Non-Symbolic AI lecture 6

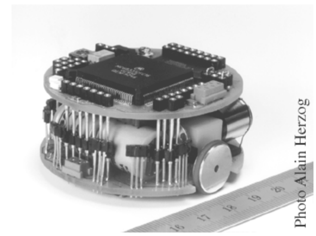
Last lecture we looked at non-symbolic approaches to robotics – new methods of approaching **Engineering** problems

This lecture we look at simulated agents as a Non-symbolic AI (or Alife) way of asking and answering scientific issues.

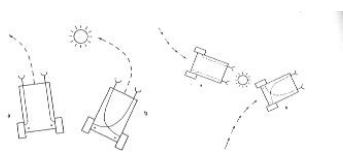
Braitenberg vehicles

Braitenberg vehicles have sensors (typically few, and simple) and motors (typically 2, driving left and right wheels independently).

A bit like Khepera robots



Braitenberg (2)



Then very simple direct connections between sensors and motors can, for example, produce light-seeking or light-avoiding behaviour.

Above examples may have direct linear connections, plus some bias term on the motors so as to move even when no sensory input.

More complex Braitenberg vehicles

These are feedforward ANNs, feeding from sensory inputs to motor outputs. They can be made more complex (while remaining feedforward) by

- Adding bias terms
- Adding internal nodes – multilayer ANN
- Adding non-linearities (… threshold functions, sigmoids)
- Adding extra sensors for new senses
- … etc etc

Reactive Behaviour

But as long as a Braitenberg vehicle has a feedforward ANN (however many sensors, however many layers, however non-linear) from sensors to motors, then it is merely a **Reactive System.**

Warning: different people use the term 'reactive' in different ways, here I am explicitly using it to mean 'no internal state'. If the robot reacts to the same inputs the same way on Mon, Tues, Wed …, then it is reactive under my definition.

A coke machine that requires 2 coins to deliver a can is *not* reactive – since reactions to inserting a coin differ.

Non-reactive behaviour

Subsumption architecture produces non-reactive behaviour (remember the 'alarm clocks' in AFSMs)

DRNNs produce non-reactive behaviour.

All really complex behaviour is non-reactive (in this sense) – even though Braitenberg vehicles demonstrate how surprisingly interesting one can get merely with reactive.

Non-reactive means that behaviour changes according to previous experience – which brings up notions such as **memory** and **learning**.

What is Learning ?

EASy

Learning is a **behaviour** of an organism – animal or human or robot, or piece of software behaving like these.

More exactly, it is a **change of behaviour** over time, so as to improve performance at some task.

On Monday I could not ride a bicycle – my behaviour was 'falling-off-the-bike'.

By Friday my behaviour had changed to 'successfully riding the bike'.

Strictly, ANN weight changes are **plasticity**, not learning – though the whole system may learn via this plasticity.

Evolution and Learning

EASy

Exploring Adaptive Agency II: Simulating the Evolution of Associative Learning
Peter Todd and Geoffrey Miller, pp. 306-315 in
From Animals to Animats, J-A Meyer & S. Wilson (eds)
MIT Press 1991 (Proc of SAB90)

Looking at some aspects of the relationship between *evolution* and *learning*.

Why bother to Learn - 1?

EASy

Todd & Miller suggest 2 reasons:

(1) Learning is a cheap way of getting complex behaviours, which would be rather expensive if 'hard-wired' by evolution.
Eg: parental imprinting in birds –

'the first large moving thing you see is mum, learn to recognise her'.

Why bother to Learn – 2?

EASy

(2) Learning can track environmental changes faster than evolution.

Eg: it might take millennia for humans to evolve so as to speak English at birth, and would require the English language not to change too much.

But we have evolved to learn whatever language we are exposed to as children -- hence have no problems keeping up with language changes.

Learning needs Feedback

EASy

Several kinds of feedback:
Supervised -- teacher tells you exactly what you should have done
Unsupervised -- you just get told good/bad, but not what was wrong or how to improve.
(..cold...warm...warmer...)

Evolution roughly equates to unsupervised learning
-- if a creature dies early then this is negative feedback as far as its chances of 'passing on its genes' are concerned -- but evolution doesn't 'suggest what it *should* have done'.

Evolving your own feedback?

EASy

However, it may be possible, under some circumstances, for evolution to create, within 'one part of an organism', some subsystem that can act as a 'supervisor' for another subsystem.

Cf.DH Ackley & ML Littman.
Interactions between learning and evolution. In *Artificial Life II*, Langton et al (eds), Addison Wesley 1991.

Later work by same authors on
Evolutionary Reinforcement Learning

EASy

The Todd & Miller model

Creatures come across food (+10 pts) and poison (-10)

Food and poison always have different *smells,* sweet and sour. **BUT** sometimes smells drift around, and cannot be reliably distinguished. In different worlds, the reliability of smell is x% where 50 <= x <= 100.

Food and poison always have different *colours,* red and green. **BUT** in some worlds it is food-red poison-green, in other worlds it is food-green poison-red. The creatures' vision is always perfect, but 'they dont know whether red is safe or dangerous'.

EASy

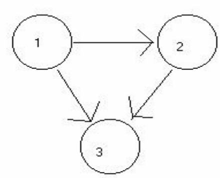
The model – ctd.

