An Entropy-Based Approach to the Hierarchical Acquisition of Perception-Action Capabilities

David Windridge, Mikhail Shevchenko, Josef Kittler

Centre for Vision, Speech and Signal Processing,
School of Electronics and Physical Sciences,
University of Surrey, Guildford, UK
Tel: +44 1483 876043, d.windridge@surrey.ac.uk

Abstract. We detail an approach to the autonomous acquisition of hierarchical perception-action competences in which capabilities are bootstrapped using an information-based saliency measure. Our principle aim is hence to accelerate learning in embodied autonomous agents by aggregating novel motor capabilities and their corresponding perceptual representations using a subsumption-based strategy. The method seeks to allocate affordance parameterizations according to the current (possibly autonomously-determined) learning goal in a manner that eliminates redundant percept-motor context, thereby obtaining maximal parametric efficiency. Experimental results within a simulated environment indicate that doing so reduces the complexity of a multistage perception-action learning problem by several orders of magnitude.

Key Words: Perception-Action Architecture, Saliency, Subsumption Hierarchy, Affordance

1 Introduction

It is by now generally agreed (e.g. [8,19,1]) that traditional top-down symbolic approaches to autonomous cognition exhibit significant, and previously unforeseen, complexities. Within such approaches, the underlying mechanism of cognition is assumed to be the manipulation of symbolic representations of the environment via a computational system that utilizes pre-existing syntactic protocols, such as predicate logic [5].

A critical issue for top-down strategies is consequently the implied disparity between the twin processes of symbolic representation and symbolic manipulation. One significant manifestation of this disparity is the problem of symbol grounding identified by Harnad [9,18] in relation to autonomous cognitive agents (i.e. those expected to exhibit some level of self-determination with regard to their learning processes). Here, because relations between symbols are syntactically constrained via a purely internal mechanism, any connection with the outside world is of a potentially arbitrary and under-determined nature. Mechanisms for addressing this issue have hence typically attempted to constrain
permissible representation via \textit{a priori} sensorimotor linkages imposed at the design level [7, 11, 6, 22].

However, such approaches potentially limit the extent of possible representation, and run contrary to the ideal of cognitive autonomy. They also run the risk of failing to address changes in the environment. Our ideal, in creating a cognitive architecture, ought thus be to create an agent capable of spontaneously generating abstract symbols simultaneously both \textit{representative of}, and \textit{appropriate to}, its surroundings [12], overcoming any potential epistemic circularity implicit in this notion [24].

One such approach is to consider the problem within the terms of a perception-action framework. In Granlund’s formulation [8] such an artificial cognition architecture can be seen as constituting the principle that ‘actions \textit{precede} perceptions’ within the domain of exploration. Autonomous agents exploiting this idea typically attempt to characterize the relationship between their actions and the environmental changes brought about by those actions. In this way initially random exploration of the agent’s motor domain can lead to a perceptual model of the environment specified in terms of the \textit{affordances} that it offers the agent [16, 17, 13]. Critical to the success of this approach is the overcoming of the classical artificial cognitive notion that environmental representation needs to be fixed \textit{prior} to the specification of action schemata. A great deal of representational redundancy can thereby be avoided.

In order to generate a cognitive architecture with comparable abstract processing abilities to classical top-down architectures, it is possible to graft a symbol processing system onto the percept-action learner in a manner consistent with this ‘representationless’ action principle. However, in an ideal system, this symbol processing ability would arise naturally in the context of the perception-action framework (Various approaches that that can be considered consistent with this idea are set out in [14, 4, 20, 21]).

Our approach to achieving this ideal solution is via the notion of progressive hierarchical perception-action \textit{reparameterization}. In a logical, task-based context such as learning the motor-manipulations required to solve a puzzle, the acquisition of high-level action capabilities (such as the ability to move a puzzle piece into the solution-state) implicitly characterizes environmental affordance in a symbolic fashion. Thus, if the perceptual representation of the environmental affordance possibilities at the apex of a spontaneously-generated subsumption hierarchy [3] can be efficiently reparameterized so as to remove perceptual context irrelevant to that action capability, then a symbolic representation of the environment is implicitly generated.

