Hanna S. (2005) Where Creativity Comes From: the social spaces of embodied minds, *Proceedings of HI'05, Computational and Cognitive Models of Creative Design*. University of Sydney. pp. 45-70. ISBN: 1-86487-789-8

WHERE CREATIVITY COMES FROM

The social spaces of embodied minds

SEAN HANNA
University College London, United Kingdom.

Abstract: This paper explores creative design, social interaction and perception. It proposes that creativity at a social level is not a result of many individuals trying to be creative at a personal level, but occurs naturally in the social interaction between comparatively simple minds embodied in a complex world. Particle swarm algorithms can model group interaction in shared spaces, but design space is not necessarily one pre-defined space of set parameters on which everyone can agree, as individual minds are very different. A computational model is proposed that allows a similar swarm to occur between spaces of different description and even dimensionality.

1 Introduction

Where does creativity originate? Does the push to innovate come from the individual, as our own point of view suggests? Or is it the group?

It is the natural assumption that we are the instigators. Margaret Boden (1990) draws a distinction between the psychological ‘P-creativity’ of the individual, and the historical ‘H-creativity’ of ideas that are fundamentally novel for the whole of a culture. Referring to Alan Turing, Friedrich von Kekulé, Mozart, and other historical innovators, she states the capacity for P-creative ideas means there is a good chance for H-creativity. Implied in the usual view is both that the creativity of the group is an accumulation of creative leaps of individual people, and that these individuals are varied in ability. H-creativity is wholly dependent on P-creativity.

This paper proposes an alternative: that creativity at a social level is not a result of many individuals trying to be creative at a personal level, but emerges naturally from the social interaction between comparatively simple minds embodied in a complex world. While individuals may differ in ability, such differences are not necessary for variety and innovation to occur. Nor is group innovation based on randomness, or even an internal drive to generate novelty, but a continual change in how we perceive the world around us and how we are affected by others.

A computational model is proposed that takes after the coordinated social dynamics of bee swarms, ant colonies, schools of fish and flocks of birds, all of which appear in motion to have a sophisticated group mind, and the swarm algorithms inspired by these. The interactive creativity of large groups of designers can be seen as a type of swarm behaviour involving the thoughts of agents (Kennedy & Eberhart, 2001), but models to explain this require the predefinition of a single uniform problem space in which to interact. Designers' minds can be vastly different, and private, so what is this space? In this proposed model swarming can occur, but takes place via three distinct, abstract spaces aligned with Czikszentmihalyi's (1988) systems model of creativity, and linked in such a way that creativity happens in the process of mapping from one space to the next.

The practical aims of this work are twofold. The first is to show how the parts of a creative system (based on Czikszentmihalyi's model) can be represented as a set of three abstract spaces with a structure that allows swarming to occur. The second is to use these to model a social system that appears to exhibit creative behaviours: namely cultural innovation, coalescence to socially established norms, and clique differentiation.

The overall goal is to show that the first is sufficient cause for the next: that is, to show that the structure of the spaces yields an emergent behaviour resembling creativity. In doing so, the model proposed and tested is minimal, and omits features often associated with a system of creative agents. The agents are not goal-motivated, novelty-seeking or of varied abilities. They have no intentionality other than the differentiation of perceptions required by neurons or machine classifiers. There is no objective measurement of utility to rate one example of creative output against the next. Also, there are no random processes to introduce innovation internal or external to the agents. This does not deny the fact that these features may exist in real, human designers, but it suggests that these are not strictly necessary to produce the creative behaviours mentioned.

The design context is a basic architectural one, with each agent able to contribute to an overall building pattern in a shared world. As there is no objective measure of utility, *novelty* is measured relative to agents' prior work, and *innovation* is simply defined as a difference between an agent's or a group's building pattern and what has come before. The overall movement of group activity in the model ensures that styles and cultural norms are always changing, and as this happens, the innovation of individuals is a result of trying to make sense of the new world. In trying to choose what they see as the norm, they create novelty unintentionally.

The following section of the paper gives a brief background to swarm models and creativity in a social context. This is followed by an example design situation to illustrate the process of single individual learning and designing to cultural norms. The model of multi-agent behaviour is next introduced, and the spaces in which it occurs described in detail, to outline how embodiment in a shared environment allows differently constructed minds to swarm together. Finally the behaviour of these agents over time is reviewed to show how the diversity of agents' perception can result in cliques, interdisciplinary activity and cultural innovation.

2 Background

2.1 DESIGN AS EXPLORATION AND SELECTION OF AFFORDANCES

While optimisation problems typically begin with a preset objective, design is acknowledged as an exploratory process, without necessarily fixed goals. Rosenman (1997) uses the lack of predetermination to define creative design, suggesting "the lesser the knowledge about existing relationships between the requirements and the form to satisfy those requirements, the more a design problem tends toward creative design". Gero (1993) goes further to suggest "… exploration in design can be characterised as a process which creates new design state spaces", changing the framework in which optimisation occurs. But optimisation methods can satisfy this creative requirement. Maher and Poon (1996), for example, use a genetic algorithm to co-evolve goals and solutions simultaneously as changing 'problem' and 'solution' spaces.

2.1.1 Embodiment and affordances in creative systems

The alternative to representing problem spaces is to accept the space of the environment as its own representation. The act of being 'in the world' described as *embodiment* (Quick et al. 1999, Dourish 2001) or *structural coupling* (Maturana and Varela, 1987) requires that there is two way perturbation between an individual and the environment. As the designer makes a design, the world is also affecting his or her brain.

Design for an embodied individual consists of the creation of a product that will also be embodied in the environment. Realisation of the design in the world is not only for communication to others, but is an intrinsic necessity of the creative act, in that the internal representation of a concept in the mind of an individual is so unique to that individual, so different from the external world, that the concept itself can not be said to fully exist until it is embodied. The idea in an artist's mind of a painting is not the same as what is actually painted, as this may be affected by external events during the act of painting.

This ongoing negotiation can be seen as a process of constantly choosing between the perceived *affordances* of the work at each point in its evolution. Related to Heidegger's notion of *zuhanden* (ready-to-hand), Gibson (1979) coined the term *affordance* to refer to the properties the environment offers an animal in terms of action, such as the support afforded by a flat, horizontal surface. Norman (1988) discusses affordances in design, but chiefly in relation to how designs are used, as it is a principle of good design practice to be aware of the affordances that will suggest themselves to the user: a chair affords sitting, a handle lifting, pulling, opening, etc. Affordances can also be considered in terms of how we interact with the world in the process of making a design. Tang and Gero (2001) suggest the act of sketching, with constant drawing and re-evaluation is such a process, and the explicit representation of choices as decision trees has also been implemented in CAD environments (Brockman and Director 1991). At any stage of the process there are only certain possibilities open to the designer, and the act of design can be seen in this sense as a selection from the afforded alternatives.

2.2 SWARM AND GROUP INTELLIGENCE

2.2.1 Simple agents in abstract spaces

Particle swarm algorithms have also been applied to the exploration of design problem spaces. Several varieties exist, all sharing the principle that an emergent intelligence arises from the interaction of groups of simple agents, each of which behaves according to simple rules and has no knowledge of the global behaviour of the group.

