Perceptually grounded meaning creation.

Luc Steels
Artificial Intelligence Laboratory
Vrije Universiteit Brussel
Pleinlaan 2, B-1050 Brussels, Belgium
E-mail: `steels@arti.vub.ac.be`
Draft, submitted to ICMAS, Kyoto 1996.

May 1, 1996

Abstract

The paper proposes a mechanism for the spontaneous formation of perceptually grounded meanings under the selectionist pressure of a discrimination task. The mechanism is defined formally and the results of some simulation experiments are reported.

Keywords: origins of meanings, self-organization, distributed agents, open systems.

1 Introduction

The research reported here is part of a larger research program to understand the origins of language and meaning using complex systems mechanisms such as self-organisation, co-evolution, and level formation [5]. This paper focuses on the meaning creation process. A theoretical model is proposed to explain how an autonomous agent may originate new meanings. The agent is autonomous in the sense that its ontology is not explicitly put in by a designer, nor is there any explicit instruction.

For the purpose of this paper, meaning is defined as a conceptualisation or categorisation of reality which is relevant from the viewpoint of the agent. Meanings can be expressed through language, although they need not be. Meaning takes many forms depending on the context and nature of the situation concerned. Some meanings (such as colors) are perceptually grounded. Others (such as social hierarchies) are grounded in social relations. Still others (such as goals or intentions for actions) are grounded in the behavioral interaction between the agent and the environment. This paper focuses on perceptually grounded meanings, although the proposed mechanism could also be used for other domains.

The proposed model is theoretical in the sense that no claim is made or evidence given that it is empirically valid for humans or animals. The goal is only to outline and validate possibilities. Independently of such a validation, applications where agents (software agents or robotic agents) autonomously have to make sense of their environment are already possible.

The present paper focuses on meaning creation in a single agent. Work is under way to also study meaning creation in multiple agents and investigate how a common language can act as a way to achieve a coherent conceptual framework between agents even though every agent individually builds up his own repertoire.

The rest of the paper is in four parts. The next section describes the approach. This is followed by a section which describes the proposed mechanisms more formally. Then some experimental results are reported. The final section contains some conclusions and a discussion of related work.

2 Approach

Agents engage in tasks relevant for their survival in a specific environment. In this paper, I focus on perceptually grounded discrimination tasks. The agent attempts to distinguish one object or situation from others using sensors and low-level sensory processes. The question is whether an agent is capable to develop autonomously a repertoire of features to succeed in discrimination and to adapt this repertoire when new objects are considered. A specific attempt to perform a discrimination and the subsequent adaptation of the feature repertoire is called a *discrimination game*.

Let us assume that there is a set of objects, or more generally situations,

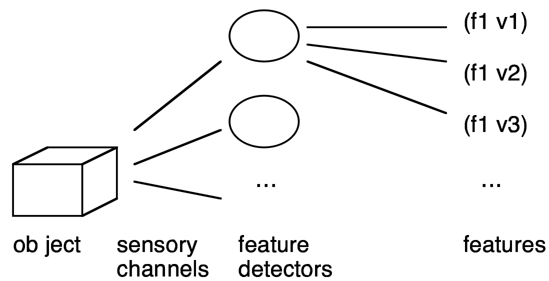

Figure 1: Feature perception is the process of going from an object to a feature set in two steps: sensory channels contain states from sensors and sensory routines, and they are transformed into features by feature detectors.

which have characteristics that are sensed through sensory channels, either derived straightly from sensors or from low level sensory routines. A sensory channel yields a value between 1.0 and 0.0. For example, the sensory channels could capture properties of moving objects like size, speed, average grey area, etc., or internal states reflecting motivations, sensations or actuator streams. We are conducting experiments in our laboratory with real mobile robots, speech, and active vision that yield a possible sensory basis for the mechanisms proposed here. In this paper, the meaning creation process is however studied abstractly without reference to specific applications.