The learning strategy outlined in section 2 of this paper thus seeks to acquire behavioral capabilities initially via the unsupervised identification of low-level \textit{goals} within the agent’s \textit{a priori} percept-space, which are then correlated with the action domain via randomized hill-climbing searches in the learning agent’s motor space. For the current investigation, these goals are identified via their \textit{information-theoretic} saliency, in contrast to previous purely stochastic approaches [23], such that a large amount of extraneous low-level context is
eliminated at the outset (with the result that the final cognitive architecture so constructed will constitute an information-based perception-action hierarchy).

Once acquired, such low-level abilities can then be concatenated and efficiently reparameterized in order to eliminate perceptual invariance in a task-dependent manner, and thereby generate novel perception-action capabilities. Thus, the proposed mechanism can learn via observation of a supervisory agent, parameterizing inferred motor capabilities in terms of salient goals identified, such that replication of both the supervisory agent's perceptions and actions becomes possible.

In this way, the perception-action hierarchy encompasses increasingly symbolic manipulation by virtue of the autonomous sub-goal specification implied in the progressive reparameterization of the subsumption hierarchy so formed. The implicit representation of the environment is thus of hierarchical set of affordances.

Section 3 will thus constitute an experimental examination of the reparameterization methodology within the environment of a shape-sorter puzzle. Results will demonstrate that the information-theoretic approach to perception-action symbolic hierarchy generation is significantly more efficient than non-hierarchical approaches, as well as being resilient to the presence of distractors. Section 4 will conclude by summarizing experimental and theoretical findings.

2 Methodology

2.1 Acquisition of Primitive Percept-Motor Capabilities via Information-theoretic Saliency

We shall initially define the architecture in generic terms, assuming the existence of an embodied agent capable of undertaking motor actions within the environment. It is further assumed that these actions may be reversed and repeated (so as to permit exploratory and learning behaviors).

A priori motor capabilities in such an agent are defined via the set of independently controllable physical motors \( \{ C_0, \ldots, C_n \} \). However, a primitive behavioral competence is configured in terms of a vector \( p \) that ranges over the agent's a priori perceptual space: ie, \( C_n = f_n(p) \). Thus, the parameters in terms of which the behavioral competences are defined are not explicitly motor parameters, but rather perceptual parameters. The corresponding a priori perceptual space is generic in nature, but will usually minimally consist of a topological label space in order to permit labeling of entities that are invariant under translation-like actions (thus enabling basic object-perception). \( p \) is hence typically a four-dimensional vector, encompassing the three ordinal directions and a label indexing parameter.

In order to generate a primitive behavioral competence it necessary to map the a priori motor capabilities onto the perceptual domain. This is achieved within an initially unsupervised learning context via the identification of salient features within the histogram of perceived features (ie the individual components of \( p \)) that are generated by randomized exploratory actions. Behavioral
competences can then be indexed via these perceptual goals. Associated with these perceptual goals are a particular set of perceptual parameters that are determined by their synchronous behavior; that is, the individual feature peaks are determined to have achieve similar levels of saliency by virtue of exceeding a particular threshold (50% of the maximum). The extraneous components of \( p \) are thus eliminated for each goal, and hence much low-level perceptual redundancy can be removed in accordance with the 'action precede perception' principle. Feature difference histograms are also calculated to capture behavioral competences that are of an explicitly relative character (for instance, the act of aligning one object with another), thus effectively doubling the \( a \ priori \) percept-space dimensionality.

![Graph](image)

**Fig. 1.** Single Feature histogram \( f(x) \) for 2 Object Environment

Perceptual saliency is determined along the scale-based lines specified in [10], rendering the hierarchical perception-action reparameterization implementation an implicitly information-theoretic one. The scale-saliency approach enables identification of salient features in the perceptual space in a manner resilient to shift/scale changes and to noise; it naturally favors isotropy and geometrical unpredictability, and is consequently suited to extraction of parametric percept elements at the most appropriate scale. We might thus plausibly expect it to favor the characteristics typically attributable to sub-goal and solution states within puzzle environments, such as the identification of key movable objects in a scene, and the identification of the conditions under which these integrate with other geometrical-matching entities. The procedure for obtaining perceptual goals, on obtaining a feature histogram \( f(x) \) (such as that given in figure 1 for exploration of a 2D scene with 2 object attractors) is thus to:
(a) Calculate the Shannon entropy $H_D(s, x)$ of local attributes of all points, $x$ of the feature space in question over a range of scales $s$;

$$H_D(s, x) = - \int p(i, s, x) \log p(i, s, x) \, di$$

(1)

where $i$ is a particular feature-value within the radius $s$ centered on $x$.