Axelrod's model of the dissemination of culture (Axelrod, 1997) allows agents in fixed locations to communicate with one another with a probability equal to their cultural similarity, resulting in a polarisation of cultures as these similarities are strengthened. Individuals split into stable non-communicating groups. The same rules applied to freely moving agents in a space result in swarm behaviour. Three broad types of swarm model have been proposed: deterministic groups resembling complex behaviour such as that in bird flocks (Reynolds, 1987), and optimisation algorithms either representing stochastic particles with a goal (Kennedy & Eberhart, 2001), or those modelled on the (not necessarily spatial) social communication of groups such as ant colonies (Dorigo, 1997). Although the methods for optimisation incorporate randomness into the agents' paths, it is purposely left out of the model proposed here.

The rules of the algorithms differ in their specifics, but interactions between agents take one of two basic forms: they either *attract* or *repel* one another in space. The update of an artificial bird's velocity in (Reynolds, 1987), for instance, would be of the form:

$$\mathbf{v}_{t+1} = \mu\mathbf{v}_t + (1-\mu)(w_{\mathrm{avoid}}\mathbf{v}_{\mathrm{avoid}} + w_{\mathrm{match}}\mathbf{v}_{\mathrm{match}} + w_{\mathrm{centre}}\mathbf{v}_{\mathrm{centre}}), \qquad (1)$$

where $\mathbf{v}_{\mathrm{match}}$ and $\mathbf{v}_{\mathrm{centre}}$ cause the agent to imitate the others and $\mathbf{v}_{\mathrm{avoid}}$ to keep away from its neighbours. Attraction allows for a focused local search and exchange of information with similarly inclined neighbours, and repulsion causes groups and individuals to explore new areas of the space.

2.2.2 More complex Agents in the world

Designers influence one another in the abstract space of their work, so it is necessary for this to be embodied in a shared world. Nehaniv and Dautenhahn (1999) suggest an algebraic framework for imitation in which dissimilar bodies can imitate one another by producing similar *effects on the environment*. Individual actions, or internal representations are not important, but rather the ability to meet a series of sub-goals such as covering a wall with paint when imitating the task of painting. This stresses embodiment in the environment, but requires that the goals – the real motivators of creativity – must be predetermined explicitly.

Luc Steels' Talking Heads project shows that even these goals can be determined, in a point and guessing game played by robots that evolve a language. (Steels, 2000) The two are also embodied, and must communicate through a shared environment consisting of coloured geometric shapes and a white board. Words and even syntax are generated from the need to express concepts the robots may hold differently in their minds, the creativity springing from an interruption in full communication between the two agents. Edwin Hutchins' research with parallel constraint satisfaction networks suggests this interruption is actually beneficial (Hutchins, 1995). In highly connected networks of individuals, his populations reached poorer solutions than networks in which individuals were connected only moderately to one another.

2.3 CZIKSZENTMIHALYI'S SYSTEMS MODEL OF CREATIVITY

A widely accepted dynamic model of the process of creativity within a broader environment of other individuals is given by Czikszentmihalyi (Figure 1). This gives an account of the flow of ideas and interaction between a *person* (the creative individual), the *field* (the group of individuals that act as arbiters of creative output) and the *domain* (the collection of embodied work and symbolic representations deemed relevant by the field). (Czikszentmihalyi 1988)

Czikszentmihalyi's model is widely accepted as describing the social structure of the creative process, and the model proposed here will suggest that the activity in each section of the triangle takes place in one of three very different spaces, each of which can be mapped into the next as indicated by the arrows. As Czikszentmihalyi makes clear, the creative act is not an occurrence within the mind of an isolated individual, but an interaction with the domain and field, both of which are spaces outside the individual's private perception, and both of which may be shared by other individuals.

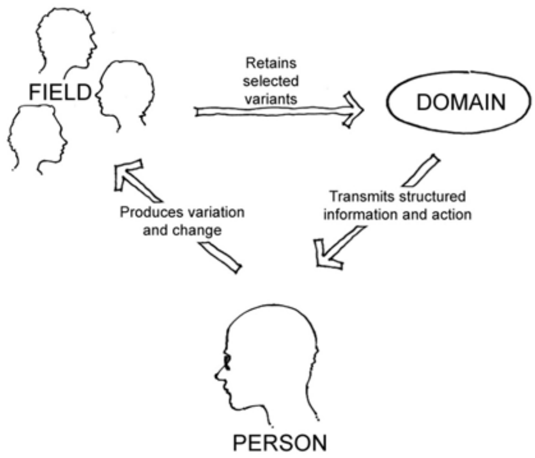

Figure 1. Czikszentmihalyi's systems model of creativity.

3 Building by an individual

3.1 ILLUSTRATION: A SAMPLE PROBLEM

A sample problem will be used to explain the proposed model with a real-world instance of collective design: that of the aggregation of buildings in towns built over time without a central plan. In this case the process is much like the swarm models introduced in Sec. 2.2.1, in that building is done by a number of people considering their own needs individually, yet a global order emerges from their effort.

Bill Hillier's analysis of hamlets in southern France is an exposition of this emergent order, in what he has termed the 'beady ring'. (Hillier and Hanson, 1984) As settlements reach a certain size a regularity appears in the shape of a ring of joined spaces around a central clump of buildings facing outward, and several inward facing groups of buildings around the perimeter. While this pattern seems to appear in every one of the hamlets studied, it need not be the result of any kind of central planning, as a simplified model makes clear.

A grid is used to simplify the geometry, and two types of objects are represented in its squares: closed cells with an orientation defined by an entrance in one side, and open voids. The minimal building unit is made up then of two face-wise adjacent squares, with a closed cell facing on to an open space in front. The model allows these pairs to aggregate so that each new pair must join its open cell to at least one other open cell already placed, and the closed cell does not join another closed cell only at the vertex. Other than these two rules, the position and orientation of each new unit is completely random, but each time the model is run, the overall structure of the aggregation forms that of the beady ring settlements studied, with a chain of open spaces onto which inner and outer groups of buildings face.

As a basic framework for town generation this is enlightening, but creative design is far more than random search, and equally important are the specific differences between towns. In Hillier's own study he notes the differences between the beady rings of France, and their counterpart villages in England that tend toward a more linear arrangement. These cultural differences in global form are also a result of the same uncoordinated local actions over time, yet the particular nuances that lead to a circular or a linear arrangement seem somehow to have been instilled into the individual members of the culture that make each building. The overall structure of this grid model will be used below to show that such cultural norms can be learned by individuals.

3.2 DESIGN BY AFFORDANCES IN TWO CULTURES

Design is an act of continually making and then examining from a different point of view. Embodiment in the world allows this. The act of design by an individual agent in the above context is simply the selection from a set of possible alternatives: "which is the most like X?" But how do we know what X is?

Assuming every act of construction is selfish and uncoordinated, the motivation behind the decisions of building placement would be to maximise some particular qualities considered to be important, such as direct access to the public space of the town and the economy of sharing a wall with a neighbour in the rules above. Each of these is an affordance of particular vacant building sites available at any given time. Rather than predetermining which of these qualities are most desirable however, suppose they are relative and change from culture to culture. Each time a unit is built, the configuration of the surrounding neighbourhood relative to the cell pair gives an ideal example for an individual to follow, another example in the domain.