A meaningful distinction takes the form of a *feature*, which decomposes into an attribute and a value. The feature is derived by a feature detector which discretises the continuous space of one sensory channel. The feature indicates that the value of a sensory channel falls within one subregion of the space (see fig 1.). There are absolute features, such as '(*color red*)', which are based on absolute values of a sensory channel for a single object, and relative meanings (such as '(*speed faster*)') which compare states of sensory channels for different objects. This paper only focuses on absolute features. A particular attribute is not necessarily relevant for each object.

The paper examines the hypothesis that the origins of meaning are based on construction and selection processes embedded in discrimination tasks. Each individual agent is assumed to be capable to construct new features, i.e. new segmentations of the continuous sensory space. The process of generating diversity and variation is subjected to selection pressure coming from the discrimination task: The agent attempts to differentiate an object

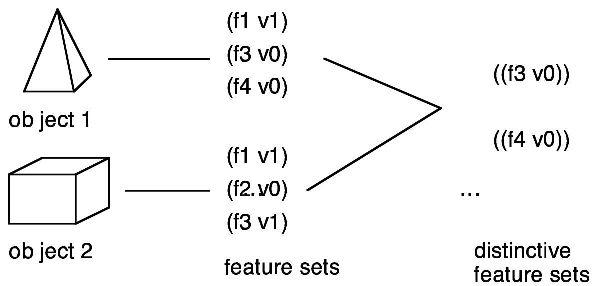

Figure 2: Discrimination is the process of comparing two feature sets to find the discriminating features.

from a set of other objects which constitute the context based on the available repertoire of features and values. A discrimination may be based on one or more features grouped as a *distinctive feature set*. There may be more than one possible distinctive feature set, but also none if not enough features are available. This happens either because no feature could be found to characterise the topic, or the attributes used to characterise the topic were not applicable to the other objects in the context, or a feature does not make a sufficiently fine-grained distinction. When there is no distinctive feature set, the discrimination fails and there is pressure to construct new feature detectors.

Feature detectors are refined in a hierarchical fashion and therefore form a kind of discrimination tree. The first detector divides the space up in some regions (in this paper always 2). This region might then later be segmented by an additional feature detector if objects that need to be discriminated fall within the same region. Thus feature-detectors form natural hierarchies, which go as deep as required.

The set of objects among which a discrimination has to take place is assumed to be open, in the sense that new objects may enter the environment that require different or more refined features.

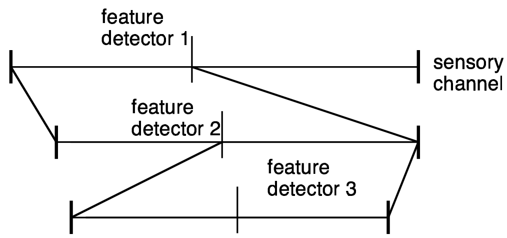

Figure 3: Feature detectors grow hierarchically as needed by the task domain.

3 Formal description of the mechanism

3.1 Terminology

Let there be a set of objects $\mathcal{O} = \{o_1, ..., o_m\}$ and a set of sensory channels $S = \{\sigma_1, ..., \sigma_n\}$, being real-valued functions over $\mathcal{O}$. Each function σ_j defines a value $0.0 \leq \sigma_j(o_i) \leq 1.0$ for each object o_i.

An agent a has a set of feature detectors $D_a = \{d_{a,1}, ..., d_{a,m}\}$. A *feature detector* $d_{a,k} = \langle p_{a,k}, V_{a,k}, \phi_{a,k}, \sigma_j \rangle$ has an attribute name $p_{a,k}$, a set of possible values $V_{a,k}$, a function $\phi_{a,k}$, and a sensory channel σ_j. The result of applying a feature detector $d_{a,k}$ to an object o_i is a feature written as a pair $(p_{a,k}\ v)$ where p is the attribute name and $v = \phi_{a,k}(\sigma_j(o_i)) \in V_{a,k}$ the value.