(cf figure 2)

![Shannon Entropy](image)

**Fig. 2.** Shannon Entropy $H_D(s, x)$ at Differing Feature-Space Scales

(b) Select scales at which the entropy over the scale function exhibits a peak, $s_p$, ie where $H_D(s, x) > \text{threshold}$ or where $d[H_D(s, x)]/ds = 0$

(c) Calculate the magnitude change of the PDF as a function of scale at each peak, $W_D(s, x)$:

$$W_D(s, x) = \left( s^2/\left\lfloor 2s - 1 \right\rfloor \right) \times \int \left( d[p(i, s, x)]/ds \right) \, di$$

(2)

(cf figure 3)

This is essentially a measure of scale self-dissimilarity.

(d) The final saliency, $S$, is then the product $H_D(s, x) \cdot W_D(s, x)$ at each peak. ($S$ is given for the whole scale-feature space in figure 4 to indicate the isolation of peak components)

Suppose, then, that we have obtained a set of perceptual goals that we wish to map onto the motor domain. We select a particular goal $g$ defined by the context-free feature vector: $f = (f_1^g, \ldots, f_n^g)$. The distance between the perceptual goal
Fig. 3. Magnitude Change in PDF Over Entire Feature Space

Fig. 4. Saliency $s$ For Differing Sales ($a$)
and the current state \( \tilde{f} \) is the Euclidean distance; 
\[
D(f, \tilde{f}) = \left[ \sum_{i=1}^{n} (f_i - \tilde{f}_i)^2 \right]^{\frac{1}{2}}.
\]
Stochastic gradient descent via the method of [15] then enables minimization of this quantity over a number of random instantiations and permutations of the motor parameter-space, thereby providing a (partially) context-free mapping between perceptual goals and the agent’s action space. (Note that the method [15] finds a global minimum for all components of \( p \) simultaneously, rather than considering the motor parameters independently).

Crucially, this method for indexing action capabilities via the perceptual domain can be iteratively generalized to enable construction of the parametric perception-action hierarchy. To do this we consider arbitrary concatenations of behavioral competences with an appropriate pruning strategy for non-contributory chains. Thus, if after carrying out the behavioral sequence;
\[
C(p) = C_{r_1}(p_{r_1}), C_{r_2}(p_{r_2}), \ldots, C_{r_m}(p_{r_m})
\]
the goal distance \( D_i = D(\tilde{f}^{(i)}, f^{(i)}) \) does not itself exhibit change for a random parametric instantiation \( i \), then a new sequence is generated. If, on the contrary \( D_i \neq D_{i+1} \), then the sequence is deemed relevant, and the gradient descent procedure continued. If \( D_i = 0 \) at any stage we thus obtain a new parametric feature \( C_{r_{K_0 + 1}}(p_{r_{K_0 + 1}}) \) where \( p_{r_{K_0 + 1}} \) has a perceptual feature domain given by the tensor product:
\[
p_{r_{K_0 + 1}} = p_{r_1} \otimes p_{r_2} \otimes \cdots \otimes p_{r_m}
\]
\[
= (f_{r_1}^{g_1} \otimes \cdots \otimes f_{r_1}^{g_{m_1}}) \otimes \\
(f_{r_2}^{g_1} \otimes \cdots \otimes f_{r_2}^{g_{m_2}}) \otimes \\
\cdots \\
(f_{r_m}^{g_1} \otimes \cdots \otimes f_{r_m}^{g_{m_m}})
\]

Such novel behavioral competences can then in principle be added to existing body of behavioral competences in an iterative manner. The subsumptive nature of the competences so formed implies a hierarchical arrangement of behavioral competences. We will later demonstrate that when such a hierarchy is formed within a supervised puzzle-based environment, this hierarchy naturally reflects the percept-motor sub-goals implicit within the scenario.

