A CURIOUS DESIGN AGENT

A Computational Model of Novelty-Seeking Behaviour in Design

ROB SAUNDERS, JOHN S. GERO
Key Centre of Design Computing and Cognition
Department of Architectural and Design Science
The University of Sydney 2006 Australia
{rob, john}@arch.usyd.edu.au

Abstract. This paper presents a “curious design agent”, i.e. an agent that uses the search for novel designs to guide its design actions. A computational model of curiosity based on a process called novelty detection is presented. The behaviour of the computational model is illustrated with a curious design agent searching the space of two-dimensional patterns generated by a simulated Spirograph is reported.

1. A Computational Model of Curiosity

Curiosity is the motivation to discover new knowledge when faced with an unfamiliar situation (Berlyne, 1971). It promotes self-directed learning by rewarding behaviour that results in the assimilation of new knowledge. Curiosity can be used to guide the search and exploration of unfamiliar design spaces to find new knowledge with the goal of gaining a better understanding of a non-routine design task. Curious search can be used to guide problem solving to find interesting design solutions. Curious exploration can be used to guide problem finding to discover interesting design problems. In this paper we concentrate on the role of curiosity in the search for interesting design solutions.

The curious design agent presented here uses a computational process called *novelty detection* to determine its level of interest in designs. The curious design agent uses the detection of novelty to guide its search during the course of a design session. Determining interestingness based on novelty depends upon the knowledge of the agent and its computational abilities; things are *boring* if either too much or too little is known about them (Schmidhuber, 1997).

A neural network called a self-organising map or SOM (Kohonen, 1995) is used in the following experiment to implement a simple novelty detector. SOMs project vectors in the high dimensional space onto a hyperplane, typically two-dimensional. The projection provides an ordered map of the high dimensional space with similar vectors clustered together into neighbourhoods.

The simple novelty detectors implemented for this experiment use a SOM to provide an adaptive representation of the design space with respect to the experiences of the agent. The SOM projects high dimensional vectors representing images onto a two-dimensional map of category prototype vectors represented by neurons. We can think of this projection as the "conceptual design space" of the agent, representing the forms that the agent has experienced often enough to learn.

The novelty detector uses the learned representation of the design space to recognise novel designs, i.e. designs that are not well represented by the categories of the SOM. For a given input the novelty detector uses the activation of the best matching neuron in the SOM to determine the typicality of the input for that neuron's assigned category. Novelty is calculated as the complement of the typicality measure, i.e. the less typical the more novel it is.

Interest in a given input is calculated by applying a sigmoid transform to the calculated novelty measure. The sigmoid transform is used to approximate the cumulative Gaussian curve observed in studies of human arousal (Berlyne, 1971). This transformation provides a more definitive judgement of interest than the linear scale of novelty.

The calculated level of interest is used to model curiosity by favouring actions that promote the learning of new representations of the design space. This is simply achieved by attenuating the magnitude of design mutations in proportion to how interesting the design is judged to be.

1.1. SEARCHING FOR NOVEL SPIROGRAPH PATTERNS

To illustrate the behaviour of a curious design agent we have implemented a computational model of a prototypical design generator from childhood: the Spirograph[1]. Spirograph sets consist of an array of plastic gears. To draw a pattern one gear is fixed to a piece of paper and a second gear is moved around it while tracing its path by pushing a pen through a hole in the interior. This simple toy has charmed children for over 30 years and has no doubt sparked an interest for geometrical patterns in budding architects and designers.

---

[1] Spirograph is a registered trademark of Hasbro.

The goal of this experiment is to examine the behaviour of a curious design agent as it searches the space of patterns that can be generated using the simulated Spirograph. A simple arrangement of circular gears can be mathematically modelled using the following equations:

$$x = (r_1+r_2) \times \cos\theta_1 - r_2 \times \cos\theta_2 \qquad (1)$$
$$y = (r_1+r_2) \times \sin\theta_1 - r_2 \times \sin\theta_2$$

where: r_1 = radius of fixed gear
r_2 = radius of moving gear
θ_1 = rotation of moving gear around fixed gear
θ_2 = rotation of moving gear

Figure 1 illustrates the variety of Spirograph patterns that is possible for a small selection of random values for r_1 and r_2. Visually, two broad categories of Spirograph patterns can be readily distinguished: simple patterns produced with a few rotations of the moving gear around the fixed gear and complex patterns produced with many rotations. The number of rotations is dictated by the greatest common denominator of the two radii. Figure 1 shows that for a random sample of the design space complex patterns are far more common.

