Team ChaGamma

Kosuke Shinoda[1 2] and NODA, Itsuki[1 2 3]

[1] Cyber Assist Reseach Center
National Institute of Advanced Industrial Science and Technology
[2] Japan Advanced Institute of Science and Technology
[3] PRESTO, Japan Science and Technology Corporation

1 Introduction

Team ChaGamma is a team of players who have models of environment, self, and other agents as hidden Markov Models (HMM). Using the models, each player can expect changes of the field and make plans of plays for the desired goal states.

In multi-agent systems, each agent needs to have a world model that handles interactions with environment and other agents. More complex the environment and strategies is, harder it is to design such a model that can be used for planning and decision-making. Our approach to overcome this problem is to apply HMM and its learning method for building the suitable world model.

We also investigate coupled HMM (CHMM) to model cooperation with other agents. In order to reduce the number of parameters of CHMM, we try to introduce various assumptions that reflect agreements of joint intentions among agents.

2 Segmentation of Environment through HMM Learning

Agent Model and Learning Method: In general, an autonomous agent is modeled as a Merly-type HMM (shown in Figure 1(a)), in which, the agent's behaviors are formalized as follows: The next state $(s^{(t+1)})$ is determined only by the previous state $(s^{(t)})$, and the agent' action $(a^{(t+1)})$ is determined by the current state transition $(s^{(t)} \rightarrow s^{(t+1)})$. This formalization, however, lacks the effect of interaction between the agent and the environment. So, we introduce an assumption, "the internal state and the environment has a probabilistic relation". In other words, an agent and its environment can be defined as a Moore-Merly-type HMM (MM-HMM) (Figure 1(b)).

Segmentation of Environments: Suppose that a learner can observe a sequence of demonstrator's actions $\{a^{(1)} \dots a^{(T)}\}$ and changes of an environment $\{e^{(0)} \dots e^{(T)}\}$. The purpose of the learner is estimate an HMM that can explain the given action and environment sequences most likely. We can derive a learning algorithm, an extension of Baum-Welch algorithm, for the MM-HMM[7, 8].
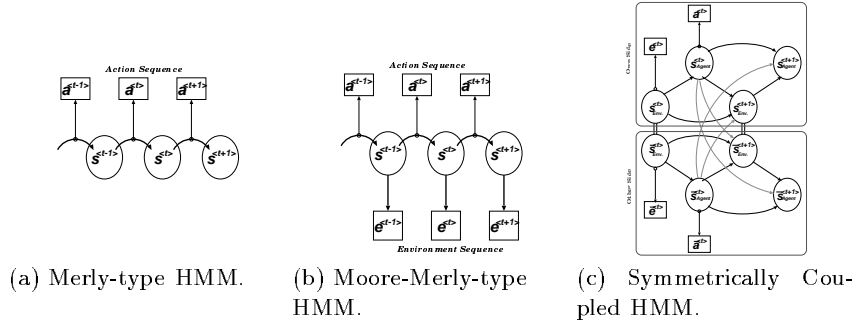
(a) Merly-type HMM. (b) Moore-Merly-type HMM. (c) Symmetrically Coupled HMM.

Fig. 1. Agent HMM.

Using this algorithm, the learner will acquire a suitable HMM, whose state transition reflects both of the internal intentions of the demonstrator and segmentation of the environment. In the HMM, each state corresponds to a combination of an intention and a segmentation for it. In other words, the representation of the intention and the segmentation are mixed in a set of states. While such representation is enough to imitate demonstrator's behavior, it is also useful to know how the acquired HMM segments the environment.

The segmentation of the environment can be represented by a probability function $Pr(s|e)$ where e is an environment data and s is an internal state of HMM. This probability can be calculated by the following equation:

$$Pr(s|e) = \frac{Pr(e|s)Pr(s)}{Pr(e)}$$

Using this equation, we can know how the acquired HMM segments the environment.

Figure 2 shows the separation of the environment the learner acquires by learning two types of dribble plays. These graphs shows correspondence between the environment (distance and direction to the ball) and states (each line in the graph) as changes of probabilities $Pr(s|e)$. In these graphs, we can see that the HMM separate the environment into more fine states in the case elegant dribble than the case of rough dribble. This separation is reasonable because elegant dribble need fine-tuned control of the ball that needs fine separation of the environment.

3 Symmetrically Coupled HMM

Symmetricity Assumption: In order to extend the HMM for the agent, we must tackle a problem of complexity of calculation and memory of the HMM, because the number of states increases exponentially when the agent has more interactions among agents. One of important issues of the problem is *generalization performance*, as the number of states or learning parameters increase, the huge number of examples are required to guarantee the generalization performance. In order to avoid the problem, I introduce *symmetricity assumption* among agents as follows [6]:
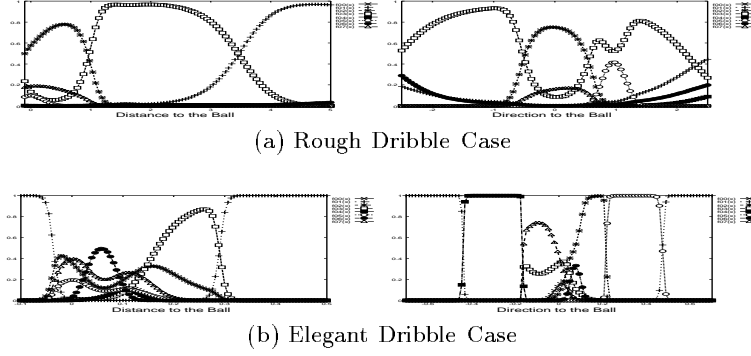
(a) Rough Dribble Case