Furthermore, in typical operational scenarios, the concatenation \( C(p) \) has the potential to be efficiently reparameterized (reflecting the fact that goal sub-tasks are not typically independent of each other). It is this property that is key enabling the hierarchical redefinition of the perceptual domain in a manner consistent with the notion of defining the environment in terms of the affordances it offers the active agent.

**Parametric Generalization of Behavioral Competences** At the action-level, parametric generalization seeks to make behavioral competences both invariant to environmental configuration and maximally efficient (in the sense of involving no extraneous actions). At the percept level, on the other hand, parametric generalization aims to reorganize the perceptual space associated with
behavioral competences in order to define the minimal number of perceptual parameters associated with it. It achieves this by eliminating constant or derived parameter values, and by reindexing multi-dimensional goal parameters into single vectors (if sufficiently few in number).

The former process may be illustrated via an example deriving from the experimental scenario of Section 3: a robotic manipulator-arm equipped with a gripper in a 2D shape-sorter puzzle environment. Suppose, therefore, that there currently exists just three behavioral capabilities; moving the manipulator-arm from any given initial position to \((x, y)\); aligning the manipulator-arm with an object indexed by the parameter \(n\); and the act of closing the gripper on an object within its grasp. These are respectively designated \(M(x, y), A(n)\), and \(G(s)\) (the latter parameter \(s = \{\text{\textquotesingle}grasp\text{\textquotesingle}, \text{\textquotesingle}ungrasp\text{\textquotesingle}\}\) encompasses the binary states of the gripper).

A typical (potentially autonomously-derived) sub-task in learning the overall puzzle competence is the acquisition of the ability to move an entity with index \(n\) to a 2D location \((x, y)\). Suppose that a redundancy-free sequence of previous capabilities has already been established that is capable of achieving this:

\[ A(n_1), G(s_1), M(x_1, y_1), G(s_2). \]

By the tensor product formulation of Equation 4, this implies a perceptual feature domain ranged over by the vector: \((n_1, s_1, x_1, y_1, s_2)\). However, it is apparent that only three of these features, \((n, x, y)\), function as parametric variables within the act of moving an entity \(n\) to the location \((x, y)\); the variables \(s_1\) and \(s_2\) are always set to constant values; \(s_1 = \text{\textquotesingle}grasp\text{\textquotesingle}, s_2 = \text{\textquotesingle}ungrasp\text{\textquotesingle}\). It is consequently not necessary to (externally) designate them as parameters within the behavioral competence 'moving an entity \(n\) to the location \((x, y)\)'. Such variable constancy can always be straight-forwardly determined via sequential random instantiation of parameters.

We are thus able to remap the perceptual space \(p_{r_{K_0}+1}\) of novel behavioral competence \(C_{r_{K_0+1}}(p_{r_{K_0}+1})\) (i.e. 'place object \(n\) at \((x, y)\)') onto a parametrically smaller perceptual domain \(p_{r_{K_0}+1} \rightarrow p'_{r_{K_0}+1}\), where \(|p_{r_{K_0}+1}| = 5\) and \(|p'_{r_{K_0}+1}| = 3\). Furthermore, if there is any redundancy within the randomly-generated sequence (for instance, if we had obtained an inefficient, but goal-equivalent sequence;

\[ G(s_1), G(s_2), A(n_1), G(s_3), M(x_1, y_1), G(s_4), \]

with spurious initial grasping movements \(G(s_1), G(s_2))\), then a similar reduction may be achieved by randomly instantiating random parameter subsets and removing any unnecessary ones. Equally, this procedure can establish whether there exist functionally identical variables (such as, for instance, when a spatial parameter \(X_1\) requires the same input as a second, apparently differing, instantiation of a singular spatial variable; \(X_2\)). Thus, the previous random instantiation procedure enables dimensionality reducing projections of the form \((X_1, X_2) \rightarrow X_1\) (although only for variables of the same type - eg spatial ordinates).

The second major procedure for reduction of the parameter space dimensionality of acquired behavioral competences involves establishing whether con-
The experimental results show that the robot's fine-grained task performance was not significantly affected by the number of objects or the complexity of the task environment. The robot was able to learn and perform the tasks in a consistent manner across different trials, indicating the feasibility of using fine-grained task representations in robotic control systems.
the differing shading characteristics of raised and sunken entities, respectively. However, it is not yet the case that the distinction between pieces and holes has any semantic content; this is what we aim to accomplish with the entropy-based hierarchical perception-action reparameterization.