The *feature set* of a for o_i is defined as $F_{a,o_i} = \{(p_{a,k}\ v) \mid d_{a,k} \in D_a, d_{a,k} = \langle p_{a,k}, V_{a,k}, \phi_{a,k}, \sigma_j \rangle, v = \phi_{a,k}(\sigma_j(o_i))\}$. Two features $(a_1\ v_1)$, $(a_2\ v_2)$ are *distinctive* iff $a_1 = a_2$ *and* $v_1 \neq v_2$. A distinctive feature set D^C_{a,o_t} is a set of features distinguishing an object o_t from a set of other objects C. $D^C_{a,o_t} = \{f \mid f = (p\ v) \in F_{a,o_t}$ *and* $\forall o_c \in C$ *either* $\nexists f' = (p'\ v') \in F_{a,o_c}$ *with* $p = p'$ *or* $\exists f' \in F_{a,o_c}$*with* f *and* f' *distinctive*$\}$. Clearly there can be several distinctive feature sets for the same o_t and C, or none.

3.2 Discrimination games

A discrimination game $d =< a, o_t, C >$ involves an agent a, a topic $o_t \in \mathcal{O}$, and a context $C \subset \mathcal{O} \bigcap \{o_t\}$. The outcome of the game is twofold. Either a distinctive feature set could be found, $D^C_{a,o_t} \neq \emptyset$, and the game ends in success, or no such feature set could be found, $D^C_{a,o_t} = \emptyset$, and the game ends in failure.

As part of each game the repertoire of meanings is adjusted in the following way by the agent:

1. $D^C_{a,o_t} = \emptyset$, i.e. the game is unsuccessful. This implies that there are not enough distinctions and therefore $\forall o_c \in C, F_{a,o_t} \subseteq F_{a,o_c}$. There are two ways to remedy the situation:

   (a) If there are still sensory channels for which there are no feature detectors, a new feature detector may be constructed. This option is preferred.

   (b) Otherwise, an existing attribute may be refined by creating a new feature detector that further segments the region covered by one of the existing attributes.

2. $D^C_{a,o_t} \neq \emptyset$. In case there is more than one possibility, feature sets are ordered based on preference criteria. The 'best' feature set is chosen and used as outcome of the discrimination game. The record of use of the features which form part of the chosen set is augmented. The criteria are as follows:

   (a) The smallest set is preferred. Thus the least number of features are used.

   (b) In case of equal size, it is the set in which the features imply the smallest number of segmentations. Thus the most abstract features are chosen.

   (c) In case of equal depth of segmentation, it is the set of which the features have been used the most. This ensures that a minimal set of features develops.

The whole system is selectionist. Failure to discriminate creates pressure to create new feature detectors. However the new feature detector is not guaranteed to do the job. It will be tried (next time) and only thrive in the population of feature detectors if it is indeed successful in performing discriminations.

4 Implementation

The discrimination game defined above has been implemented and encapsulated as an agent. The programs create a set of sensory channels and an initial set of objects which have arbitrary values for some of the sensory channels. A typical example is the following list of objects and associated values for channels:

```
o-0: [sc-3:0.73] [sc-4:0.82] [sc-5:0.07]
o-1: [sc-0:0.89] [sc-3:0.02] [sc-4:0.56] [sc-6:0.48]
o-2: [sc-0:0.74] [sc-1:0.92] [sc-2:0.22] [sc-3:0.56]
        [sc-8:0.52] [sc-9:0.03]
o-3: [sc-2:0.36] [sc-3:0.09] [sc-4:0.14]
o-4: [sc-1:0.47] [sc-2:0.61] [sc-3:0.69] [sc-5:0.67]
         [sc-6:0.14] [sc-9:0.43]
o-5: [sc-1:0.84] [sc-4:0.82] [sc-5:0.70] [sc-8:0.81]
o-6: [sc-1:0.40] [sc-2:0.32] [sc-3:0.68] [sc-4:0.96]
        [sc-5:0.41] [sc-7:0.14] [sc-8:0.76]
o-7: [sc-1:0.84] [sc-2:0.89] [sc-3:0.63] [sc-8:0.41]
o-8: [sc-0:0.72] [sc-1:0.02] [sc-3:0.92] [sc-4:0.44]
        [sc-5:0.04] [sc-7:0.29]
o-9: [sc-1:0.35] [sc-2:0.72] [sc-3:0.58] [sc-4:0.34]
```

A feature detector is a function assigning a feature-value to a certain attribute. The name of the attribute indicates its nature. It is of the form $sc_i - n_1 - ...$ where i is the sensory channel followed by the number of segments of each consecutive segment. For example, sc-5-2 is the name of an attribute whose feature detector operates on sc-5 and divides it in 2 regions. sc-5-2-2 would be the name of an attribute that is a further refinement. (sc-5-2 v-0) is a feature combining this attribute with the value v-0.

