Decentralized Multi-agent POMDPs for Human-Robots Team Coordination

Hend Al Tair
Robotics Institute
Khalifa University
Abu Dhabi, UAE

Tarek Taha
Robotics Institute
Khalifa University
Abu Dhabi, UAE

Mahmoud Al-Qutayri
College of Engineering
Khalifa University
Abu Dhabi, UAE

Jorge Dias
Robotics Institute
Khalifa University
Abu Dhabi, UAE

Email: hend.altair@kustar.ac.ae Email: tarek.taha@kustar.ac.ae Email: mqutayri@kustar.ac.ae Email: jorge.dias@kustar.ac.ae

Abstract—Despite the fact that robots have reached a high level of autonomy in recent years, the need for human presence in certain situations is still essential especially in search and rescue scenarios. The human extends the robots capabilities beyond what they are capable of with current technologies. While current robotic devices are able to navigate, locate, and map search and rescue areas, some interventions require high degree of dexterity and information exchange that implies cooperation between the agents intervening in the area - human and/or robots. This paper presents a framework for modelling the coordination between human responders and robots in search and rescue scenarios using Decentralised Multi-agent Partially Observable Markov Decision Processes (Dec-MPOMDP). In this framework the human is modelled as an intelligent agent with observations and actions that are communicated with the remain team (human and robots) to reach the level of synergy required to accomplish joined tasks.

I. INTRODUCTION

Utilising robotic systems in hazardous situations continues to expand and hence research effort is being put to enhance the efficiency and performance of these systems. Robots involvement in hazardous situations not only saves lives by reducing the human exposure to dangerous environments, but also increases the efficiency of responding to these incidents. Examples of those incidents might vary from small fires [1], oil spillage incidents [2], to major nuclear incidents [3]. Most current robotic systems tend to either be autonomous or work under the control of a human operator. However, to deal with emergencies, particularly the large scale ones, systems that allows seamless collaboration between teams of robots and humans is highly desirable. Collaboration between robots and humans is challenging since they are heterogeneous [4]. Robust decisions and coordination among agents are required so that they can interact, collaborate and form one team based on shared observations and actions. However, in search and rescue missions it is normal that information can be uncertain which affects the decisions taken. The work proposed in this paper focuses on the deployment of a group of robots and humans as intelligent agents in a search and rescue scenario.

II. MOTIVATION AND RELATED WORK

Direct physical human intervention within search and rescue scenarios is very risky and dangerous, and the operating conditions in these environments are beyond the limits of humans (heat, radiation, narrow openings, visibility, smoke ...). Modern robotic systems are being deployed in such domain. However, it is challenging to inject high degree on intelligence in the robotic system in order to be able to work safely and naturally along side humans. The decision making is critical since it requires sharing relevant information among agents [5]. In uncertain and unpredictable environments of search and rescue scenarios some information can be lost. Planning under uncertainty motivates the usage of Partially observable Markov Decision Processes (POMDPs). There are two main models of extending POMDP: Decentralized POMDP (Dec-POMDP) and centralized POMDP (MPOMDP). The difference between them is that in case of decentralized each agent observes its local observation only o_h while in the centralized each agent observe the full observation (local and global from other agents) which requires high communication o [6]. However, in search and rescue communication is limited. Therefore, a suitable approach is based on reducing the MPOMDP which is proposed in this paper as the Decentralized Multi-agent Partially Observable Markov Decision Processes (Dec-MPOMDP). The focus of the work in this paper is to use Dec-MPOMDP in the planning process of taking actions in an uncertain environment to find and rescue victims by cooperative team of robots and humans. The remainder of this paper is organized as follows. The paper begins with a brief review of some related work using POMDP with multi-agents and more specifically in search and rescue scenarios, followed by description of the Dec-MPOMDP. Finally, the proposed multi-agent HRI collaborative scenario is presented.

III. DEC-MPOMDP

The Dec-MPOMDP is a Dec-POMDP with reduced MPOMDP. In order to define the formal model Dec-MPOMDP, a definition of Dec-POMDP will be presented.

A. Model Parameters

Referring to Oliehoek et al. work [7] standard Dec-POMDP for n agents is defined as a tuple \( \langle n, S, A, T, \mathcal{R}, O, o, h, b^o \rangle \) consists of:

- a finite set of \( n \) agents;
- \( S \) is a finite set of world states;
- \( A = \times_i A_i \) is the set \( \{a_1, ..., a_J\} \) of \( J \) joint actions. \( A_i \) is the set of actions available for agent \( i \). At each time step one \( a = (a_1, ..., a_n) \) is taken.
• $T$ the transition function which defined the probability of going to state $s'$ when in state $s$ under action $a$, $p(s'|s, a)$.

• $R$ is the reward function. A reward is given for taking an action $a$ when in state $s$, $R(s, a)$.

• $O = \times_i O_i$ is the set $\{o^1, ..., o^K\}$ of $K$ joint observations. Every agent $i$ has a set of observations $o_i = \langle o_{i1}, ..., o_{ih} \rangle$ at each time step.

• $O$ is the observation function, represents the probability of an observation $o$ after an action $a$ and ending up in state $s'$ experienced, $p(o|a, s')$.

• $h$ is the horizon, the number of time steps that agents will interact with their environment.

• $b^0 \in \mathcal{P}(S)$, is the initial state distribution at time $t$. $\mathcal{P}(S)$ donates the set probability distributions over $S$.

The planning problem aims at finding a policy $\pi$ for every agent that is optimal for a particular number of steps $h$. Joint policies is a tuple of policies $\pi = (\pi_1, ..., \pi_N)$. Each policy $\pi_i$ is a mapping from histories of observations, that are received by agent $i$, to an action $\pi_i(o^1_i, ..., o^K_i) = a_i$. The joint belief $b$ is a probability distribution over states. It is computed after the action-observation history which consists of all observations received and actions taken up to time step $t: = (o_0^i, o_1^i, a_1^i, ..., o_{t-1}^i, a_t^i)$. The joint belief is calculated:

$$b^{t+1}(s') = \frac{P(o|a, s')}{{\mathcal{P}}(o|b^t, a)} \sum P(s'|s, a)b^t(s) \quad (1)$$

where $P(o|b^t, a)$ is used for normalization.

IV. DYNAMICS OF THE PROPOSED MULTI-AGENT FRAMEWORK

As illustrated in Figure 1 each agent (robots and humans) receives its local observation from the environment and communicate it to the other agents. A central agent receives the joint observations and within State Estimator (SE) unit it combines the joint observations with previous joint actions and joint belief to update the joint belief and compute joint policies. Joint actions then are sent to the agents (robots and humans). Each one of the agents will perform its own task.

V. SCENARIO DESCRIPTION

In response to a search and rescue mission, a team has been dispatched in order to locate and extract a victim present in the rescue area. This team consists of two robots and a human. The two robots are capable of scanning the area, locating the victim, locating any source of danger, and clearing the danger. The human has the capability to locate the victim, and extracting him/her to the evacuation area. The problem is modelled in a two slice Dynamic Bayesian network (DBN) shown in Figure 2. Further details about states, observations, transitions of the model is in [8].

![Fig. 2. POMDP representation using DBN](image)

VI. CONCLUSIONS

This paper presented a human-robot collaboration framework for a search and rescue scenario. Robots and humans are treated as intelligent agents that share observations. Joint actions are taken based on Dec-MPOMDP in which MPOMDP are reduced since communication is assumed not to be perfect within such domain.

REFERENCES


