Guidelines for Preparing a Paper for the European Conference on Artificial Intelligence

First Author[a,*,1]**, Second Author**[b,1] **and Third Author**[b,c]

[a]Short Affiliation of First Author
[b]Short Affiliation of Second Author and Third Author
[c]Short Alternate Affiliation of Third Author
ORCID (First Author): https://orcid.org/....-....-....-...., ORCID (Second Author): https://orcid.org/....-....-....-....,
ORCID (Third Author): https://orcid.org/....-....-....-....

Abstract. >>> To be completed. . .

1 Introduction

>>> To be completed. . .

2 Related literature

>>> To be completed. . .

3 Experimental setup

In this Section we empirically explore coachable searching as an approach to preference elicitation, by having a machine learner iteratively seek advice by a proxy coach, as in [], with the aim of iteratively improving its search capacity, as per the coach's explicit or implicit search policy. Two experimental procedures, one aiming to assess the efficacy and scalability of the coaching process and one to explore a more realistic and open–ended environment.

3.1 Sorting coaching

In this series of coaching sessions we have implemented two variants of a machine coach, aiming to explain to an initial ignorant agent about how to sort a list of numbers, as a rough equivalent to linear preference elicitation. The first coach, $\mathcal{C}_{\mathrm{bubble}}$, uses bubble sort as its underlying policy while the second one, $\mathcal{C}_{\mathrm{quick}}$, uses quicksort to provide advice from. We explored different state sizes, n, ranging from 1 to 20, running $m = 100$ iterations for each value of n. We have also explored two different types of advice for each coach: (A1) *full advice*, where each piece of advice depends on the entire state, and; (A2) *partial advice*, where each piece of advice depends on the specific part of the state to be improved. Moreover, for each advice type we have considered three learner configurations regarding advice memory: (C1) *no memory* across different iterations for the same state size, n, i.e., coaching sessions are pairwise independent; (C2) *short memory* across different iterations for the same state size, n, i.e., coaching sessions across different values of n are independent but not within the same value for n, and; (C3) *long memory*

across all values of state size, n, so all coaching sessions are correlated, building on knowledge from previous test cases.

In Figures 1 we present the number of coaching interactions required to reach the goal state in each case as a function of the state size, n. Regarding the full advice condition (Fig. 1a), the utilization of either short of long term memory does not appear to have any particular effect on the number of coaching interactions, which was expected given the overspecific nature of provided advice, being fully tailored to the state it triggered the learner to seek for that piece of advice. On the contrary, partial advice (Fig. 1b) offer a significant advantage to the learner, in both short and long memory conditions. As for the short memory case,
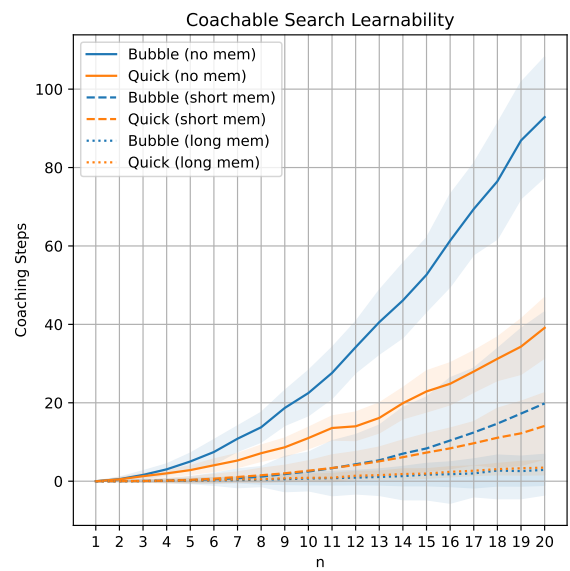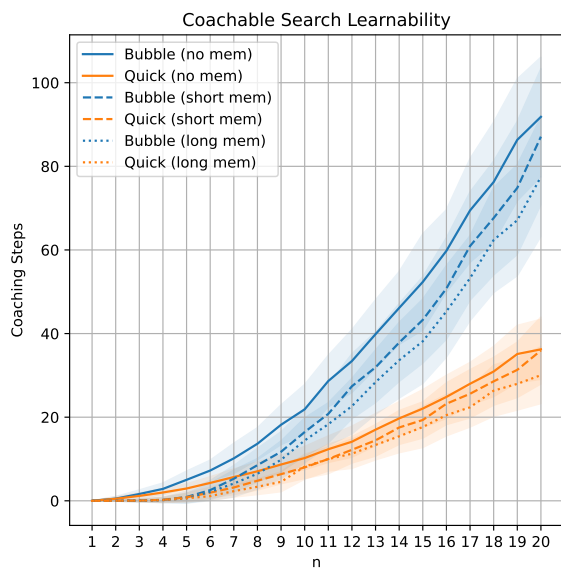
4 Results

>>> To be completed. . .

5 Conclusions and future work

>>> To be completed. . .

Add a few words about how it appears the text searching, i.e., how from a sub-state coach vides piece of advice corre- to (im- plicit) ex- policy that im- pro the state to- wards a(n im- plicit / ex- plicit) goal.

* Corresponding Author. Email: somename@university.edu.
1 Equal contribution.

(a) >>> To be completed. Figure 1: >>> To be completed... (b) >>> To be completed...