In normal operation, the agent continuously goes through a loop performing the following activities:

1. A context is delineated. The context consists of the objects currently in the field of attention of the agent.

2. One object in this context is chosen randomly as topic.

3. The feature sets of the topic and the other objects in the context are derived.

4. An attempt is made to find possible discriminating feature sets.

We now show some typical situations for an agent a-5, which starts from no features at all. In the first game, a-5 tries to differentiate the object o-5 from o-3. The agent does not have a way yet to characterise the topic and creates a new attribute operating on sc-5.

```
a-5: o-5 <-> {o-3 }
Topic: NIL
Not enough features topic
New attribute: sc-5-2
```

The next game to distinguish o-5 from o-9 and o-1 is already successful, because o-5 is again the topic. The context contains objects that do not have any response for sc-5, and thus no features can be constructed:

```
a-5: o-5 <-> {o-9 o-1 }
Topic: ((sc-5-2 v-1))
Context: (NIL NIL)
Success: ((sc-5-2 v-1))
```

The next game is also sucessful because o-6 has value v-0 for sc-5-2, o-2 has nothing and o-5 has v-1.

```
a-5: o-6 <-> {o-2 o-5 }
Topic: ((sc-5-2 v-0))
Context: (NIL ((sc-5-2 v-1)))
Success: ((sc-5-2 v-0))
```

In the following game the attributes are not sufficiently distinctive and therefore a new attribute is created. As long as there are possibilities to focus on additional sensory channels, existing attributes will not be refined. The new attribute operates on sc-3.

```
a-5: o-7 <-> {o-1 o-2 }
Topic: ((sc-1-2 v-1))
Context: (NIL ((sc-1-2 v-1)))
No distinctive features but new one possible: (sc-2 sc-3 sc-8)
New attribute: sc-3-2
```

When uncovered sensory channels are no longer available, more refined feature detectors for existing attributes start to be made. In the following example, o-0 fails to be distinguished from o-8 and 0-1, even though a set of features is available to characterise each object. A refinement of the attribute operating over sc-5 is chosen.

```
a-5: o-0 <-> {o-8 o-1 }
Topic: ((sc-3-2 v-1)(sc-4-2 v-1)(sc-5-2 v-0))
Context: (((sc-0-2 v-1)(sc-1-2 v-0)(sc-3-2 v-1)
           (sc-4-2 v-0)(sc-5-2 v-0)))
          ((sc-0-2 v-1)(sc-3-2 v-0)(sc-4-2 v-1)))
No distinctive features but refinements possible.
Refining attribute: sc-5-2 => sc-5-2-2
```

After a sufficient number of discrimination games the set of features stabilises. For the set of objects given above, the following is a stable discrimination tree. For each attributes the possible values are listed with their corresponding regions as well as the number of times a feature has been used.

```
sc-5-2:
  v-0: [0.00 0.50] 358.
   sc-5-2-2:
     v-0: [0.00 0.25] 31.
      sc-5-2-2-2:
        v-0: [0.00 0.12]
         sc-5-2-2-2-2:
           v-0: [0.00 0.06] ; v-1: [0.06 0.12] 3.
        v-1: [0.12 0.25]
     v-1: [0.25 0.50] 22.
  v-1: [0.50 1.00] 309.
sc-1-2:
  v-0: [0.00 0.50] 651. ;  v-1: [0.50 1.00] 628.
sc-3-2:
  v-0: [0.00 0.50] 713. ; v-1: [0.50 1.00] 733.
sc-8-2:
  v-0: [0.00 0.50] 15. ; v-1: [0.50 1.00] 8.
sc-2-2:
  v-0: [0.00 0.50] 99. ; v-1: [0.50 1.00] 112.
```

```
sc-0-2:
  v-0: [0.00 0.50] ;  v-1: [0.50 1.00] 42.
sc-4-2:
  v-0: [0.00 0.50] 223.
   sc-4-2-2:
     v-0: [0.00 0.25]
     v-1: [0.25 0.50] 1.
      (att a-5 sc-4 2 2 2):
        v-0: [0.25 0.37] 5.;  v-1: [0.37 0.50] 5.
  v-1: [0.50 1.00] 215.
   sc-4-2-2:
     v-0: [0.50 0.75] 1.
     v-1: [0.75 1.00] 2.
      sc-4-2-2-2:
        v-0: [0.75 0.87] 5. ; v-1: [0.87 1.00] 2.
sc-6-2:
  v-0: [0.00 0.50] 2. ; v-1: [0.50 1.00]
```

