.1

Framework Specification

In this appendix, we will start to analyse questions related to data acquisition in these scenarios. In total there are three aspects related to this analysis.

Firstly, how to collect data is an important question. For example, in first space invader example, robot can only interact with human when one session is finished. This brings us a question that is the data collected enough for the training of our multi-armed bandit algorithms ? As a consequence, a critical analysis is more than needed. Simultaneously, what kind of data should be collected should also be considered carefully. Several questions need to be raised here. Can we learning some interesting behaviours of human in this process ? is the data of high quality ? is the data user-specific ? is collection process of data repeatable ? These analyses are crucial for our selection of scenarios.

Finally, when to collect data is an art itself. In our vision, robot is not a simple machine that tries to follow whatever is happening in its surroundings especially when it is working with human. A robot with social intelligence should be able to interact with human diplomatically. That is it knows when to show respect and when to acquire the knowledge it needs.

We will talk about following scenarios in consideration of previous three aspects and we also need to analyse how these three aspects are considered in other cases/papers.

Two extra subtopics need to be discussed for each scenario. One is about the learning frameworks of each scenario and another one is how should measure learning performance of human. These topics are important because the first question is related to how should we use the data and second one is related to performance of our algorithms.

.1.1

Scenario one: Space Invader

In this scenario, robot need to learn supportive strategies in order to maximize the users feelings and improve the efficiency while learning. A model of multi-armed bandit algorithms should be considered. In this scenario, we would like to include context and consider contextual algorithms.

Regarding the learning framework, we will start with a learning algorithm called Exp3 and extend this algorithm with contextual information. At first stage, we propose a hypothesis that user has a stationary model and we use the affective feedback of the user to be the feedback of the algorithm. More concretely, the algorithm execute an action $a \in A$, where A is a set of all actions and receive a reward r from environment E. We indicate the process of generating r from E as $e(a) \to r$, where e is a function associated with E and takes an action a as parameter.

In the first stage, we assume the user model is stationary, as a consequence, function e is stationary. Now we consider how can $e(a)$ be defined. In a related

work by [1], the reward is defined by valence provided by an affect recognition software. We will use the same setting at the beginning. Later, we consider the non-stationary case and combine MBA with contextual information.

Consider we have a contextual vector $\vec{s}_t$

[1] The Eyes cant See, but the Heart Feels! Affect-Aware Robots that Learn and Adapt to the User Over Time