Two artificial cultural norms were established that were easily distinguishable from one another, and a simple algorithm written to aggregate open/closed pairs of units in the manner of each (Figure 2). The first is a strict arrangement of straight rows rather like highly planned settlements such as Manhattan, and the second is a completely random arrangement of units joined open cell to open cell. To learn the two ideals, a classification algorithm can be trained on the units as they are built. Each time a new pair is placed in the plan, the 7×7 cell square surrounding the open half of the doublet is taken as its neighbourhood, and oriented such that it is always seen by an agent looking from the open cell toward the closed. The 49 cells, each containing either a closed building (indicated by a filled cell, or 1), a public open space (a dot, or -1) or yet unbuilt (an empty cell, or 0) are used as an agent's sensory experience of that particular example in the domain.

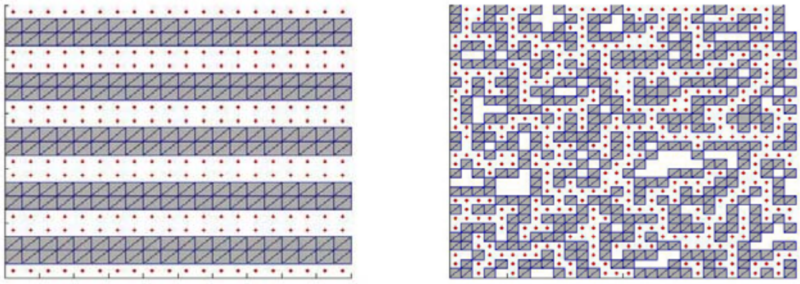

Figure 2. Building patterns of two cultures: strict rows and random aggregation.

The algorithm to build based on this is simple. The two swarm rules as applied to design say something like 'try to make something like what is accepted by your culture or group' (i.e. attraction toward the perceived mean) and 'try to distinguish yourself as far as possible from the other culture' (i.e. repulsion). Each building decision is an exercise in maximizing the qualities seen as desirable based on other examples in the domain. At every step a given number of positions and orientations are available to be built, and the decision is simply the act of choosing which one of these affordances best fits the ideal the agent has learned.

A support vector machine (SVM) was used as the agent's initial classifier due to its easily tuneable parameters. SVMs operate by finding a maximally separating hyperplane between the two labelled classes in a higher dimensional representation of the input, given in this case by a weighted sum of the non-linear Gaussian kernel function

$$\varphi(\boldsymbol{x}, \boldsymbol{\mu}) = \exp[\, -\, \| \, \boldsymbol{x} - \boldsymbol{\mu} \, \|^2 / \, \sigma^2 \,] \qquad (2)$$

with a parameter σ^2 that can be adjusted – in this case the variance of the Gaussian. Figure 3 shows the results for $\sigma^2 = 5$, 15 and 25 respectively. The SVM output is plotted (left column) with the vertical axis indicating the output values, and 450 row examples followed by 450 random aggregation examples positioned along the horizontal axis. The resulting agents' construction over time is shown for each, with the agent's attempt at replicating the rows (centre) followed by the random aggregation (right). At each construction step, the possible construction sites and orientations are evaluated by the SVM, and the one closest the mean of either culture as learned is selected. It is evident from the results that as σ^2 increases there is both better separation between the two groups by the SVM, and also a clearer construction result – more obvious in the rows than in the random arrangement. But this separation is never quite enough, and the classifier can only be seen to produce really adequate rows with an artificially created set of 'perfect' examples of row neighbourhoods is used, all identical so that each is exactly perceived as the ideal mean (Figure 4.)

Although there is a vast difference between the ability of the four agents above, their interaction with the environment always entails a similar kind of choice. It is this common environment that allows different agents to interact with one another, as investigated in the following section.

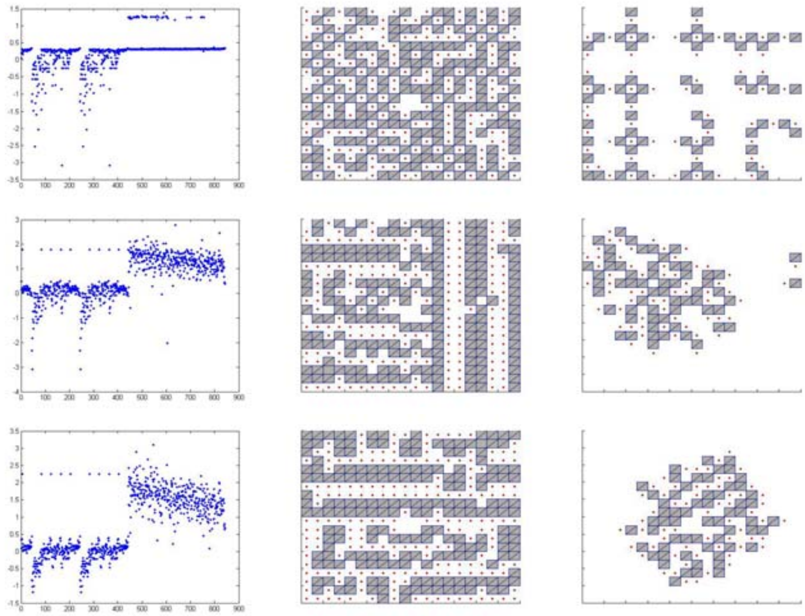

Figure 3. Building results for agents trained with a SVM: σ^2 = 5, 15 and 25. SVM output on 800 examples is shown at left, building patterns based on rows in centre, and building based on random aggregation at right.

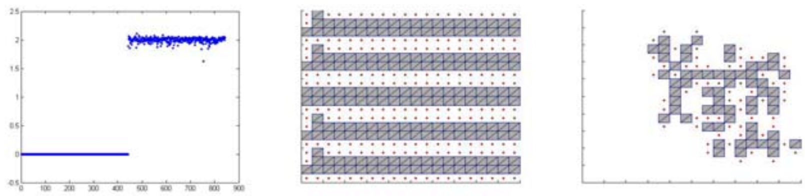

Figure 4. The same training on a set of 'ideal' examples.

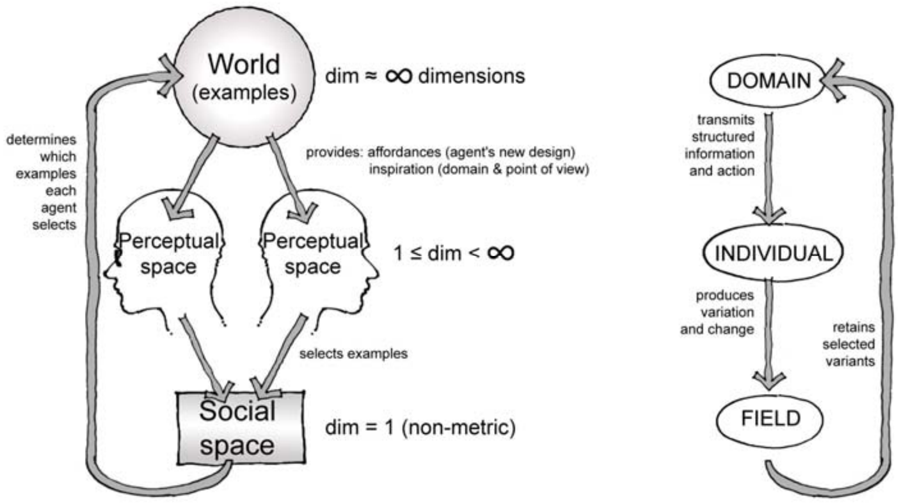

Figure 5. Three spaces of the model (left) paired with Czikszentmihalyi's (right).