We see that more abstract features, like *(sc-1-2 v-0)*, are used more often. For some, like *(sc-5-2 v-0)*, there is a deep further discrimination. For others, like *(sc-5-2 v-1)*, there is none. Some features, like *(sc-6-2 v-1)*, have not been used at all and could therefore be eliminated. Another experiment with the same objects but for a different agent a-6 yields a different discrimination tree. In one example, some sensory channels (such as sc-6) were not used, sc-4 was no longer refined, etc. Usually there are indeed many different possibilities and an important question for further study is how optimal the discrimination trees obtained with the proposed mechanism are.

When new objects enter the environment, the agent should construct new distinctions if they are necessary. This is effectively what happens. If new sensory channels become available, for example because a new sensory routine has become active, then it will be exploited if the need arises.

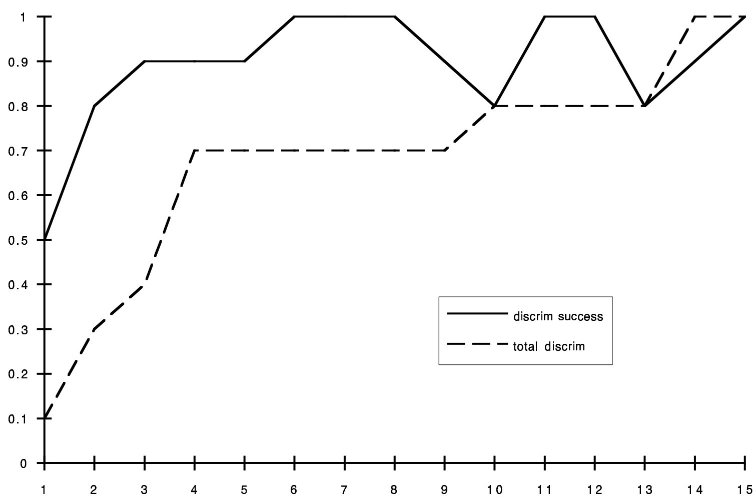

Figure 4: The graph shows the evolution of the discriminatory capacities of a single agent. The total number of objects (10) is fixed. There are 5 sensory channels. The average success in discrimination games as well as the global success is shown on the y-axis. The number of discrimination is mapped on the x-axis (scale 1/10). All objects can be discriminated after 150 discrimination games.

5 Experimental Results

5.1 Fixed set of objects

Fig 4. shows a typical example where an agent builds up a repertoire of feature detectors, starting from scratch. The graph shows the increasing discrimination success as experienced by the agent in discrimination games. It also shows the global success with the features so far, i.e. all objects are compared to all other objects only based on their features. Progress in finding more discriminatory features depends on encountering those objects that require more discrimination. Because context and topic are set probabilistically, this is not predictable.

