A is written very well and presents some very interesting ways to simplify the real space to a manageable representational space, e.g. generation of task spaces and speech-to-feature. The results section is also written well and the analysis is very informative.

Also the paper is riddled with Typos (here is a small set):

Abstract: "...whose associated unknown meaning..."

Abstract: "... the meaning of unknown and noisy ..."

P. 2, left column, second par. from end, 2nd line: "...known...", n is missing

P. 5, right column, second par from end, 4th line: Sentence begins with "To can ...".

P. 7, left column, Par. beginning with 4), 5th line from end: "...known..." n is missing.

P. 7, left column, second par from end, 3rd line from beginning: "...led..." → "... lead ..."

Reviewer 7: Strong Accept

The paper addresses an interesting problem - that of a robot simultaneously learning a task and the particular way in which a given human will provide teaching instructions. Some assumptions and provisions needed to be made; in particular the task must already be known to the robot (in the sense that that robot can, at a minimum, approximate it with tasks it already knows). I am not aware of similar work and the approach appears thorough and correct.

I have no strong criticisms of the paper in its current form given the targeted venue. The paper does however contain a large number of typos and linguistic mistakes. I would suggest that the authors have the text checked by a native English speaker.

Reviewer 9: Accept

Interesting work that proposes an algorithm to simultaneously learn the meaning of verbal instructions given by a human to a robot and the task to be performed. A very important point of this work is that it does not rely on classical speech recognisers but on a distance metric to some prototype words. Anyway, the paper could be improved in a number of ways, mainly to improve readability and better explain the assumptions that underlie the work. The authors mention several times that the algorithm is general and does not make particular assumptions on how the task and the instructions are represented. This creates a lot of expectation in the reader that soon becomes disappointed because there are hard assumptions that must be made about the algorithms: the task must be represented as a goal state in discrete state-space, the robot must know the mode of interaction (instruction vs feedback), the feedback model for each task must be known, the possible tasks are known and finite, the robot has prior knowledge about the distribution of possible tasks, the robot has planning skills, other? The authors should define very clearly, already in the introduction, the assumptions and domain of validity of their work. Also in the introduction the authors should explain the two different modes of operation: feedback or instruction. This only becomes clear very late in the paper.

Other points for improvement:

- 1 - The explanation of the algorithm (section II) should be accompanied with examples as in the provided video. Also, a graphical model describing the independence assumptions would facilitate the comprehension.
- 2 - Interaction data is provided by three signals: state (s_i), action (a_i) and instruction (n_i) but it is called a quadruplet?
- 3 - f and z are both called the "meaning" of the instruction. Please clarify?
- 4 - Task symbol (\hat{x}_i) sometimes wears a hat, sometimes don't. Why?
- 5 - In eq (2) what represents index i ? Time? Experience (combination state-action) index?
- 6 - In eq (3), what is $z_{i^{\{x_i\}}}$? Should it be $z_{ij^{\{x_i\}}}$? The same undefined symbol appears just before algorithm 1 and later in the paper.
- 7 - In the guidance case, what is the role of f ?