Int. Journal of Math. Analysis, Vol. 6, 2012, no. 30, 1453 - 1455

Robot Self-Awareness: Occam's Razor for Fluents

Anna Gorbenko

Department of Intelligent Systems and Robotics
Ural Federal University
620083 Ekaterinburg, Russia
gorbenko.ann@gmail.com

Vladimir Popov

Department of Intelligent Systems and Robotics
Ural Federal University
620083 Ekaterinburg, Russia
Vladimir.Popov@usu.ru

Abstract

In this paper we consider robot self-awareness from the point of view of temporal relation based data mining. In particular, we use rational function approximations and consider Occam's razor for fluents.

Mathematics Subject Classification: 41A20

Keywords: robot self-awareness, Occam's razor, fluents

A robot with self-aware system has the possibility of dealing with complex novel situations more effectively than a robot without self-awareness (see e.g. [1, 2]). In [1], the authors considered robot self-awareness from the point of view of temporal relation based data mining. Temporal patterns can be expressed using fluents (see e.g. [1, 3]). A fluent is a proposition with temporal extent. For example, "drinking-coffee" can be defined as a fluent that is true whenever I am drinking coffee. This fluent can be represented as a binary time series x, where $x[t]$ is 1 if and only if I am drinking coffee at time t. Respectively, temporal relations needed only to express relations of fluents.

To build a system of self-awareness the robot needs a system of automatic selection of fluents and prediction of values of selected fluents. Using some kind of system of self-learning and accumulation of knowledge (see e.g. [1, 3]) the robot can obtain fluents as experimental measured dependencies. Therefore, we can use the rational function approximations for prediction of values of fluents. Since the function $x[t]$ is given by measured data pair, the rational function approximations can be made by following manner. For every $t[i]$, $1 \leq i \leq n$, we can write following expression $x[t[i]] = \frac{\Sigma_{j=0}^{p} a_j t^j[i]}{1+\Sigma_{l=1}^{r} b_l t^l[i]}$, where $x[t[i]]$ and $t[i]$ are known data and a_j, b_l are required unknown coefficients. This expression we can write in following form

$$x[t[i]] = \Sigma_{j=0}^{p} a_j t^j[i] - x[t[i]]\Sigma_{l=1}^{r} b_l t^l[i].$$

Note that we can suppose that $n = p + r$. Therefore,

$$(a_0, \dots, a_p, b_1, \dots, b_r)^T =$$

$$\left(\begin{pmatrix} 1 & t[1] & \dots & t^p[1] & -x[t[1]]t[1] & \dots & -x[t[1]]t^r[1] \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & t[n] & \dots & t^p[n] & -x[t[n]]t[n] & \dots & -x[t[n]]t^r[n] \end{pmatrix}^T \cdot\right.$$

$$\left.\begin{pmatrix} 1 & t[1] & \dots & t^p[1] & -x[t[1]]t[1] & \dots & -x[t[1]]t^r[1] \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & t[n] & \dots & t^p[n] & -x[t[n]]t[n] & \dots & -x[t[n]]t^r[n] \end{pmatrix}\right)^{-1} \cdot$$

$$\begin{pmatrix} 1 & t[1] & \dots & t^p[1] & -x[t[1]]t[1] & \dots & -x[t[1]]t^r[1] \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & t[n] & \dots & t^p[n] & -x[t[n]]t[n] & \dots & -x[t[n]]t^r[n] \end{pmatrix}^T \cdot$$

$$(x[t[1]], \dots, x[t[n]])^T.$$

The thesis "Entities should not be multiplied beyond necessity" articulated by the 14th century logician William of Occam and known as the Occam's razor states that the simplest theory should be chosen from among all theories which equally well explain observed data [4]. Occam's razor is obviously relevant to machine learning [5]. In particular, it can be used for selection of fluents. However, the exact way it translates to this context has been subject to a long-lasting debate (see e.g. [6, 7, 8]).

Let $length(z)$ denote the number of bits which are required for representation of z. We can consider

$$(p + r + 1) \max\{\max_{0 \leq j \leq p} length(a_j), \max_{1 \leq l \leq r} length(b_l)\}$$

as a measure of complexity of fluent. In our model, robot automatically use a genetic algorithm to determine rational function approximations of fluents with best predictive capability. After this, robot consider a set of simplest fluents. Let $x[t[i]] = \frac{\Sigma_{j=0}^{p} a_j t^j[i]}{1+\Sigma_{l=1}^{r} b_l t^l[i]}$ and $y[t[i]] = \frac{\Sigma_{j=0}^{p} c_j t^j[i]}{1+\Sigma_{l=1}^{r} d_l t^l[i]}$. Robot use the function

$$\Sigma_{j=0}^{p}(a_j - c_j)^2 + \Sigma_{l=1}^{r}(b_l - d_l)^2$$

as the measure of similarity of fluent $x[t[i]]$ and $y[t[i]]$. After this, robot use the following assertion for further selection of fluents: "Models selected from a larger set of tested candidate models overfit more than those selected from a smaller set" (see e.g. [8]).

References

[1] A. Gorbenko, V. Popov, and A. Sheka, Robot Self-Awareness: Temporal Relation Based Data Mining, *Engineering Letters*, 19 (2011), 169-178.

[2] A. Gorbenko, V. Popov, and A. Sheka, Robot Self-Awareness: Exploration of Internal States, *Applied Mathematical Sciences*, 6 (2012), 675-688.

[3] T. Oates and P. R. Cohen, Learning planning operators with conditional and probabilistic effects, *Proceedings of the AAAI Spring Symposium on Planning with Incomplete Information for Robot Problems*, (1996), 86-94.

[4] D. Nolan, Quantitative parsimony, *The British Journal for the Philosophy of Science*, 48 (1997), 329-343.

[5] D. Gamberger and N. Lavrač, Conditions for Occam's razor applicability and noise elimination, *Proceedings of the 9th European conference on machine learning*, (1997), 108-123.

[6] P. Domingos, The role of Occam's razor in knowledge discovery, *Data Mining and Knowledge Discovery*, 3 (1999), 409-425.

[7] S. Esmeir and S. Markovitch, Occam's razor just got sharper, *Proceedings of the 20th international joint conference on artificial intelligence*, (2007), 768-773.

[8] J. Zahálka and F. Železný, An experimental test of Occam's razor in classification, *Machine Learning*, 82 (1996), 475-481.

Received: January, 2012