The graph in fig 5. shows for the same experiment the increasing number of features (as a percentage of the final total (22) features reached at the end of the experiment), and the percentage of features that is effectively used. We see that many features created earlier on are only gradually used and

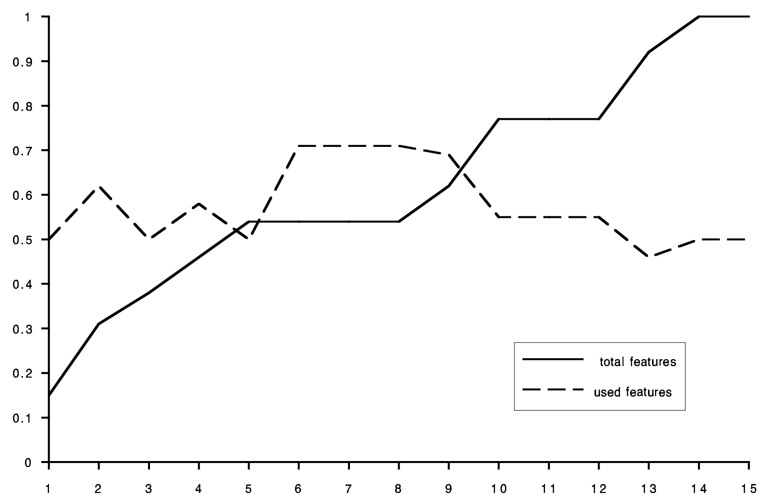

Figure 5: The graph plots data for the same experiment as in fig 1. The total number of features and the percentage of features used of the total available at each time moment.

there are still many cases that have not been encountered.

5.2 Increasing the set of objects

In the next experiment (fig 6.) we start from a set of 10 objects and gradually add new objects in a probabilistic fashion, to reach a total of 50 objects. We see that the feature repertoire is extended occasionally. The average discrimination success remains close to the maximum (1.0) because new objects are only encountered occasionally and the feature detectors already constructed are general.

Fig 7. shows for the same experiment the relation between the total number of features that are available and the features that are used. We see that the repertoire of features created in the beginning is used much more extensively, clearly showing

Initially not many new features are introduced but the available repertoire is used better. Later on new features are indeed necessary.

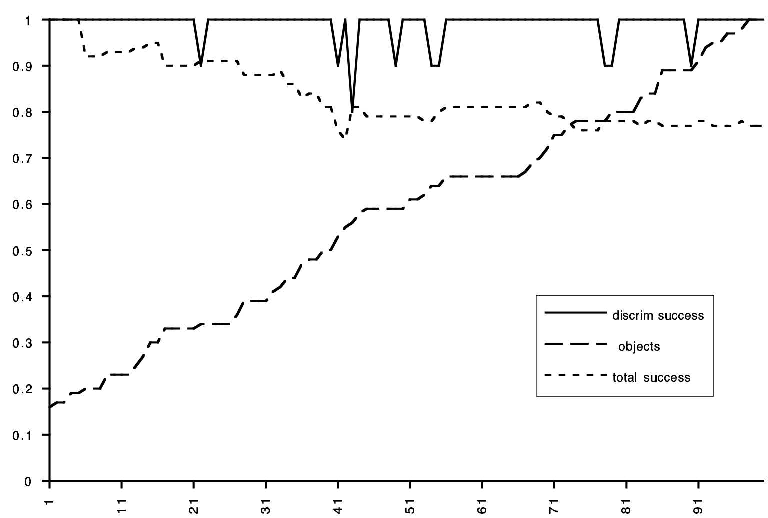

Figure 6: Graph showing a steady increase in the number of objects. The graph shows on the y-axis the number of objects (as a percentage of the total reached at the end, i.e. 50), the discriminatory success which remains close to the maximum, and the number of features (as a percentage of the total reached at the end, i.e. 35). The x-axis plots the number of discrimination games (scale 1/10).