4 Building by groups: a swarm in three spaces

4.1 STRUCTURE OF THE MODEL

Three distinct spaces form the arena for the collective swarm. The nature of these spaces and the way they are connected are the basis of this model (Figure 5). Creativity will not occur in any one of them, but in the cycle of mapping from one to the next. We can refer to these via Czikszentmihalyi's model, as they correspond roughly with its *domain, person* and *field,* and communication occurs between them in the same direction, but we will define the spaces somewhat more broadly in terms of what they can contain.

- *Space of the world*: Beginning with our shared, physical reality, this contains the set of embodied artefacts or events that can be experienced by all. These examples are available to any agent also embodied in the universe, and are anything that is objective, out there in the world, including even sounds and instances of symbols. This space includes any object or communication made by an individual, so it will be called the *world*, leaving *domain* to refer to selected sets within it.
- *Perceptual space*: Next, a subset of this reality is experienced by a given individual as subjective experience, in a second space of lower dimensionality we can call the *perceptual space*. This individual's perceptual space is unique, and therefore yields a unique picture of objective reality.

- *Social space*: The third space is the space in which the swarm dynamic can be seen: the social space of a shared culture. This is not necessarily a metric space, but can be represented as a graph of distances between individuals. The group of individuals closest to an agent would be that agent's *field*.

As communication occurs from one to the next, a point in one space can be mapped to the next space in the model. But due to the structure of the spaces, this can only occur in one direction. Most importantly, for one individual to communicate with another they must complete a full cycle, using the conventions of their social space to make an example in the world that the second individual can see with a different perceptual framework. Design as an act of selection from affordances as introduced in Sec. 3.2 allows this. This may seem a laborious process, but as with Steels' robots or in Czikszentmihalyi's model, it provides the framework for innovation.

4.2 SPACE OF THE WORLD

The example units placed in the above exercise constitute the domain in the space of the real world for our design agents, and a subset of all possible dimensions is used to represent the examples. Because the sensory input of each agent is an identical 49 square grid, each unique example can be represented by a point in a 49-dimensional space, a 2-d projection of which is shown in Figure 6. All example neighbourhoods are projected onto the first two principle components of the set: neighbourhoods of the straight rows are indicated by '×', and the random culture by '○' markers in the centre.

The choice of a 49-dimensional world for these examples is one of computational tractability, but in our own experience, each example from which a designer can be inspired can be represented as a point in a potentially infinite-dimensional space of physical reality.

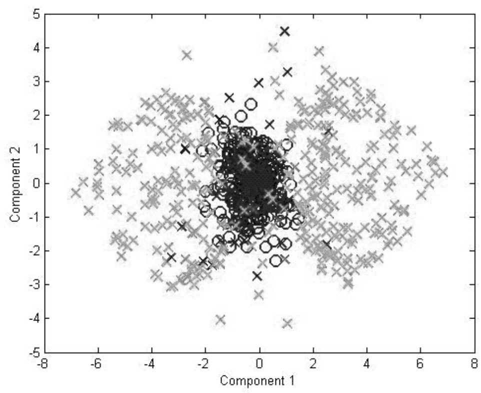

Figure 6. A projection of examples in a 49-dimensional world space.

The dimensionality chosen for the world is also, by necessity, higher than the agent's perceptual space. While the point that represents a given example has a potentially limitless set of qualities that may be experienced, our senses limit us to only a small set. But this inexhaustibility of dimensions is the chief advantage of being embodied in the world – an object in the domain such as a painting or sculpture can be revisited again and again and always be seen as something new, and the same is true of intangible arts such as music and literature.

The higher dimensionality of the space of the world compared to the structure of our perceptual space does not allow a thought to be uniquely mapped into reality. This might seem to pose a problem for designers, but not if the act of design is seen as the repeated selection from a set of affordances as they present themselves. The outline of design as given in section 3.2 requires only a choice from the possibilities as they appear in the agent's perceptual space.

4.3 PERCEPTUAL SPACE

Perceptual space, at any one time, is necessarily a lower-dimensional abstraction of the world space, which determines what we actually *see* in the world. In the case of the design agents in section 3.2, it is the one-dimensional output of the SVM. Any example in the real world can be mapped (via brain or learning algorithm) to a corresponding point in the space of an *individual's* perceptual space, which is different from every other individual's. Unlike the shared space if the world, this perception is completely private.

A perceptual space is represented in us by the state of our brain when experiencing the example. The nature of this space is like what has been termed *semantic space* by linguists. Semantic space research has shown that words and concepts can be mapped in a multidimensional space such that words with similar meanings are located near one another. Burgess and Lund (1988) have built a model of such a space based on the proximity of words to one another in Usenet discussions in which nouns such as 'dog' and 'cat' fall into one cluster, 'china' and 'america' together in another in a manner that coincides intuitively with many people's experience of the concept. This perceptual framework is the subjective counterpart to the space of external examples, the interior space of qualia or private meaning. A thing may exist in the world independently, but is actually experienced in the perceptual space of the mind.

Perception is an act of differentiation. Neurons react to *changes* in stimulus. It has long been known that the intensity of sensation is proportional to the frequency of neural activity, and that this decreases with time after the change (Adrian, 1928). Thus white noise, or the hue and intensity of ambient light, etc. are only perceived against a contrasting other. When we see, we group the continuous spectrum of visible light into distinct colours, so perceptions of colours closer to green are seen as green, closer to blue as blue.

This act of perception is analogous to the two swarming rules in that perceptions shift closer to a perceived mean (e.g. green), and away from the other label (blue or yellow). This model proposes that the equivalent of agent movement in a traditional particle swarm occurs in this perceptual space, by changing the perceptual mapping. The swarm rules of attraction and repulsion are rooted here in the perceptual system of the individual simply to allow that individual to differentiate or classify effectively. In the SVM examples in section 3.2 it was seen that the most effective mappings present examples of one group closest to the ideal mean (attraction) and those of the other farther away (repulsion), resulting in greater differentiation between the two sets of examples, and a clearer reconstruction of that culture. The same was found of the neural networks used to map to this perceptual space.

4.3.1 Representing the individual's point of view

The behaviours of attraction and repulsion were implemented in the perceptual space via an artificial neural network. Training of the network serves to usefully illustrate this movement, as the function of every neuron:

$$\mathbf{y} = \mathbf{wx} + w_{0,} \tag{3}$$

is a simple linear function, and can be visualised as a hyperplane constantly moving in a high dimensional space in an attempt to separate the two classes of input examples. The agent's point of view can be pictured as aligned to this hyperplane and moving with it.

A three layer neural network was used to map to the agent's perceptual space, with 49 input nodes corresponding to the state of the neighbourhood, 50 nodes in the hidden layer, and a single, linear output that rates each example in a single dimension that represents the internal perception of the agent. Training was conducted by exposing the network to 450 examples from each of the two cultures and backpropagation of errors.