Maybe they can learn, using their unreliable smell?
Todd & Miller claim that simplest associative learning needs:
(A) an input that (unreliably) senses what is known to be good or bad (smell)
(B) another sensor such as that for colour, above
(C) output such that behaviour alters fitness.
(D) an evolved, fixed connection (A)->(C) with the appropriate weighting (+ve for good, -ve for bad)
(E) a learnable, plastic connection (B)->(C) which can be built up by association with the activation of (C)

EASy

Their 'Brains'

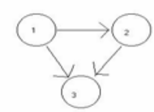
Creatures brains are genetically specified, with exactly 3 neurons connected thus:



EASy

Genetic specification

The genotype specifies for each neuron whether it is
- ✓input sweet-sensor
- ✓input sour-sensor
- ✓input red-sensor
- ✓input green-sensor
- ✓hidden unit or interneuron
- ✓output or decision unit: eat/dont-eat



- ✓and for each neuron the bias (0, 1, 2, 3) (+/-)
- ✓and for each link the weight (0, 1, 2, 3) (+/-)
- ✓and whether weight is fixed or plastic

EASy

Plastic links

Some of the links between neurons are fixed, some are plastic.

For plastic weights on a link from P to Q, Hebb rule:
Change in WEIGHTPQ = $k * A_P * A_Q$
where A_P is the current activation of neuron P.

I.e.:- if the before and after activations are the same sign (tend to be correlated), increase strength of link
If opposite sign (anti-correlated), decrease strength

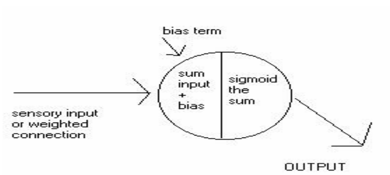
EASy

(n.b.)

(Actually, if you check the details of the Todd and Miller paper, it is a bit more complicated.

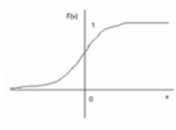
The Hebb rule assumes outputs can be positive or negative. The sigmoids used (see a couple of slides later) are always positive, but these are then rescaled to allow outputs of hidden and motor units to fall within range –1 to +1. So we are OK for the Hebb rule to make sense.)

Within each neuron



The sigmoid is a 'squashing function' such as

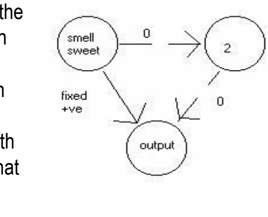
$$f(x)=\frac{1}{1+e^{-x}}$$



Possible brains

So one possible genetically designed brain would be this: **colour-blind eater** – this is **not** a learner.

Whatever neuron 2 is, the links are 0, hence it can be ignored.
This depends purely on smell, and has the connection wired up with the right +ve sign, so that it eats things that smell sweet.

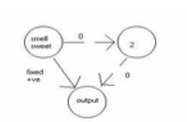


Is a colour-blind eater any good?

This creature depends purely on smell, and has the connection wired up with the right +ve sign, so that it eats things that smell sweet.

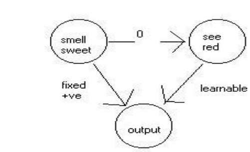
Though it may occasionally ignore food that (noisily) smells sour, and eat poison that (noisily) smells sweet, on average it will do better than random.

So evolving within a population, this (non-learning) design will do better than random, and increase in the population.



A learning brain

This different possible design is a **colour-learner**



If this one is born in a world where smell is 75% accurate, and food is red, then (more often than not) seeing red is associated with the positive output.

Is a colour-learner any good?

… seeing red is associated with the positive output. So, with Hebbs rule, the RH connection gets built up +ve more strongly, until it fires the output on its own -- it can even over-rule the smell input on the 25% of occasions when it is mistaken.

Contrariwise, if it is born in a world where food is green, then Hebbs rule will build up a strong –ve connection -- with the same results.

So over a period, this will do better than the previous one

Evolutionary runs

Populations are initialised randomly.

Of course many have no input neurons, or no output neurons, or stupid links
- these will behave stupidly or not at all.

The noise level on smell is set at some fixed value between 50% accurate (chance) and 100% accurate.

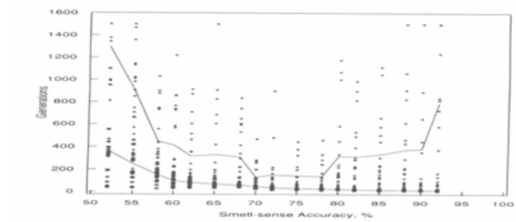
Then individuals are tested in a number of worlds where (50/50) red is food or red is poison.

Keeping statistics

Statistics are kept for how long it takes for **colour-blind eater** to turn up and take over the population; and for **colour-learner** likewise.
[Statistics: multiple runs to reduce chance elements, keep record of standard deviations]

- When smell noise-level is 50%, then there is no available information, no improvement seen.
- When smell is 100% accurate, then colour-learning is unnecessary.
- What happens between 50% and 100% ?

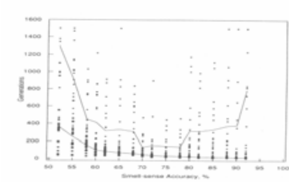
Results on a graph



Colour-learner= top line colour-blind eater=bottom line

U-shaped curve

As smell-sense accuracy goes from 50% (left) to 100% (right), average number of generations needed to evolve **colour-blind eater** decreases steadily (lower line)



But average number of generations needed to evolve **colour-learner** goes through the U-shaped curve (upper line)

Baldwin Effect

A much more subtle and interesting interaction between Learning and Evolution is the Baldwin effect -- named after Baldwin (1896).

Roughly speaking, this is an effect such that, under some circumstances, the ability of creatures to learn something guides evolution, such that in later generations their descendants can do the same job innately, without the need to learn.

Is it Lamarckian?

This sounds like Lamarckism -- "giraffes stretched their necks to reach higher trees, and the increased neck-length in the adults was directly inherited by their children".

Lamarckism of this type is almost universally considered impossible -- why ??

The Baldwin effect gives the impression of Lamarckism, without the flaws.
Be warned, this is tricky stuff !