Procedia Computer Science
Volume 88, 2016, Pages 58–63
7th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2016

A Generative Probabilistic Model for Leaning Complex Visual Stimuli

Alexey Potapov[1,2] and Vita Potapova[2]
[1]*Department of Computer Photonics and Videomatics, ITMO University, St. Petersburg, Russia*
[2]*Faculty of Liberal Arts and Sciences, St. Petersburg State University, St. Petersburg, Russia*
pas.aicv@gmail.com, elokkuu@gmail.com

Abstract
The problem of representing and learning complex visual stimuli in the context of modeling the process of conditional reflex formation is considered. The generative probabilistic framework is chosen which has been recently successfully applied to cognitive modeling. A model capable of learning different visual stimuli is developed in the form of a program in Church (probabilistic programming language). NAO robot is programmed to detect visual stimuli, to point at selected stimuli in a sequence of trials, and to receive reinforcement signals for correct choices. Conducted experiments showed that the robot can learn stimuli of different types showing different decision-making behavior in a series of trial that could help arranging psychophysiological experiments.

Keywords: Cognitive robotics, Probabilistic programming, Visual stimuli, Conditional reflex

1 Introduction

Biologically inspired robotics pays attention to different aspects of living things including both bodies and behavior. The latter is studied on very different levels starting from neurons and ending with higher cognitive abilities. However, most works are focused either on low-level or high-level aspects. The latter are usually modeled with cognitive architectures [1–3], which capabilities, however, are highly dependent on low-level control and perception modules. At the same time, these low-level modules usually perform some fixed operations. This results in limiting intermediate level control to some set of standard tasks as picking up an object or navigating to a destination.

This can also be seen on the problem of semantic grounding in sensomotoric data. The basic idea consists in detection of co-occurring semantic categories and linguistic units [4]. These investigations have high theoretical and practical significance. However, they have certain limitations, because of restricted perceptual information representations, which can capture restricted (and predefined) set of regularities in sensorial and linguistic input. The most common and strong restriction consists in consideration of only objects (ball, cup, etc.) and their features (shape, color, etc.) as semantic categories or visual concepts [4, 5]. In more recent works [6], spatial relations are also used to ground

Selection and peer-review under responsibility of the Scientific Programme Committee of BICA 2016
© The Authors. Published by Elsevier B.V.

doi:10.1016/j.procs.2016.07.407

adverbs (in addition to verbs grounded in actions), but this possibility is defined *a priori*. It is interesting to point out that utilized learning techniques do not exceed those used in early classical cybernetic models of conditional reflexes [7, 8]. Here, a semantic category and a linguistic unit are just two stimuli, between which a conditional link is established.

Capabilities of animals to form conditional reflexes are far beyond that of robots. Without these capabilities, robots will be highly limited in their higher cognitive functions. Surprisingly, detailed models of general mechanisms of conditional reflex are rare. At the same time, some authors believe [9] that classical (Pavlov) and operant conditioning are key mechanisms in the learning of any adaptive behavior of animals, and robot training methods are still very far from this.

Models of operant (not classic) conditioning are developed in more details in robotics. However, they mostly address the problem of complex structure of motor acts [7]. Traditionally, robotic conditioned reflex systems are implemented in supposition that stimuli (or events such as ‘hearing a bell’, ‘seeing a light’, etc.) are already extracted and identified (in different trials) [9]. At best, visually observed objects are used as stimuli [10].

At the same time, experiments with animals reveal their ability not only to respond to any circle or red object, but also to learn to select the darkest object, object with certain number of dots, and so on [11, 12]. Interestingly, chickens (just as most cognitive robots) cannot learn to choose stimulus based on its relative brightness. How can we supply robots with abilities similar to that of higher animals? At least, such diverse stimuli should be representable in sensory systems of robots.

In order to investigate this problem, we placed a robot in conditions similar to those used in some experiments on animal’s conditioning. The robot repeatedly observes several “feedboxes” with different marks (we use geometric figures). It should point at a correct one to receive reinforcement. To do this, it should learn a correct rule, which can be based on relative or absolute size, position, shape, or brightness of stimuli.

This task states such questions as how to represent these rules and how to search for them. Previously [13] we used logic rules, which were inferred by genetic programming. This solution is technically admissible, but is neither flexible, no biologically plausible. More appropriate approach is to use probabilistic generative models, which are successfully applied in cognitive modeling [14].

The main contribution of our work consists in application of the framework based on generative probabilistic models to intermediate-level perception (involved in conditioning) in addition to high-level cognition.

2 Methodology

We considered the task of selecting “feedboxes” with rewards on the base of marks (stimuli) placed on them. In each trial, three “feedboxes” were given, and strictly one of them “contained” a reward. Marks on “feedboxes” with rewards followed the same rule within a series of trials.

We programmed a NAO robot to select a stimulus from several visually observed stimuli and to point at the selected stimulus after its head is touched, and then to interpret the next touch as the positive feedback. Processing of images was carried out by our previously developed algorithms [14]. Derived descriptions of each stimulus contain coordinates (x, y) within its bounding box, brightness (b), size (sz), position of the bounding box in the image (pos = left, middle, right), and shape (sh = circle, triangle, square). Thus, on each trial, the robot has a history of previous trials including the correctness of the selected stimuli, and it should determine the most probably rewarded stimulus.

Representing rules for selecting correct stimuli in predicate logic is straightforward [13], however inductive inference of these logic rules is not supported by deductive inference engines of such logic languages as Prolog, and is needed to be separately implemented, e.g. on the base of genetic algorithms or some other metaheuristic search techniques. Probabilistic programming allows for representing rules for selecting stimuli and inferring these rules from.

Download English Version:

https://daneshyari.com/en/article/4962246

Download Persian Version:

https://daneshyari.com/article/4962246

Daneshyari.com