
Possible computational ways to represent goals

Xingping Yu
Peking University
2100017812@stu.pku.edu.cn

Abstract

Intentionality can sometimes be regarded as goals. In our daily life, goals can be a long-term picture or a short-time motivation. Observational studies show that infants selectively encode the goal object of an actor's reach [4]. Also, infants can understand intentions of some certain actions. When it comes to AI agent, we need to represent the goal and its reward in a computational way before we train it to recognize and reason. So how to represent the goals and there advantages and disadvantages? This essay talks about some possible ways to represent goals and the relative rewards in a computational way. They are: single-goal representation, multi-goal representation, and hierarchical goal representation. The ideas are based on existing works and researches. The essay also discusses the advantages and disadvantages of these ways.

1 Introduction

Intention and action prediction is an important part in recognitive and reasoning model. Before we train the model to learn to understand the intentions, we should define the target sets and represent the goals. Research shows that infants can take the "intentional stance" in interpreting the goal-directed spatial behavior of a rational agent [3]. To make model more like human, we need to define the goals in a computational way. That means the model can search the goals in the target set and find ways to get the largest reward. The ways to represent the goals are expected to be time-saving to search and take less space. In the coming content, I will propose some possible ways to represent the goals and discuss their advantages and disadvantages.

2 Related works

In fact, there are already some ways to represent the goals in existing studies. Most of them studied the representation of goals and the intention and action prediction at the same time. On the one hand, they tried to find an efficient way to represent the goals. On the other hand, they use the ways to predict the agent's intention and future action.

In Csibra and Gergely(2007)'s research [2], the relationship between goal attribution and action understanding was discussed. The research proposed three mechanisms and discussed their applications in infants and adults. They are action-effect association, simulation procedures and teleological reasoning.

The action-effect association is based on the agent's ability to connect the action and the results. It infers the goals of the actions from the results. This is a familiar mechanism to our human kind. It is often believed that the results of the action can represent the intention and goal. It indeed represent the goal if succeeded. The simulation procedures are based on agent's ability to empathize. In another word it need the agent to put itself in others' shoes. Then the agent can generate the possible mental conditions to understand others. Teleological reasoning is the third mechanism. It transfers the agent's decision policy into others to understand and predict the intentions. These mechanisms play different roles in goal attribution and action understanding. The action-effect association can predict the results quickly. The simulation procedures can predict the goals and actions. The teleological reasoning can

help understand the observed actions. For AI agents, the first and third mechanisms are more practical and possible to realize.

However, Csibra's research mainly focused on goal attribution and action prediction. It talked little about ways to represent the goals. In Baker(2009)' research [1], researchers made some experiments to measure the predictions of the agents. The results provide some requires for the representation of the goals.

The experiment 1 shows that in online target inference, the goal representation must be more flexible than the single fixed location goal representation. The changing goals or subgoals based on conditions represent better than single fixed representation. And the former is more closed to human's judgments. The experiment 2 provides clear evidence for the use of representations if changing goals in retrospective judgment tasks. And the experiment 3 shows that given enough evidence, people will use more complex subgoal-based representations. These representations can account for multiple trajectories that start at different locations but are assigned to the same result.

Generally speaking, the goal representation should be complex and changable based on certain conditions. No matter how the goals are represented, the agent must be able to change and fix the goals based on the judgment of the situation.

3 Naive thoughts about possible ways

Based on the researches above, I would like to talk about three possible ways to represent the goals and discuss their advantages and disadvantages. They are single-goal representation, multi-goal representation, and hierarchical goal representation.

First, the single-goal representation defines the goal as a long-term picture or a big-scale picture. It means that the goal is the end. This method is quite simple, which is easy to realize and apply. It nearly doesn't cost anything to search and change the goal. However, the representation may be too abstract to understand and pursue. It has no idea how the goal can be realized step by step. When the goal is potentially large and hard, this representation will not perform well.

The multi-goal representation represents the goals as a set. Each goal has a weight of priority. That means the set may be in order. This representation can help the agent understand the priority of each goal. The agent then decide which action should be taken first based on the decision policy. However, it can not represent the relationships between different goals. The agent may not understand the goal and subgoal relationships. This will have a negative effect on the action and result.

The hierarchical goal representation uses a tree to represent the goals and subgoals. The root is the ending goal. And each branch represents the relationship between goals and subgoals. There are also weights of priority on the branches. The advantage of this representation is obvious. It is very clear and easy to understand. The relationships of goals and subgoals and different goals are represented clearly. The disadvantage is that the tree may be too complex to change and search. It costs much time and space to represent the goals.

4 Conclusion

In this essay, I concluded some results of existing works about goal attribution and action prediction. Inspired by them, I talked about three different ways to represent the goals in computational ways. Finally, I discussed their advantages and disadvantages. In my opinion, the multi-goal representation and hierarchical goal representation are more practical and complex. The representation of goals should consider the complexity, the time and space consuming, and the efficiency of the results. And the representation of goals should change based on the situation.

To sum up, choosing the appropriate goal representation depends on the specific application scenario and requirements. Different representation methods have their advantages and limitations, which need to be weighed and selected according to the specific situation.

References

- [1] Chris L Baker, Rebecca Saxe, and Joshua B Tenenbaum. Action understanding as inverse planning. *Cognition*, 113(3):329–349, 2009. 2
- [2] Gergely Csibra and György Gergely. ‘obsessed with goals’: Functions and mechanisms of teleological interpretation of actions in humans. *Acta psychologica*, 124(1):60–78, 2007. 1
- [3] Gergely Csibra Szilvia Biro Gyorgy Gergely, Zoltan Nadasdy. Taking the intentional stance at 12 months of age. *Cognition*, 56(1):165–193, 1995. 1
- [4] Amanda L Woodward. Infants selectively encode the goal object of an actor’s reach. *Cognition*, 69(1):1–34, 1998. 1