Several classification methods were tested. The typical error function

$$J = ½ \, || \, \mathbf{t} - \mathbf{z} \, ||^2 \tag{4}$$

operates on the difference between the neural output **z** and a specified target **t**, set at 0 for the domain examples and 1 for the others. Alternatively no target was set for examples outside the domain, and for these either the normal neural weight updates were subtracted rather than added, or the reciprocal 1/**z** is used in (eq. 4), causing the error to fall as examples appear farther away. All three methods performed well; the results of the third are below (Figure 7).

Plotting the output of the trained neural network reveals how the agent sees the world. Each of the examples is shown as a single dot in the vertical axis corresponding to the value of the network's single node output (Figure 7, left). This agent was set in the culture of the straight row builders, as differentiated from the random aggregators, and as such has learned to see most of the first 450 examples along the horizontal axis (the row units) as 0, and most of the others (the random aggregations, to the right) as far away (note the extreme scale of the output axis). If this agent's neural network is used to place an aggregation of open and closed cells according to the algorithm described (Sec. 3.2), the result is one that very closely resembles that of the original rows in global arrangement (Figure 7, right).

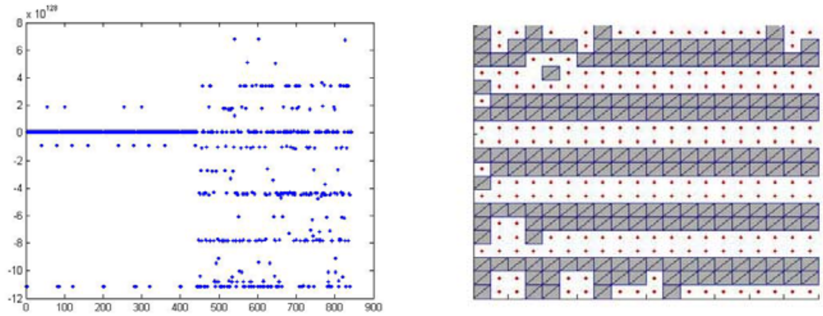

Figure 7. The agent's perceptual space (left), and its resulting building pattern.

Not every unit is placed perfectly, and not every example is seen by the network as clearly in its group, but every agent is different. In fact this difference will be crucial to the overall swarm dynamic when groups of agents interact. Various structures of network can be used representing different artificial minds, different numbers of internal connections or methods of measuring error, even learning algorithms that do not rely on neural networks. One of the main assertions of this paper is that differently constructed minds can share in the same overall process of social creativity. They are connected in the next space.

4.4 SOCIAL SPACE

The group of other agents with which one is associated, and which determines examples in the domain, constitutes the *field*. It is not a single, constant group of judges, but a fluid, changing collection of individuals selected based on the similarities of their perceptual mappings. To measure these connections between individuals the model proposes a third, *social space*, actually a simple one dimensional measure of distance between agents.

Its essential function is to enable communication between the otherwise private perceptual spaces of different agents. The similarity between two perceptual spaces can be measured by the degree of correspondence between how each sees the world. The perceptual spaces display a measure of distance between what the agent judges to be the current mean of its culture (0) and any given sample. After normalising these to have an identical variance of 1, the distance between any two agents' perceptions can be measured based on the mean of squared differences between each of the example points:

$$D(a_1, a_2) = \Sigma_{i = 1 : numExamples} [p_{a1}(ex_i) - p_{a2}(ex_i)]^2 / numExamples \qquad (5)$$

All learning algorithms that result in a perceptual mapping can be compared in this way, regardless of their internal workings. To illustrate, Figure 8 shows the result of several very different learning algorithms exposed to the same set of examples. Although each may differ in the details, each individual shares with the others the ability to perceive examples in the world. At the top is a neural network similar to the one in Figure 7, except that only a fixed number of examples closest to the mean are used in training. Below this, a different technique is used to train the network: errors from both

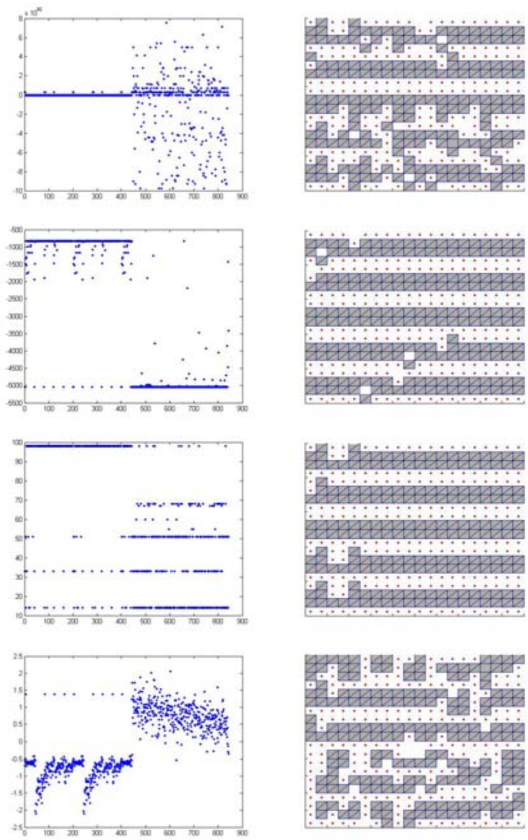

Figure 8. Agents trained with completely different algorithms (two three-layer neural networks, one Kohonen network and one SVM) have different perceptual spaces (left) but can make similar evaluations and similar constructions (right).

TABLE 1. Social space distances measured between the above perceptual spaces.

	neural net 1	neural net 2	Kohonen	SVM
neural net 1	-	1.47	1.54	1.62
neural net 2	1.47	-	0.44	1.95
Kohonen	1.54	0.44	-	1.90
SVM	1.62	1.95	1.90	-

groups are measured from the mean, but rather than adding the weight updates at each step for the examples of the random aggregation, they are subtracted. The last two examples are different algorithms entirely: a Kohonen self-organising feature map, and a support vector machine. All of these different algorithms, when trained on the same examples, result in different outputs in their individual perceptual spaces (which may even have different numbers of dimensions) but each is alike in that the resulting perceptual framework allows the individual to make similar decisions about examples regarding their distance from the mean.

The group of other agents that each considers to be its *field* is determined by the distances measured as in (eq. 5), each selecting the agents perceived to be the closest. Table 1 shows the distances between the perceptual spaces of the four agents in Figure 8. Based on these distances, the second neural network agent and the Kohonen agent would each select the other as a member of their fields, being the closest of the possible choices. The similarity is also evident in the visual appearance of their construction outputs (Figure 8, right).

Kuhn (1962) emphasizes the role of a field's lexicon in both facilitating internal communication, and isolating it from outsiders. As the boundaries of a field are defined by shared examples that define it, we can actually speak of the field itself as having a perceptual framework of its own: the perceptual space that would be defined by those examples (Figure 9). Suppose we take a set of eight examples in the world. If you and I both select a set of six or seven as representative of our individual ideals we might each have one or two unique samples, but there is a general overlap in the remaining 5 which will define our shared field, and its agreed perceptual space. Also, because the field ignores examples outside those chosen, its distance to the perceptual spaces of each individual could be zero, even if these individuals would differ in mapping examples out of the field. This field's perceptual space therefore coincides to a fairly high degree with the perceptual spaces of each individual in the field, providing the common ground that makes communication and mutual understanding possible.