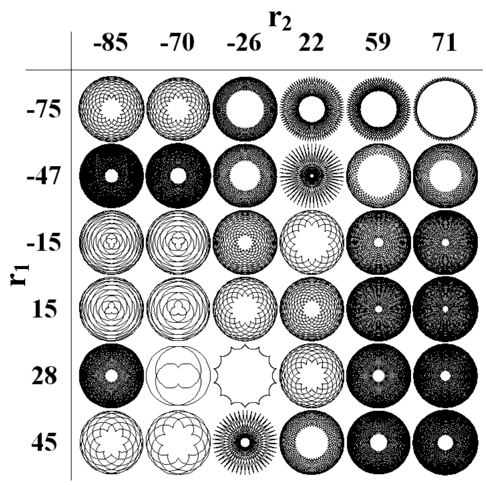

Figure 1. A random sample of Spirograph patterns with a small selection of random values for the fixed gear radius (r_1) and the moving gear radius (r_2).

We set a curious design agent to search a sub-space of possible Spirograph patterns bounded by gear ratios r_1:r_2 from -100:1 to 100:1. The agent explored the space of possible Spirograph patterns by changing the value of the ratio directly, rather than the values of the gear radii, as this provides a more predictable space of patterns because similar gear radii can produce wildly different patterns, whereas similar ratios generally produce similar patterns. The patterns were analysed by the agent as 32x32 pixel greyscale images. At this resolution many of the finer details of the patterns are lost and the classification of the patterns is based on larger scale features.

1.1.1. Results

The novelty detected by the agent during a search of the design space for the first 200 time steps is shown in Figure 2. The chart shows that the curious design agent performs an initial search of the design space, up to time step 54, when little novelty is detected. The agent then repeatedly finds interesting patterns, indicated by peaks in novelty and will remain with a novel pattern to learn a category for it, producing a tailing-off in the novelty detected as the category is constructed.

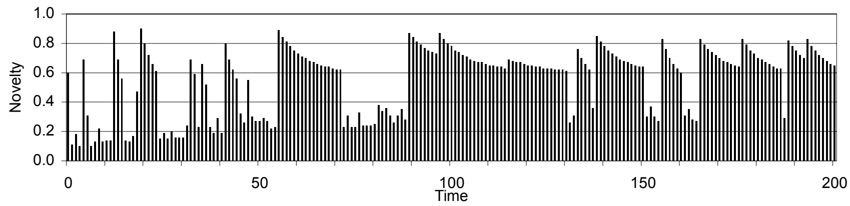

Figure 2. A chart showing the novelty detected over time for the first 200 time steps of a typical run. The chart shows an initial period where little novelty is detected followed by a series of peaks as novel patterns are discovered. The chart also shows the tailing-off of the novelty detected after each peak as the novel pattern's category is learned.

SOMs are often used to visualise complex multi-dimensional vector spaces and we have used the SOMs constructed by a curious and a non-curious agent to visualise the "conceptual design spaces" constructed by these agents. The non-curious agent trained its SOM on a random sample of patterns. The curious agent trained its SOM by generating patterns as described above. Each agent explored the design space for 400 time steps.

Figure 3 shows the representations learned by the SOMs of the two agents. Each grid shows the two-dimensional map of the design space that has been learned. Each cell corresponds to a neuron in the lattice of the SOM and shows the prototypical image of the category it represents. The prototype is an average of the images of the patterns that are contained within the category.

Although the networks have mapped the design space differently, some correspondences can be found, e.g. the categories found in the bottom half of the non-curious agent's map (D1–F6) roughly correspond to those found in the top-left corner of the curious agent's map (A1–D4).

