A finite state machine robot (Brooks' Genghis)

Rodney Brooks took the idea of a finite state machine and used it to build layered levels of behavior into insect-like robots, one of the earliest being Genghis, a small six-legged robot that can walk, climb over obstacles, and follow people, all with no central program to tell it to do these things.

Genghis is now in the Smithsonian Air and Space Museum but can also be seen on the web at:

http://www.ai.mit.edu/projects/genghis/genghis.html

How Genghis works:

Genghis consists of a frame, six legs, motors to move the legs, and 51 augmented finite state machines (AFSMs). They were augmented in that the AFSMs could communicate with each other over wires. The Six legs each could move back and forth (to a position referred to by alpha) or up and down (to a position referred to as beta). The alpha motor could move the leg with values from +25 (maximum forward) to -25 (maximum backward), 0 straight out. The beta motor would make a leg stick out parallel to floor (0) with values from maximum positive (down) to maximum negative (up).

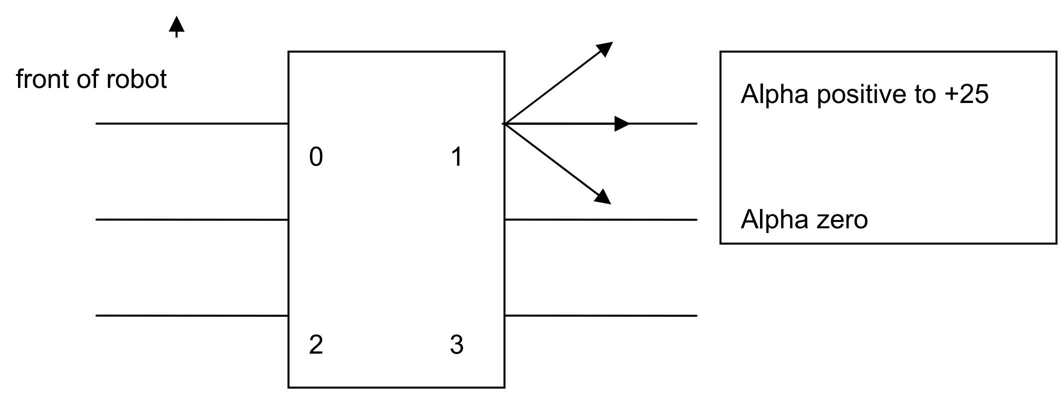

Figure 9-1 - Genghis viewed from the top

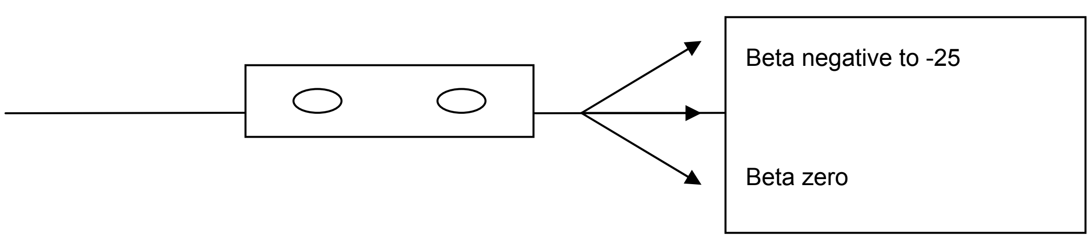

Figure 9-2 - Genghis viewed from the front

There were also four sets of sensors:

- Two whiskers in front: ON or OFF to indicate whether something was touching whisker with enough force to bend whisker.
- Force sensors on motors: range of 0 to 15 indicating how much resistance a leg was feeling to movement.
- Inclinometer: Ball bearing in a trough that detected whether the insect was tilted up in front or back, giving value from 15 (extreme up angle in front) to 0 (extreme up angle in back).
- IR sensors in front tuned to 10 micron wavelength (temperature of people)

The figure above shows the alpha front-to-back angle of a Genghis leg and the beta up down angle. Now, carefully follow the definitions of the behaviors of the finite state machines. The first set of AFSMs controlled the legs – each leg had a set of these functions.

- <u>Beta posn</u>: **input** δ (value to change motor setting β, β initially set to 20), continuously set motor value to β. If a new δ value received, change β by that amount. **output**: new β.
- <u>Leg down</u>: **input** β, **output** 20 – beta (actual code a bit more complex to make smoother leg action)
- <u>Alpha advance</u>: **input** α and β. If beta < 5, **output** 15 - α as new α. (if leg is in the air, swing it forward)
- <u>Alpha posn</u>: **input** δ (value to change motor setting α, α initially set to 0), continuously set motor value to α. If a new δ value received, change α by that amount. **output**: new α.