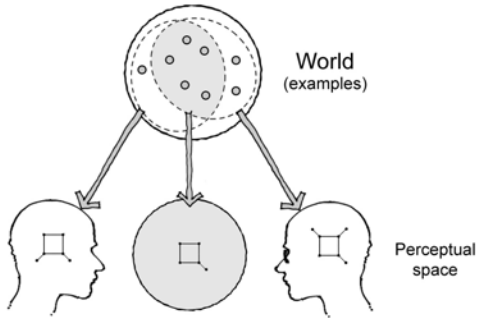

Figure 9. The field's shared perceptual space is defined by shared examples.

5 Completing the loop: putting agents together

At the macroscopic level, this model is concerned with both the changes in design over time, and the behaviour of large communities. The behaviour of a group of agents would therefore be expected to exhibit certain characteristics associated with creative social systems: innovation, local coalescence toward accepted ideals, and clique formation.

Innovation has two requirements, the most obvious being the generation of novelty in the world space of possibilities, a prerequisite for both P and H-creativity (Boden 1993). Over long spans of time, this inevitably results in a slow shift in cultural norms, fashions or styles, as an accumulation of many small innovations. If one imagines a theoretical space of 'all possible designs', (illustrated schematically in Figure 10) then the artefacts thus far produced in the history of humanity would fill only a small cloud, leaving vast expanses of the space still unexplored. But that cloud is always expanding around its periphery, and over time the system should expand in the world space of possibilities, to explore it broadly.

The second requirement of creative innovation is that the results are not just new, but also unexpected or surprising (Boden 1993, p.30). Many swarm algorithms (Dorigo 1997, Kennedy and Eberhart 2001) and cultural simulations (Axelrod 1997, Saunders and Gero 2001, Sosa and Gero 2002) incorporate randomness to generate novelty, whereas others (Reynolds 1987, Wolfram 1994) produce a complex or chaotic overall behaviour from the interactions of deterministic agents. Because this model is proposing an emergent creativity in the interactions *between* groups rather than explicitly novelty-seeking agents, it follows the second approach. Although no stochastic algorithms are used for sampling or training, viewed over time the system should display apparent randomness and unpredictability, both in the

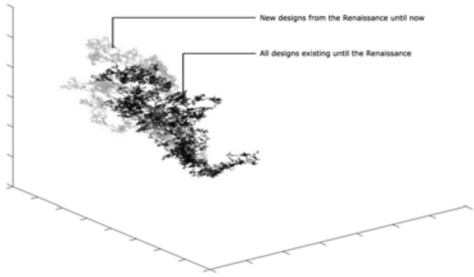

Figure 10. Examples in the space of all possible designs.

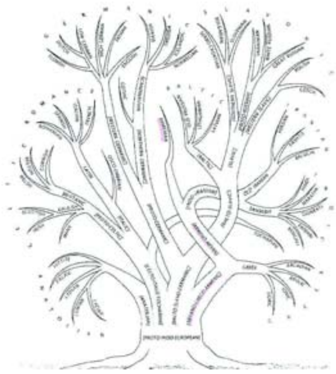

Figure 11. Indo-European languages. Image © 1990 Gamkrelidze and Ivanov

examples chosen (or created) by each agent, and in changes of field in the social space.

Like the polarisations seen in Axelrod's and other swarming models, subgroups in a social system will be expected to coalesce toward some local consensus. This is the basic behaviour of swarm models. This swarming tendency among people is necessary because we are social – there is a biological need to coalesce and an intellectual need to understand one another.

Finally, cliquing tendencies also split us up, often causing the bitterest of disagreements between seemingly close individuals. Saunders and Gero (2001) demonstrate the influence of groups of artificially creative agents on one another in the creation of fractal art, showing that agents with similar desires for novelty tend to form cliques. The effect is also revealed in genetic evidence: Bodmer and Cavalli-Sforza (1976) note that 93% of genetic differences occur within races, and the 7% genetic differences between them are weighted toward visible characteristics. These visible differences are explained by a sexual selection that penalizes traits identified in one's neighbours outside the group, thereby accentuating the visible differences of adjacent groups and increasing the cultural divide.

This divergent behaviour the creative social system should display is often pictured as a branching tree, such as that of languages (Figure 11) or the increasing specialisation of scientific disciplines from a common trunk of enlightenment natural philosophy, but there is also the possibility of merging branches. English, for instance, has had major contributions from neighbours on the Latin, Britannic and Northern Germanic branches as well as influence from many others, and there are the so called interdisciplinary fields like biochemistry, which combine major branches of science. Rather than the clearly defined branches of the diagram then, the whole is a diverging and converging collection of loosely connected individuals clustered around individual foci, and this is how a system of agents is expected to behave over time.

5.1 TESTING THE SWARM

Agents interact by sharing examples in a common world, and indicating their perceptions to others with whom they are close in the social space. The interaction of a group of agents is tested in this section. Each agent cycles through the following steps, one in each of the spaces mentioned:

1. Get a new set of n_d examples from the *world* that determine the agent's domain, and a set of n_o examples outside the domain.
2. Train the neural network so that it distinguishes the chosen n_d domain examples from the other n_o examples as the agent sees them. This adjusts the *perceptual space* of the agent. It can communicate its current position to all others by selecting an example from all available affordances in the world space nearest to what it perceives is its current ideal.
3. The similarity in the *social space* determines which agents communicate with one another. An agent selects the new set of samples for step 1 from the set pointed to by this group (the field), including the other agents' newest designs in its new domain.

The above steps accomplish the mapping between the three spaces as described in section 4: examples in the world are mapped to a lower-dimensional perceptual space, and distances between perceptual spaces are mapped to a single dimension in the social space. The cycle reiterates as proximity in the social space determines, or points to, high-dimensional examples in the world. This repeated process causes agents to constantly adjust their perceptual framework to accommodate new examples in the domain. The results of the model test indicate this motivates an overall cultural change, as represented by variation in the building output of a particular field of agents.

5.1.1 System behaviour: innovation and cliques

Figure 12 shows the result of a group's interaction over time according to the above rules. It plots an arbitrary (one-dimensional) projection of the domain means for each agent in their shared world on the vertical axis against time on the horizontal. The details of this appear quite different depending on the axis of projection chosen (just as they would appear different again in each agent's perceptual space), but the overall characteristics are the same. There is a gradual expansion from a common start, as agents' work explores the space of options.

One can see the same branching into cliques as occurs in Axelrod's model, and the trees of languages or disciplines (Figure 11), but in fact each is made up of many units produced by several agents. There is individual movement between them, and no strict definition of membership, but after 30 cycles two major groups appear, typified by agents 3 and 4. The building patterns produced by each of these agents in isolation at cycle 30 are shown in Figure 13. All agents in the clique represented by agent 3 display a tendency to build in a similar (but not identical) radial network, and those in the other clique in rough horizontal rows. An examination of the *fields* chosen by agents 3 and 4 also reveals that they are mutually exclusive, sharing no agents in common.

