A Multidimensional Scaling Approach to Indexing by Metric Adaptation and Representation Upgrade*

Rodrigo Ventura and **Carlos Pinto-Ferreira**
Institute for Systems and Robotics
Instituto Superior Técnico, TULisbon
Av. Rovisco Pais, 1
1049-001 Lisbon, PORTUGAL
{yoda,cpf}@isr.ist.utl.pt

Following recent neurophysiological research, one important role of emotions consists in providing a mechanism for adequate and efficient response to relevant stimuli (Damasio 1994). In this paper we propose a methodology for implementing such a mechanism, based on a previously presented emotion-based agent model (Ventura & Pinto-Ferreira 1998). This model is founded on a double knowledge representation paradigm: a stimulus reaching the agent is processed under two different and simultaneous perspectives — a simple (termed perceptual) and a complex (termed cognitive) — from which two differing representation schemata are derived. The perceptual representation is oriented to capturing the relevant aspects of the environment, aiming at a quick response to urgent situations, while the cognitive one is directed towards high-level cognitive processing. These two representations are associated and stored in the agent memory in such a way that, in the future, if the agent is confronted with a similar situation, the perceptual representation will help the retrieval of the cognitive one in an efficient way. This indexing mechanism provides a quick algorithm to find cognitive matches.

The indexing mechanism addressed in this paper was previously formulated and theoretically analyzed, under the assumption that the matching of the cognitive and perceptual images are performed in metric spaces (Ventura & Pinto-Ferreira 2002). Given a stimulus $s \in \mathcal{S}$, the agent extracts two kinds of representations: a perceptual image $i_p \in \mathcal{I}_p$, and a cognitive one $i_c \in \mathcal{I}_c$. Each one of the sets $\mathcal{I}_p$ and $\mathcal{I}_c$ of all possible perceptual and cognitive images is equipped with a metric function, denoted by $d_p : \mathcal{I}_p \times \mathcal{I}_p \to \mathbb{R}_0^+$ and $d_c : \mathcal{I}_c \times \mathcal{I}_c \to \mathbb{R}_0^+$ respectively, mapping pairs of images to distances, interpreted as degrees of mismatch. The memory is assumed to be formed by pairs of cognitive and perceptual images $\langle i_c^k, i_p^k \rangle$ $(k = 1, \ldots)$. The goal of the indexing mechanism is then to find the memory pair which cognitive image minimizes its distance to the one extracted from the stimulus, employing the perceptual representation to do so in an efficient manner.

The research presented here concerns the following problem: how to construct a perceptual representation (and metric) with the goal of optimizing the indexing efficiency. In other words, the ideal perceptual representation and metric are the ones that yield small perceptual distances iff the corresponding cognitive distances are also small. To do so, two strategies are explored. One corresponds to adapting a perceptual metric, via a set of parameters, such that cognitive proximity implies perceptual nearness:

$$d_c(i_c^1, i_c^2) < d_c(i_c^1, i_c^3) \Rightarrow d_p(i_p^1, i_p^2) < d_p(i_p^1, i_p^3) \qquad (1)$$

for all possible image pairs $\langle i_c^k, i_p^k \rangle$ $(k = 1, 2, 3)$ extractable from stimuli. The second strategy addresses the improvement of the perceptual representation, in the following sense. Assuming that the perceptual representation is a vector of features extracted from stimuli, when these features are not sufficiently representative to satisfy (1), the goal is to upgrade the perceptual representation with new, more representative, features. Both of these strategies are approached here using Multidimensional Scaling (MDS) techniques (Cox & Cox 1994).

In this framework, a good perceptual representation is one which satisfies the implication (1) for all image pairs. Note that this goal is similar to the MDS one, once one considers the cognitive distances to be the *dissimilarities,* and the perceptual ones, to be the *distances* among objects, in the MDS terminology. However, there are differences. In the case of the MDS, the metric is given while the object coordinates are sought. In the case of the indexing, the object coordinates (perceptual images) are given, while the (perceptual) metric is subject to adaptation.

We propose to perform a gradient descent, within the framework of MDS, w.r.t. a parametrization of the perceptual metric, instead of w.r.t. the point coordinates. Regarding the construction of additional perceptual features, we propose to append each perceptual image with a prespecified amount of additional components. These components represent the values that the new features ougth to take, for each one of the perceptual images in the training set. Their values are randomly initialized, and subject to gradient descent as in the nonmetric MDS. Concerning how to obtain those added components for new stimuli, the idea we advance is to utilize the obtained values to construct a regression model. That regression model can then be used to obtain the new features values for new stimuli.

Let a perceptual image i_p^r, consisting of the concatenation

*This work was partially supported by FCT (ISR/IST plurianual funding) through the POS_Conhecimento Program that includes FEDER funds.

of q numerical features $x_{r1}, \ldots, x_{rq}$, extracted from a given stimulus, with p additional components $y_{r1}, \ldots, y_{rp}$, be denoted by the vector $i_p^r = (x_{r1}, \ldots, x_{rq}, y_{r1}, \ldots, y_{rp})^T$. These additional components correspond to the values that the new features ought to take for that particular perceptual image. The perceptual metric employed here is parametrized by q coefficients $\theta_1, \ldots, \theta_q$, taking the form

$$d_{rs} = \sqrt{\sum_{i=1}^{q} \theta_i^2 (x_{ri} - x_{si})^2 + \sum_{i=1}^{p} (y_{ri} - y_{si})^2} \qquad (2)$$

This parametrization corresponds to assigning a weight (relevance) to each perceptual feature before calculating the Euclidean metric. Moreover, when the algorithm assigns a zero weight to a feature, that feature can be deleted from the perceptual representation, since it is irrelevant (w.r.t. the cognitive matching). The additional components are not weighted since it would just add redundant degrees of freedom.

