Robot Self-Awareness: 
Formulation of Hypotheses 
Based on the Discovered Regularities 

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 propose an approach to formulation of hypotheses based on the discovered regularities. 

Keywords: robot self-awareness, the farthest string problem, robot 

Robot self-awareness is extensively studied in robotics (see e.g. [1] – [3]). We can use different regularities for formulation of hypotheses (see e.g. [4] – [9]). The ability to formulate hypotheses is critical to the creation of a system of self-awareness. To formulate hypotheses based on the discovered regularities the notion of contrast used (see [10]). In [10] presented two ways in which the fluency of a hand-over robot-human interaction can be improved. First, in [10] supposed that humans will be more responsive to the robot if they can easily interpret its intentions. In [10] proposed to achieve this by making the robot’s hand-over poses distinct from poses that the robot might have during a different action with the object. Such actions in [10] considered as spatial contrasts. Second, in [10] supposed that the coordination of the hand-over can be improved by making the timing of the hand-over predictable for the human
using an intuitive signal. In [10] proposed using the robot’s movements to
signal the moment of hand-over by transitioning from a pose that is perceived
as non-handing to a pose that is perceived as handing. Such transitioning in
[10] considered as temporal contrast.

Note that in [10] the concept of contrast is used to evoke anticipation in
humans. However, these concept is of considerable interest also for robots. In
order to the concepts of spatial contrast and temporal contrast could use the
robot we need an efficient method for automatically localization and extraction
of instances of spatial contrast and temporal contrast.

For localization and extraction of instances of spatial contrast and temporal
contrast the model of farthest string used. Note that in case of the model of
farthest string consider sequences already proper aligned. Therefore, Hamming
distance is sufficient. To use the model of farthest string it is needed to solve
the following problem.

**THE FARNEST STRING PROBLEM (FS):**

**INSTANCE:** Given a set \( S \) of strings of length \( n \) over an alphabet \( \Sigma \), a
positive integer \( D \).

**TASK:** Find a string \( X \) of length \( n \) over \( \Sigma \) such that \( \delta(X, S) \geq D \) for any
\( S \) in \( S \).

The FS problem is \textbf{NP}-hard for strings over any alphabet \( \Sigma \) with \( |\Sigma| \geq 2 \) [11]. However, there is a PTAS for FS based on a linear programming
relaxation technique. Using satellite models and FS models the robot can
learn sequences of effects of its actions and changes of the environment and
find spatial and temporal contrasts. These abilities the robot can use for
formulation of hypotheses. Efficiency of satellite model depends critically from
the length of data sequence (see e.g. Table 1).

**Table 1: The dependence from the length of data sequence.**

<table>
<thead>
<tr>
<th>Model</th>
<th>( 10^3 )</th>
<th>( 10^4 )</th>
<th>( 10^5 )</th>
<th>( 10^6 )</th>
<th>( 10^7 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prefix</td>
<td>23 %</td>
<td>31 %</td>
<td>43 %</td>
<td>56 %</td>
<td>71 %</td>
</tr>
<tr>
<td>Consensus</td>
<td>38 %</td>
<td>59 %</td>
<td>78 %</td>
<td>89 %</td>
<td>95 %</td>
</tr>
</tbody>
</table>

On one hand, the hypothesis should reflect a situation which is relatively
close to reality. On the other hand, hypothesis should contain an element of
novelty. These two conditions can be interpreted using the following model.
Let \( S = \{S_1, S_2, \ldots, S_k\} \) be the set of input words, \( S_i \in \Sigma^* \), \( 1 \leq i \leq k \). Let
\( |S| \) denote the total length of all words in \( S \). Let \( \#occ(U,V) \) be the number of
occurrences (as a factor) of the word \( U \) in the word \( V \). Consider the following
problem.

**MULTIPLE OCCURRENCES SHORTEST COMMON SUPERSTRING PROBLEM**

(MOSCS):
 Instance: A fixed alphabet $\Sigma$, a positive integer $k$, a set of input words $S = \{S_1, S_2, \ldots, S_k\}$, nonnegative integers $m_1, m_2, \ldots, m_k, n_1, n_2, \ldots, n_k$ and positive integer $m$.

 Task: Find a shortest word $U$ such that $m_i \leq \#occ(S_i, U) \leq n_i$ for all $1 \leq i \leq k$ and $\sum_{i=1}^{k} \#occ(S_i, U) \geq m$.

References


Received: September, 2012