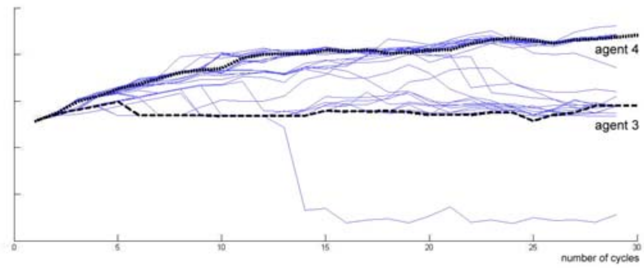

Figure 12. A projection of the agents' ideal means in the world space over time.

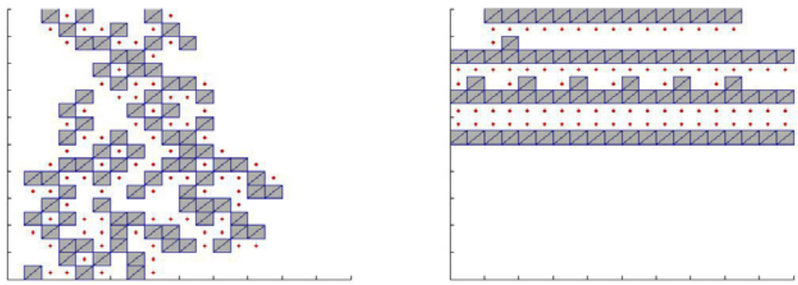

Figure 13. Examples of building patterns from two different cliques: agents 3 (left) and 4 (right).

Cultural change has occurred, in that the preferences dictated by the changing perceptual spaces of the agents cause them to build differently over time. At cycle 1 the building pattern of every agent is identical, but at cycle 30 neither of these building patterns is exactly like the others, or like the initial starting point.

5.1.2 The effect of different points of view

The hypothesis that different perceptual spaces, different ways of seeing the world, are responsible for the group's creativity was tested in the model in three different runs from the same initial domain. A simpler model was used, with only five dimensions in the world space and the agents' building affordances unrestricted (i.e. their output can be any point in this five-dimensional space), to look purely at the dynamics of the group over time.

Figure 14(a) shows the result of all agents locked to identical perceptual spaces. There is no difference to the way they see the world, and their ability as a society to explore a broad region of the world space is extremely low. (The ticks on the vertical axis in this case are actually single units rather than hundreds.) All agents have the same internal complexity as one another, and

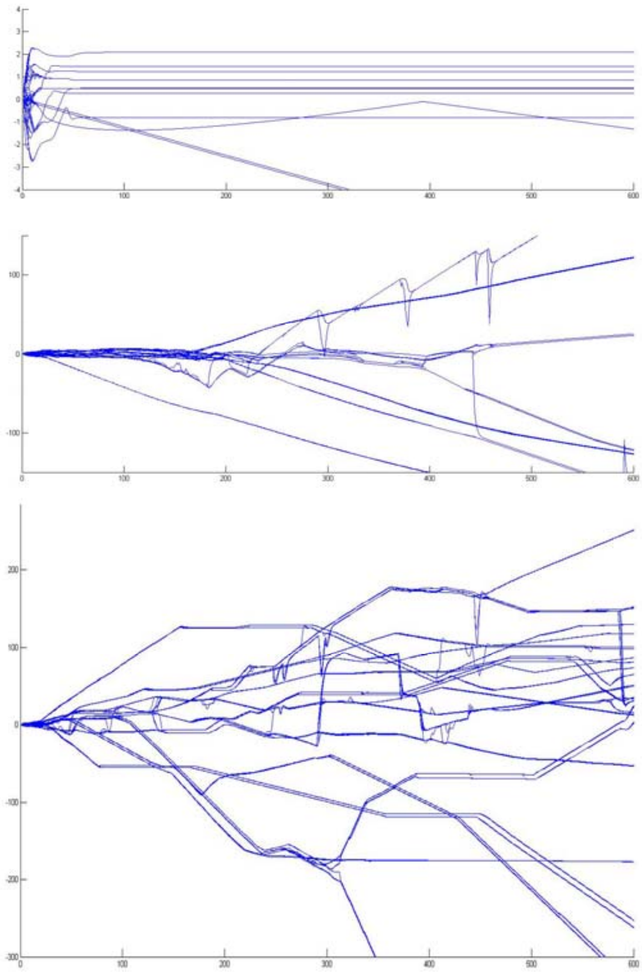

Figure 14. Agents' ideals in the real world (vertical) over time (horizontal):
a) (top) Identical perceptual framework in all agents,
b) (centre) Each agent with different perceptual frameworks,
c) (bottom) New agents added over time.

those in the other runs. The only difference in Figure 14(b) is the fact that the agents view the world differently from one another, but this produces both overall change and innovation of the group, and far more internal variation and complexity. The individual lines on the plots were found at any given time to contain between one and ten agents. Some can cross, particularly in the dense first 300 cycles. In doing so there is both a branching and merging of a loose collection of individuals. Even when these seem to be in clear fields an apparently interdisciplinary individual can be seen to cut across several branches by being near examples of both in its own perceptual space.

In Figure 14(c), the complexity is seen to further increase with the addition of new agents to the culture over time. Every ten time steps a new agent designer was added to the population. To avoid direct introduction of novelty, it was specifically placed in an existing field by being given a perceptual space and a domain of examples identical to an existing agent. Even with this initial similarity, there is a far greater branching and innovation of work than seen in the previous run. Fixed parameters of n_d and n_o ensure that a given field will eventually saturate when the number of examples exceeds their capacity to view, causing a split. This process occurs continuously, then, as long as the population increases.

5.2 THE INTERDISCIPLINARY INDIVIDUAL

An agent seen moving across fields only appears so from outside its own perceptual framework. It does not follow a middle line half-way between the two branches, because in some dimensions these two are very closely aligned. Such agents are found to include examples of both in a single domain, although they may look distinct from other agents' points of view.

This suggests that what we tend to call interdisciplinary work is often this same occurrence. Different disciplines have their own unique languages and customs that allow for ease of communication and rapid exchange within the community, but misunderstandings and general lack of exchange between groups. The more a discipline becomes 'specialised', the greater is its ability to make progress (movement in the swarm) on the problem at hand, but less its ability to communicate to outsiders. Interdisciplinary work results when individuals from different disciplines approach one another in dimensions that are not part of the general perceptual framework of either group. A multidisciplinary individual can bridge between the two groups not because of a *greater* breadth of understanding, but because of a *different* perceptual framework that includes the dimensions in which the two groups produce work of similar features. The work of a particular artist and a particular biologist can seem highly relevant to one another from this new point of view. Interdisciplinary work can itself sometimes develop into a new discipline as more individuals adopt similar views and form a new field.

5.3 SUBSTITUTES FOR RANDOMNESS, AND SOURCES OF INNOVATION

The behaviour of individuals within the creative system modelled appears unpredictable over time, but is not the result of a stochastic process. The outcome is completely deterministic given the initial conditions of domains and perceptual spaces. Instead, the apparent randomness appears to arise from two different processes.

The first is the act of design as a process of selection from affordances in the world. The details of the work are input to the agent, rather than output, and a constantly changing world (due to new building) affords different possibilities at each given point in time. Thus even an agent that maintains a constant state in terms of domain and perceptual space would produce different output at times t and t+1, and appear to be making random choices.

