

Inference Model for Heterogeneous Robot Team Configuration based on Reinforcement Learning

Xueqing Sun, Tao Mao, Jacob C. Jurmain, Laura E. Ray

Thayer School of Engineering, 8000 Cummings Hall, Dartmouth College, Hanover, NH 03755, USA. {xueqing.sun, tao.mao, jacob.c.jurmain, laura.e.ray}@dartmouth.edu. Tel: 603-646-1243, Fax: 603-646-3856

Introduction

In many practical robotics problems, including search and rescue, hunting, foraging, and patrolling, knowledge of the team configuration and capabilities is crucial in coordination of multiple heterogeneous robots. In a challenging environment with costly, sporadic, or absent communication, inferencing based on observed spatio-temporal state transitions is necessary for learning and reasoning. A wide variety of probabilistic models are available for inferencing, and prominent among them are Bayesian networks, Bayesian knowledge base [1, 2, 3] and Hidden Markov Models [4, 5]. However, for each new application or environment, setting up the application specific model, deriving the probabilistic parameters and implementing them in an environment full of uncertainties is both error prone and computation costly.

In this paper, we present a general purpose inference engine that takes sparse observations of state transitions made during multiple robot team execution of a task as input and dynamically inference the team configuration through a rational decision making process using Reinforcement Learning (RL) techniques[6, 7]. We demonstrate the operation of this approach using a multi-robot foraging task. The method takes into account both probabilistic observation and environmental uncertainties, and is thus robust to incomplete information and dynamic changes in team composition.

Inference Model

The inference engine is derived from brain circuit models [8] and is based on compact coding of the team configuration combined with sparse reinforcement learning. Simulation experiments are provided in the use of this inference engine in foraging tasks performed by heterogeneous multi-robot team. The objectives of our approach are that this inference engine should generalize with a high degree of scalability to size and composition of a robot team and low computational cost during the inference process. Additionally, the inference engine should be robust to the incomplete and uncertain information from state observation. Figure 1 shows a schematic of the framework consisting of primary components for Environmental Sensation (ES), Observation Measurement (OM), Inference Process (IP), and Error Estimation (EE). Primary pathways connect between these individual modules through input and output signals.

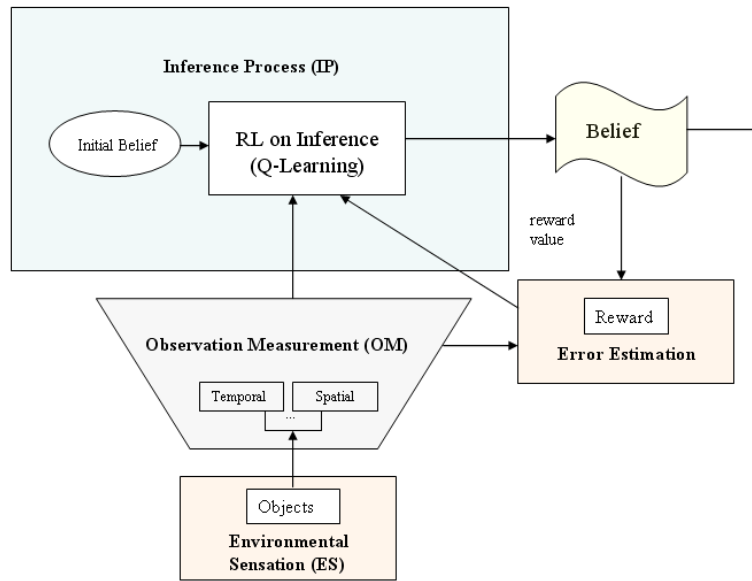


Figure 1: Schematic diagram of an inference framework

ES

In the Environmental Sensation (ES) module, robots sense the outside world. Based on raw sensory output, the sensation signals are packaged and passed to the Observation Measurement module for spatio-temporal representation of the environmental change. The ES module is roughly analogous to the sensory processing areas of the brain, such as the early processing areas of visual cortex of mammals.

OM

Inspired by research on how the human brain receives, processes and represents the surrounding world based on its perceptions [9], the Observation Measurement (OM) module takes the raw object sensation and constructs the observation vectors to represent the temporal or spatial changes caused by the robot behaviors. This module is important for the later inference process because it records the actual effects from the actions of the robot team, for which the configuration is the hidden state. If communication is unavailable or unreliable, the output of the observation measurement module is required for the inference module to infer team composition through reinforcement learning.

IP

After the temporal/spatial representation of the team configurations are being constructed, Inference Process (IP) is conducted in this core module which is closely coupled with the Belief and Reward Sensation module through Reinforcement Learning for inferring the team configuration. A recursive Q-learning algorithm “smears” the initial belief and generates (*action*) a set of near-neighbor hypotheses (*state*) for different team configurations and their expected output. The Error Estimation module then calculates the observation residuals between a series of expected output and actual observations to generate the feedback signal (*reward*). Based on the reward, the Q-learning algorithm updates the state-action value Q at the end of a sequence of distinct episodes and records the best policy or belief. The actual observation can be sparse, noisy or incomplete, and can be taken at irregular intervals; however, simulation experiments show that this inference approach is capable of accommodating uncertainty and sparse observation.

The paper is divided into four sections covering a brief overview of the reinforcement learning techniques and Q-learning in particular; Detailed description of the inference framework and algorithm based on reinforcement learning; Presentation of experimental results for various robot-team configurations in foraging tasks. The paper will conclude with summary remarks and recommendations for future work.