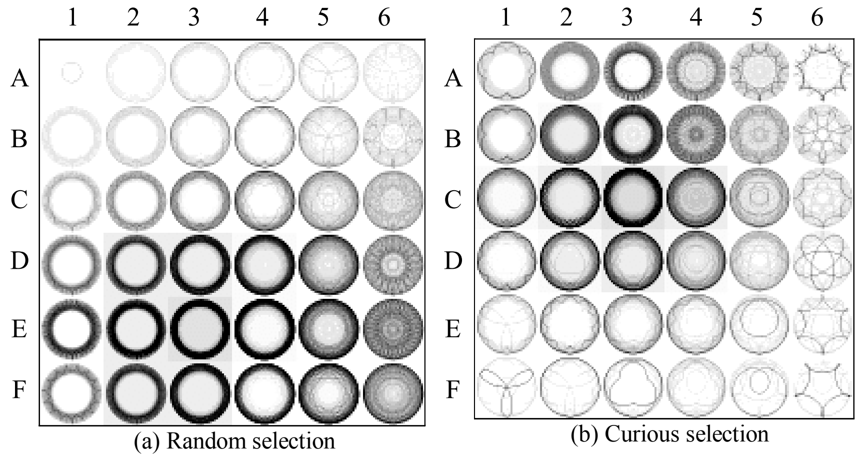

Figure 3. The above maps show the bitmap images of the prototype patterns represented by the neurons of two 6x6 SOMs trained with (a) a random selection of patterns chosen by a non-curious agent, and (b) a set of 'interesting' patterns chosen by a curious design agent.

1.1.2. Analysis

Figure 2 shows that little novelty is detected in the first 50 time steps of searching. During this phase the lack of novelty means that mutations of the design parameters, i.e. the ratio of gears, will be high. Consequently, the agent begins by learning from a fairly random sample of patterns and as Figure 1 illustrates the agent is consequently exposed to far more complex patterns than simple ones. The result is that the agent learns a set of categories for typical, complex patterns first.

Once an initial set of pattern categories has been learned, novel patterns, i.e. simple patterns, can be recognized. Figure 2 shows that the agent spends most of its time beyond the time step 55 learning novel patterns in some cases spending as long as 10 time steps learning a single pattern, indicated by the slow decay in the novelty detected.

The two phases of learning are reflected in the differences between the SOMs produced by the curious and non-curious agents shown in Figure 3. The maps share similar representations for typical patterns, e.g. compare the prototypes E1, E5, F5 in the non-curious agent's map with A3, C4, D3 in the curious agent's map respectively. However, the maps differ considerably in their representation of novel patterns. For example, the non-curious agent's map has only a few categories in rows A and B that are similar to those found in rows E and F and columns 5 and 6 of the curious agent's map.

In fact, the curious agent's map devotes nearly 50% of its categories to different types of novel patterns whereas only ~15% of the categories in the non-curious agent's map represent unique novel patterns. This marked difference between the maps reflects the prolonged learning of novel patterns as a consequence the actions of the curious agent. As a result of this improved map of some initially novel patterns the curious agent will no longer find the simple patterns represented to be novel and will have to search for new patterns in order to maintain interest in the design space.

The neural networks used in the above experiment are relatively small compared to the design space, and so are never likely to accurately represent all of the possible Spirograph patterns. However, larger neural networks and more powerful learning systems may have the capacity to learn the space of Spirograph patterns so well that they can become “bored” with the entire space, triggering the exploration of new design spaces.

2. Discussion

The results from the above experiment, and similar experiments in other domains (Gero and Saunders, 2000; Saunders and Gero, to appear), suggest that providing design agents with a sense of curiosity confers significant advantages in the search of ill-defined design spaces. Computational models of curiosity provide general-purpose, knowledge-lean heuristics to guide the search for potentially interesting, and possibly even creative, designs.

Future work will need to develop the role of curious design agents in the user interface of design tools such as CAAD systems. The benefit for future CAAD systems that incorporate curious design agents is that these agents will be able to assist designers exploring unfamiliar design spaces. Computational models of curiosity promise to provide agents that can usefully engage in collaborative, non-routine design by reducing the set of generated designs requiring a designer’s attention to those that are potentially interesting.

References

Berlyne, D. E.: 1971, *Aesthetics and Psychobiology*, Appleton-Century-Crofts, New York.

Gero, J. S. and Saunders, R.: 2000, Constructed representations and their functions in computational models of designing, *in* Tang, Tan and Wong (eds), *CAADRIA 2000*, CASA, Singapore, pp. 215–224.

Kohonen, T.: 1995, *Self-Organizing Maps*, Springer–Verlag, Berlin.

Saunders, R. and Gero, J. S.: to appear, Designing for interest and novelty: Motivating design agents, *CAAD Futures 2001*, Eindhoven.

Schmidhuber, J.: 1997, What’s interesting? *Technical Report TR-35-97*, IDSIA, Lugano, Switzerland.