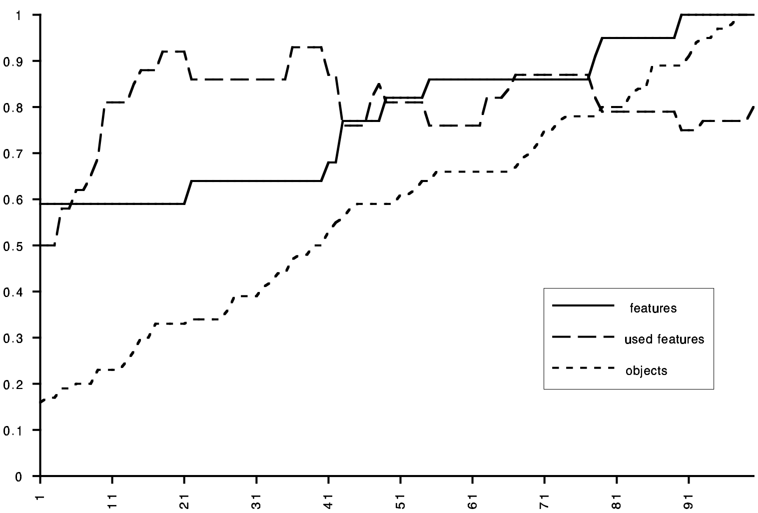

Figure 7: Graph showing (on the y-axis) the relation between the increasing total feature repertoire and the percentage of the available repertoire that is used. The x-axis plots the number of discrimination games (scale 1/10).

6 Conclusions

The paper proposed a mechanism for the creation of perceptually grounded meaning giving a set of sensory channels and a series of objects among which discrimination has to take place. The mechanism is based on selectionist principles. There is a generator of variety and selection pressure coming from success or failure in discrimination. It was shown that the system arrives quite rapidly at a set of possible features for discriminating objects. Most interestingly, the system remains adaptive when new objects are added or when new sensory channels become available. Further work is obviously required, particularly in the context of concrete applications where the sensory channels are linked to visual, auditory, or internal sensors.

There has been a lot of other work on the problem of meaning creation, particularly in the connectionist literature [4]. A perceptron for example can be seen as a device that acquires a set of distinctions as relevant for a classification task. The sensory channels constitute the inputs to the perceptron, and the weights perform the function of selecting out regions which will be input for the classification process. The most important differences between these connectionist proposals and what has been presented here is that (1) connectionist networks embed the build up of a feature repertoire within the task of classification (as opposed to discrimination) and (2) an inductive/instructional approach as opposed to selectionist approach is used. An inductive approach is based on going through a (typically large) set of examples which drives the weights stepwise to reflect the best classification. In a selectionist approach a structure comes into existence by variation or construction and is then tested as a whole for fitness in the environment. Inductive approaches result in gradual generalisation. Selectionism gives immediately generalisations which might be refined gradually.

The selectionist approach followed here is more in tune with work on feature generation in genetic algorithms research [3], unsupervised learning as exemplified by the Kohonen network [2], and most importantly proposals made by Edelman known as Neural Darwinism [1]. Edelman assumes that neuronal growth processes yield a primary repertoire stabilised by developmental selection, which is then subjected to experiential selection, yielding a secondary repertoire of categories. Using re-entrant maps and degeneracy, categorial perceptions of different objects can be compared and generalised to classes. Meaning creation and classification are clearly distinct here. The selectionist pressure in the Edelman case comes from statistical signal correlations (for the formation of the secondary repertoire) and similarity matching (for the formation of classes). In this work, the selectionist pressure comes from a discrimination task. Nevertheless, the neural machinery proposed by Edelman (spontaneous variation, selection, re-entrant mapping) is probably adequate for a neural implementation of the mechanisms proposed here.

7 Acknowledgement

The research and writing of this paper has been financed by the Belgian Federal government FKFO project on emergent functionality (FKFO contract nr. G.0014.95) and the IUAP 'Construct' Project (nr. 20) of the Belgian government, with additional support from the external researcher program of the Sony Computer Science Laboratory in Tokyo.

References

[1] Edelman, G.M. (1987) Neural Darwinism: The Theory of Neuronal Group Selection. Basic Books, New York.

[2] Kohonen, T. (1988) Self-Organization and Associative Memory. Springer Series in Information Sciences. Vol 8. Springer Verlag, Berlin.

[3] Koza, J. (1992) Genetic Programming. MIT Press, Cambridge Ma.

[4] McClelland, J.L. and D.E. Rumelhart (eds.) (1986) Explorations in Parallel Distributed Processing. MIT Press/Bradford Books, Cambridge Ma.

[5] Steels, L. (1996c) Synthesising the origins of language and meaning using co-evolution and self-organisation. In: Hurford, J. (ed.) Evolution of Human Language. Edinburgh Univ. Press. Edinburgh.