Simulation Results

The purpose of simulation experiments is to show that inferencing with sparse, incomplete and finite observations can be performed using reinforcement learning in finite time with good results. Figure 2 shows an example of inferencing performed to determine a robot team configuration during a foraging task where robots must retrieve 50 entities. Each entity can contain one or multiple objects, and the robot team is comprised of heterogeneous robots of type “P2”, which can retrieve either one or two objects at a time, or “S_P1”, which can sweep two objects into a cluster (a single entity of two objects) or can pick up a single object. There exists an unknown number of robots of each type, and the inference task must determine the number of robots of each type. The inference module converges to the correct team configuration within three observations. Experiments conducted to test scalability and effectiveness in environmental uncertainty will be reported in the full paper.

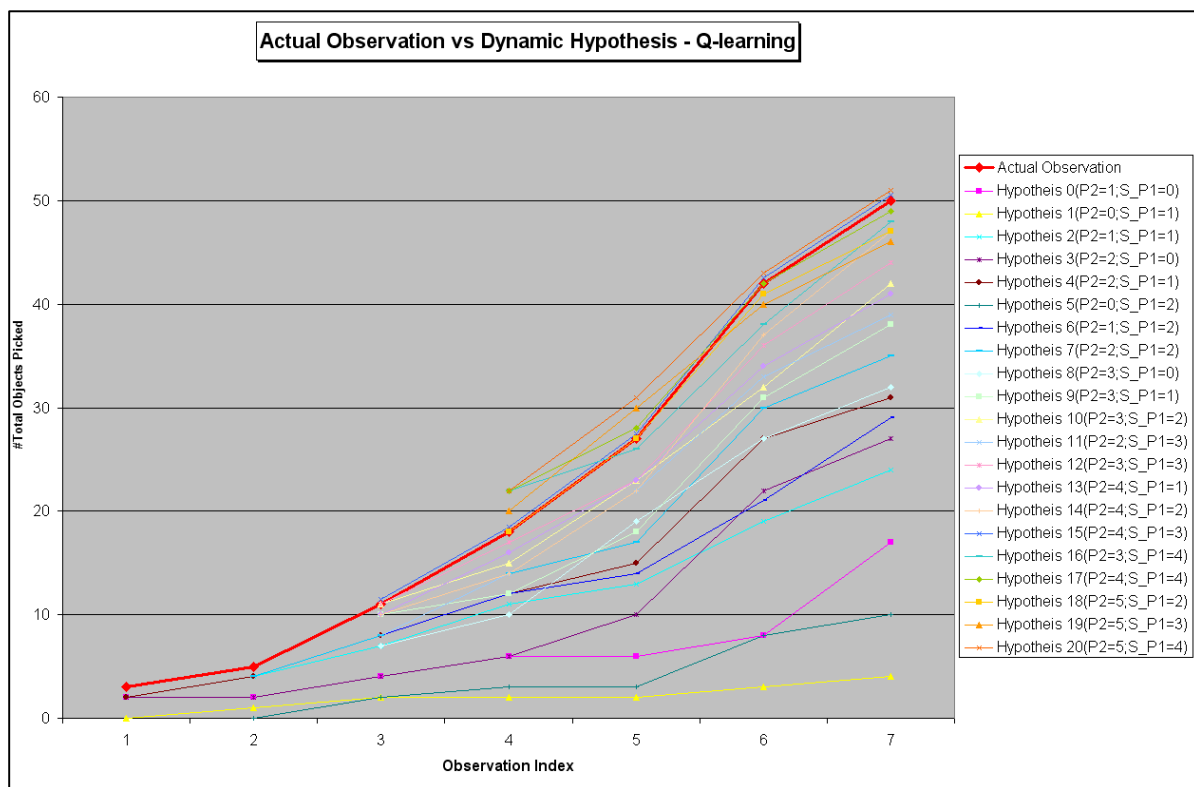


Figure 2: Inference on robot team configuration for a foraging task consisting of 50 entities and unknown number of heterogeneous robots

References

- [1] Milch B., Marthi B., Sontag D., Russell S., Ong D. L., and Kolobov A. (2005). Approximate Inference for Infinite Contingent Bayesian Networks. In Proc. of the 10th International Workshop on Artificial Intelligence and Statistics.
- [2] Charniak E. (1991). Bayesian Networks without Tears. *AI Magazine*, 12(4).
- [3] Santos, E., Jr. and Santos, E.S. (1999). A Framework for Building Knowledge-Bases Under Uncertainty,” *Journal of Experimental and Theoretical Artificial Intelligence* 11, 265-286.
- [4] Rabiner L. R. (1989). A Tutorial on Hidden Markov Models and Selected Applications in Speech. Recognition. *Proc. IEEE*, 77, 257-86
- [5] Fox M., Ghallab M., Infantes G., and Long D. (2006). Robot Introspection through Learned Hidden Markov Models. *Artificial Intelligence*, 170(2). 59–113.
- [6] Sutton R. S. & Barto A. G. (1998). *Reinforcement Learning*. MIT Press, Cambridge, MA.
- [7] Watkins C. J.C.H. & Dayan P. (1992). Technical Note: Q-Learning. *Machine Learning*, Volume 8, Numbers 3-4, May, pp. 279-292.
- [8] Granger R. (2006). Engines of the brain: The computational instruction set of human cognition, *AI Magazine* 27(2), pp. 15-31.
- [9] Hearn R. & Granger R. (2008). Learning Hierarchical Representations and Behaviors. Association for the Advancement of Artificial Intelligence.