Initial determination of the primitive behavioral goals is accomplished via the scale-saliency algorithm, followed by gradient descent in the motor parameter space. Distractor entities are included in the perceptual domain (ie they have an index $n$ and corresponding $x$, $y$ and $\theta$ values), but which cannot be moved by the gripper arm (ie they are 'glued' to the table). In order to permit maximum generalizability, these are assumed to be visibly distinguishable from pieces and holes, and so have a different type allocation, $h_d$.

A human supervisor is observed solving the puzzle over a large number $O(10^3)$ of trials conducted via a 'drag and drop' mouse interface. The solution involves four principle stages of motor-competence (given the initial motor parameter space), which are given an explicit emphasis during training, by emphasizing their sequentially-applied nature. (i.e., the supervisor is instructed not to perform continuous composite movements during training, much as when training a human child). Other than this, no a priori knowledge is given to the system. The four key stages of competence are: Moving the gripper to, and aligning it with a given object; Moving a given object to a specified location, Inserting a given object into the appropriate hole, Solving the shape-sorter puzzle.

It is apparent that these competences form a subsumption hierarchy, with each level critically dependent on the level immediately beneath it. Each of the stages consequently requires a progressively complex concatenation of the primitive motor capabilities. Under normal circumstances, it would therefore be the case that the hierarchy of competences would have associated with it a progressively complex perceptual domain (at the first stage, the system observes oriented entities; at the second-stage, the system observes movable objects; at the third stage the system observes puzzle pieces; at the final stage the system simply observes a puzzle). Without a modification of the perceptual parameter spaces, the key perceptual entities at each stage would require highly complex description in terms of the a priori parameters.

The saliency-based perceptual-action hierarchy formation mechanism should therefore be able to identify these stages of behavioral competence autonomously, and make the appropriate modification to the perceptual space at each level of the hierarchy. In doing so, we expect that it will significantly reduce the parameter space associated with exploratory moves (which are defined at the apex of the hierarchy, and transmitted down the perception-action hierarchy, acquiring increasing amounts of perceptual context as the sub-goals are progressively defined).

To demonstrate this, we establish the total number of perception-action cycles required to achieve the various stages of competence for both a hierarchical learner, and a similar saliency-based perception-action learner, albeit without perceptual reparameterization capability. These results are outlined in figure 5.
We also indicate in figures 6 to 20 typical perceptual goals identified by the scale-saliency mechanism for a single spatial feature at each stage of the hierarchy. In each case, it is apparent that well-defined and relevant perceptual parameters are isolated by the algorithm.
**Fig. 7.** Single Feature Perceptual Goals for the Competence 'Aligning with an Object'

**Fig. 8.** Single Feature Perceptual Goals for the Competence 'Moving an Object' (5 Objects in Scene)

**Fig. 9.** Single Feature Perceptual Goals for the Competence 'Filling a Hole' (5 Holes in Scene)
Fig. 10. Single Feature Perceptual Goals for the Competence 'Solving the Puzzle'

4 Conclusions

We have presented a novel scale-saliency based technique for building perception-action hierarchies. The mechanism is contiguous over the entire hierarchy, and aims to identify both salient parameter sets and parameter indices within the perceptual domain. This enables an affordance-based characterization of the environment in which higher levels of the perception-action hierarchy represent increasingly task-specific and de-contextualized sensorimotor competences. Equivalently, this constitutes a mechanism for autonomous nesting and scheduling of sub-goals; in essence a subsumption hierarchy in the manner of Brooks [2], although constructed spontaneously according to the task requirements.

It is evident from figure 5 that the autonomous generation of an entropy-based saliency hierarchy has the further advantage of significantly more efficient learning than non-reparameterizing methodologies. In fact, it exhibits an essentially linear scaling in performance over the task-complexity hierarchy (similar to an earlier stochastic approach to autonomous construction of perception-action hierarchies [23]). Here, however, the perceptual goals are far more clearly defined for both the primitive and derived capabilities. The method is also significantly resilient to the presence of distractors.

Acknowledgement: The research leading to these results has received funding from the European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreement no 215078.

References