(b) Elegant Dribble Case

Fig. 2. Separation of Environment Acquired by Learning Rough and Elegant Dribble

symmetricity assumption
Every agent has the same rules of behavior. In other words, every agent shares the same state transition rules with each other.

Based on this assumption, we derives a *symmetrically coupled HMM* (sCHMM) shown in Figure 1-(c).

Discussion: The Number of Parameters in the Model: As mentioned before, the number of parameters in HMM is an important factor for generalization performance of learning. In the case of coupled HMM, especially, the number of parameters increase exponentially. Actually, if we use a simple coupled HMM, the number of parameters in the state transition is $|S_{\mathbf{e}}|^2 |S_{\mathbf{a}}|^N + N |S_{\mathbf{e}}| |S_{\mathbf{a}}|^{N+1}$, where N is the number of agents, and $|S_{\mathbf{e}}|$ and $|S_{\mathbf{a}}|$ are the numbers of states of environment and agent respectively. Compared with this, in the symmetrically coupled HMM, the number of parameters is reduced to $|S_{\mathbf{e}}|^2 |S_{\mathbf{a}}|^N + |S_{\mathbf{e}}| |S_{\mathbf{a}}|^{N+1}$. We can also introduce the following additional assumption based on cooperative properties of multi-agent systems.

No Explicit Communication Assumption: In the case of real soccer, players can share their joint intention by sensing environment and other's behaviors without explicit communication. In such case, the transition of the agent state are determined only by previous states of the environment and the agent itself. As a result, the total number of the parameters is reduced to $|S_{\mathbf{e}}|^2 |S_{\mathbf{a}}|^N + |S_{\mathbf{e}}| |S_{\mathbf{a}}|^2$.

Shared Joint Intention Assumption: During a cooperation of multiple agents each agent believes that all agents completely share the joint intention. This means that each agent believes that other agents will behave as the agent wants. In this case, the transition of environment states are determined only by the previous states of the environment and one agent. This will reduce the number of parameters to $|S_{\mathbf{e}}|^2 |S_{\mathbf{a}}| + |S_{\mathbf{e}}| |S_{\mathbf{a}}|^{N+1}$.

4 Related Works and Open Issues

Brand Et al. [1,3] proposed coupled HMM and its learning method, in which several HMMs are coupled via inter-HMM dependencies. Jordan Et al. [5,2, 4] proposed factorial HMM and hidden Markov decision trees. Both of works mainly focused on reducing the complexity in EM processes. They use mean field approximation or N-heads dynamic programming to reduce the cost of the approximation of posterior probabilities. However, they does not focused on symmetricity in agent-interactions and generalization performance problem.

The proposed methods have the following open issues on the proposed method:

- The cost of calculation increase exponentially when structures of agents and environments become complicated. In order to reduce the complexity, several techniques like mean field approximation and N-head dynamic programming should be applied to these models.
- The incremental learning will suit to acquire high-level cooperative behaviors. We may be able to realize the step-by-step learning using dependency of the initial parameters.

References

1. Matthew Brand. Coupled hideen markov models for modeling interacting processes. Perceptual Computing/Learning and Common Sense Technical Report 405, MIT Lab, jun 1997.
2. Zoubin Ghahramani and Michael I. Jordan. Factorial hidden markov models. *Machine Learning*, 29:245–275, 1997.
3. Coupled hidden Markov models for complex action recognition. Matthew brand and nuria oliver and alex pentland. Perceptual Computing/Learning and Common Sense Technical Report 407, MIT Media Lab, 20 1996.
4. Michael I. Jordan, Zoubin Ghahramani, Tommi Jaakkola, and Lawrence K. Saul. An introduction to variational methods for graphical models. *Machine Learning*, 37(2):183–233, 1999.
5. Michael I. Jordan, Zoubin Ghahramani, and Lawrence K. Saul. Hidden markov decision trees. In Michael C. Mozer, Michael I. Jordan, and Thomas Petsche, editors, *Advances in Neural Information Processing Systems 9*, page 501. The MIT Press, 1997.
6. Itsuki NODA. Hidden markov modeling for multi-agent systems. In *Proc. of The Seventh Pacific Rim International Conference on Artificial Intelligence*, Aug. 2002. (submitted).
7. Itsuki NODA. Segmentation of environments using hidden markov modeling of other agents. In *Proc. of AAMAS-2002*, July 2002. (accepted).
8. Itsuki NODA. Segmentation of environments using hidden markov modeling of other agents. In *Proc. of the 16th Annual Conf. of JSAI*, 5 月 2002. (in Japanese) (accepted).