Even when the possibilities in the world are unrestricted however, as in section 5.1.2, there is still apparently random behaviour. The other source appears to be the cycle of mapping between the three different spaces of differing dimensionality. This introduces into the overall system a complex non-linearity like that seen in Reynolds' (1987) flocks or Wolfram's (1994) class 3 and 4 cellular automata.

6 Conclusion

In this work a computational model for creative design among agents was presented, in which domain and field change over time. These changes, which appear in real societies, seem to be an important part of the process. The simulation yielded global cultural innovation that occurs within several different abstract spaces. Their structure provides several advantages over a typical swarm model:

- The behaviours of attraction and repulsion that lead to swarm dynamics can arise out of the simple action of an agent classifying examples in the world with a learning algorithm.
- While agents in swarm algorithms are normally identical, differently constructed agents with different ways of learning can also swarm together, even while each exists in a different perceptual space. The point of view, sensory experience and even the dimensionality of perception may be different for every agent, but they can still form a common social dynamic by taking advantage of a shared world in which they act by selection of affordances.
- Even though the perceptual space of each individual is limited, the space in which they are able to swarm as a group is far greater in dimensionality. In this way it seems possible that a finite mind may explore any of the possibilities of an infinite-dimensional reality.

The model has been kept simple for clarity, and to show that individual innovation can result from such a structure. In particular, agents' behaviour has not been motivated by goals, either for utility or novelty for its own sake. Nonetheless, the model does produce the kind of behaviour expected of a swarm, and a seemingly random exploration through a space of possibilities, so it is likely that goals, in the form of a fitness evaluation, could be incorporated. Unlike stochastic particle swarm or optimisation algorithms however, random input would not be required.

Unusual for a model of creativity, this work proposes that the primary role of the mind is to take in, and make sense of the world, rather than stressing innovation for its own sake. But in making sense of what it takes in, it necessarily develops a new point of view that produces the innovation naturally. When two agents share the same examples in their domain, their perceptual frameworks become more aligned to one another by repeated exposure to those examples, but on examples outside the domain they may always disagree, as colleagues in work may disagree on politics or art. These variations in other dimensions continue to motivate the group, while the role of the group in this respect is to pull the individual along.

So where does individual creativity happen? Where does group creativity happen? These may not be separate questions. It may also be the case that creativity at a global level does not always come from an intentional effort on the part of the individual, but that novelty is sometimes a product of the interaction between the different parts of a creative system. Thus it is an emergent phenomenon that can happen at any level, among groups, neurons, agents or us.

Acknowledgements

I would like to thank Alan Penn for spatially related advice, and Peter Bentley for insights while developing the model presented in this paper.

References

Adrian E D: 1928, *The Basis of Sensation*. Christophers, London.

Axelrod R: 1997, The dissemination of culture: a model with local convergence and global polarization. *Journal of conflict resolution*, Vol. 41, No. 2. pp. 203-206.

Boden MA: 1990, *The creative mind: myths & mechanisms*. Weidenfeld & Nicolson,London.

Bodmer W and Cavalli-Sforza L L: 1976, *Genetics, Evolution and Man*. W H Freeman, San Francisco.

Brockman JB and Director SW: 1991, The Hercules Task Management System, *in Procedings of the International Conference on Computer-Aided design.*

Burgess C: 1998, From simple associations to the building blocks of language: Modelling meaning in memory with the HAL model. *Behavior Research Methods, Instruments, and Computers,* 30, 188-198.

Czikszentmihalyi, M: 1988, Society, culture, and person: a systems view of creativity, *in* Sternberg, RJ (ed), *The nature of creativity: Contemporary psychological perspectives*, Cambridge University Press, Cambridge, pp. 325-339.

Dorigo M and Gambardella L M: 1997, Ant colony system: A cooperative learning approach to the traveling salesman problem. IEEE Transactions on Evolutionary Computation, 1 (1), 53-66.

Dourish P: 2001, Where the *Action Is - The Foundations of Embodied Interaction*, MIT Press, Cambridge, MA.

Gamkrelidze TV and Ivanov VV: 1990, The Early History of Indo-European Languages. *Scientific American*, March 1990, P.110.

Gero JS: 1993, Towards a model of exploration in design, *in* Gero JS and Sudweeks F (eds) *Preprints Formal Design Methods for CAD,* Key Centre of Design Computing, University of Sydney, Sydney, IFIP, pp. 271-292.

Gibson J J: 1979, The Ecological Approach to Visual Perception. Boston: Houghton Mifflin.

Hillier B and Hanson J: 1984*, The Social Logic of Space*. Cambridge University Press.

Hutchins E: 1995, Cognition in the Wild. MIT Press. Cambridge, MA.

Kennedy J and Eberhart RC: 2001, *Swarm Intelligence*. Morgan Kaufmann, San Francisco.

Kuhn, Thomas S: 1962, *The Structure of Scientific Revolutions*, 3rd edition, University of Chicago Press.

Maher ML & Poon J: 1996, Modelling design exploration as co-evolution, Microcomputers in Civil Engineering (Special Issues on Evolutionary Systems in Design)

Maturana HR and Varela FJ: 1987, *The Tree of Knowledge: the biological roots of human understanding*. Shambhala Publications Inc. Boston.

Nehaniv CL and Dautenhahn K: 1999, Of hummingbirds and helicopters: An algebraic framework for interdisciplinary studies of imitation and its applications. *In Learning Robots: An Interdisciplinary Approach*, J. Demiris and A. Birk, eds., World Scientific Press, in press.

Norman DA: 1988, The *psychology of everyday things*. New York: Basic Books.

Quick T, Dautenhahn K, Nehaniv CL and Roberts G: 1999, The essence of embodiment: a framework for understanding structural coupling between system and environment, *Proc Third International Conference on Computing Anticipatory Systems*, (CASYS, '99).

Reynolds CW: 1987, Flocks, herds and schools: A distributed behavioural model. *Computer Graphics*, 21, 25-34.

Rosenman, M: 1997, The Generation of Form Using an Evolutionary Approach, *in* Dasgupta, D and Michalewicz, Z (eds), *Evolutionary Algorithms in Engineering Applications*, Springer-Verlag, pp. 69-86.

Saunders R and Gero JS: 2001, Artificial creativity: A synthetic approach to the study of creative behaviour, in JS Gero and ML Maher (eds), *Computational and Cognitive Models of Creative Design V*, Key Centre of Design Computing and Cognition, University of Sydney, Sydney, pp. 113-139.

Sosa R and Gero JS: 2002, Computational models of creative situations: Towards multiagent modelling of creativity and innovation in design, in Gero JS and Brazier FMT (eds) *Agents in Design 2002*, Key Centre of Design Computing and Cognition, University of Sydney, Sydney, pp. 165-180.

Steels L: 2000, The Emergence of Grammar in Communicating Autonomous Robotic Agents, in Horn W (ed), *ECAI2000*. IOS Press, Amsterdam. pp 764-769.

Tang H-H and Gero J: 2001, Sketches as affordances of meanings in the design process. In J. S. Gero, B. Tversky and T. Purcell (eds), 2001, Visual and Spatial Reasoning in Design II, Key Centre of Design Computing and Cognition, University of Sydney, Australia, pp. 271-282.

Wolfram S: 1994, *Cellular Automata and Complexity: Collected papers*, Addison-Wesley, Reading MA.