The cost function employed here is the sum of the MDS stress, using the cognitive distances as dissimilarities and d_{rs} as distances, with a regularization term penalizing the absolute values of the metric parameters

$$J = S + \xi \sum_{i=1}^{q} |\theta_i| \qquad (3)$$

This latter term, weighted by ξ, is included in the cost function to drive to zero any weight corresponding to an irrelevant perceptual component.

Based on the standard nonmetric MDS algorithm (Cox & Cox 1994), we propose the following one:

1. Start with an initial variables vector $\Lambda = [\theta_1 \cdots \theta_q | y_{11} \cdots y_{np}]^T$ (e.g., $\theta_i = 1$, and y_{ri} randomly distributed).
2. Normalize the metric parameter vector $(\theta_1, \ldots, \theta_q)^T$ to unit norm, since the stress is invariant to scaling of this vector. The additional components $\{y_{ri}\}$ are, however, not normalized;
3. Compute the distance set $\{d_{rs}\}$ using the parametrized perceptual metric (2);
4. Perform the isotonic regression (as in nonmetric MDS) to obtain the set of distances $\{\hat{d}_{rs}\}$;
5. Compute the cost; if its value is below a threshold ϵ, stop the algorithm;
6. Find the gradient of the cost function w.r.t. the variables vector Λ, and perform a step of the gradient descent method;
7. Go to step 2.

In order to evaluate the results, a measure of performance called *eval-order* was introduced, aiming at assessing how well the indexing mechanism would behave. This assessment is performed using a test set disjoint from the training set employed in the gradient descent (cross-validation). Inspired on the *N-best* indexing algorithm described in (Ventura & Pinto-Ferreira 2002), the eval-order is defined in the following way: given a cognitive and perceptual images pair $\langle i_c, i_p \rangle$, determine all perceptual distances from it to images in the perceptual memory (*i.e.,* the training set); then, after sorting all these images w.r.t. the perceptual distances, determine which n-th image pair $\langle i_c^k, i_p^k \rangle$ on the resulting ordered list has the minimum cognitive distance to $\langle i_c, i_p \rangle$. In the ideal case, it corresponds to the first one, and thus an eval-order of 1. Higher values correspond to worse performance.

To validate the proposed methodology, a simple testbed was devised. Random points $\mathbf{x} \in \mathbb{R}^c$ (simulating stimuli) were uniformly drawn from an hypercube of unit side length. The cognitive images $i_c \in \mathbb{R}^c$ were set to the components of $\mathbf{x}$ multiplied by some fixed coefficients $\mathbf{W} = \mathrm{diag}(w_1, \ldots, w_c)$ between 0 and 2 each: $i_c = \mathbf{W}\mathbf{x}$. These coefficients introduce different degrees of relevance to the components of i_c. The perceptual images were obtained by concatenating two vectors: the p first components of $\mathbf{x}$ weighted by a second set of fixed coefficients $\mathbf{V} = \mathrm{diag}(v_1, \ldots, v_p)$ ($p \leq c$); and a vector $\mathbf{u}$ of n random numbers (noise) between 0 and 1. Thus, the perceptual images have $p + n$ components: $i_p = [(\mathbf{V}\mathbf{x})^T | \mathbf{u}^T]^T$.

In the first phase of experimentation, no additional perceptual components were considered, and the cognitive and perceptual dimensions were made equal ($c = p, n > 0$). The results have shown that the algorithm was able to both (1) successfully determine the relevance of each components, in the form of their weights $\mathbf{W}$ regarding the cognitive distance, and (2) to discard the noise perceptual components, by setting the corresponding weights to zero, thus showing a successful capability of identifying irrelevant features. Concerning the eval-order performance measure, the results show a significant improvement of the eval-order performance after using the metric weights found by the algorithm, when compared to $\theta_i = 1$.

The second phase of the experimentation comprised the introduction of new components to the perceptual representation. To do so, the dimension of the cognitive images was made higher than the perceptual one, *i.e.,* $c > p$. Thus, the perceptual metric is performed with less components than the cognitive one. The first impact of this is that, without the introduction of new components, the final cost values were much higher than before, due to lack of fit. This happens because the perceptual representation has not enought information to be able to faithfuly replicate the cognitive distances. The experimental results show a significant reduction of the final cost values, whenever the total dimensionality of the perceptual representation, including the amount of new components, is greater or equal than c, thus showing an improved fit.

References

Cox, T. F., and Cox, M. A. A. 1994. *Multidimensional Scaling*. London, UK: Chapman & Hall.

Damasio, A. R. 1994. *Descartes' Error: Emotion, Reason and the Human Brain*. Picador.

Ventura, R., and Pinto-Ferreira, C. 1998. Emotion-based agents. In *Proceedings AAAI-98*, 1204. AAAI.

Ventura, R., and Pinto-Ferreira, C. 2002. A formal indexing mechanism for an emotion-based agent. In *Proceedings IASTED-2002*, 34–40. Malaga, Spain: ACTA Press.