Then there was an AFSM that “balanced” the six legs:

- <u>Alpha balance</u>: Add ups six α values; if the sum was greater than +5 or less than -5, divide by six and output result δ to alpha posn machines. .

There was one further AFSM for each leg, the up leg trigger.

- <u>Up leg trigger</u> – inputs τ and β. When input is received on τ, it puts out 4 - β for next ½ second.

To get a sense of what the robot is doing, imagine that Genghis is standing still, and a signal τ is sent to the right front leg’s up leg trigger. What happens? The trigger puts ot a delta value of 4 - 20 or -16. This is the input to the leg’s beta value which is changed by -16 or raised to 4, almost horizontal.

The alpha posn AFSM for the right front leg would receive this new β. Since this value is less than 5,the leg would take 15 as its new alpha value, that is, the leg would swing forward to 15.

The alpha posn machine would output this α to the alpha balance AFSM. What would be the six alpha values received by this AFSM? With a value of 15 from the right front leg, the SUM ends up greater than 5, so it is divided by 6 and the result is sent to all of the alpha leg posn AFSMs. The result is that all of the other legs move back slightly *moving Genghis forward*.

Giving Genghis some purpose

What Genghis did was follow people. It detected a person by the heat emitted by a body, infrared radiation at 10 micron wavelength. The interaction between the sensors detecting IR radiation and Genghis' mtor controls happened via three "high-level" AFSMs:

- IR Sensors: would continuously output the value of the 6 IR sensors (as 6 bit number – each sensor either detects heat or not). No input

- Prowl: Input → IR sensor list. Output: a message that would inhibit the walk mechanisms unless any sensors are listed on input – in which case don't send message for 5 seconds. So, if IR is detected, start walking.

- Steer: Input → IR sensor list. Compare number of left sensors to number of right sensors. If more on left, instruct left legs to take smaller steps.

The alpha posn AFSM was modified: now input delta and mu. If new delta received, change target motor position by that amount but never to less than mu, and reset mu to -25. Steer would count sensors on left and right. If more on right were registering, then value of mu = 7 would be sent to right-leg machines; similarly for left. The result was smaller steps on the side with more sensors active.

So, here is the tough philosophical question. If you watch the machine, it appears to prowl after humans. You could imagine that even in a simple robot like this, there would be a central program that contained the "intention" to prowl after humans, and that is probably the way we would imagine the robot at first blush. "Hey, I am going to program this robot so it has the intention of prowling after humans." But there is really no intention programmed into this robot as described. So, at what level of complexity would we say that intention appears?

This cutting edge research is now even available for hobbyists, but probably minus the AFSMS ☺:

http://www.sjlrobotics.20m.com/hex-pod.html

Robots and Human Interaction (Breazeal's Kismet)

Kismet and Emotional involvement (Cynthia Breazeal)

Anyone who has messed up a friendship because the right nuance wasn't transmitted by an email message can appreciate how much of human-to-human communication involves non-verbal cues. In email we compensate a little by tossing in emoticons, smiley faces, to gives some clues. At the other end of the spectrum, it appears that people are also eager to attribute human emotions to non-humans if we can spot any resemblance to human facial expressions. Dogs win us over by smiling, cats have that certain air of disdain. If might lessen the strain of dealing with computers and robots constantly if we could see what the machine was thinking as well as read or hear what it says. Well, Professor Cynthia Breazeal of MIT took the Brooks' finite state machine model of robotics and developed a machine that could interact with people in surprisingly human ways.

And, once again, the system displayed how complexity can arise from just a few simple ideas and rules. The robot that Prof. Breazeal and her students developed eventually consisted of 15 computers that did various jobs – picking up audio, video, moving the robot's eyes. It had no central program and so was very much like Genghis but on a grander scale. The fact that the work can be distributed among a variety of computers illustrates one of the attributes and advantages of the finite state model: the so-called cognitive load, the work that computers must do to solve a problem, does not overwhelm a single computer; instead the work is spread out as the problem grows. This leads naturally to thinking about many embedded processors breaking up complex problems, each processor doing its own little chunk of work.

The basis of the robot's seemingly human interaction with subjects is that it paid attention to just three things:

- Moving things
- saturated colors (like kids' toys)
- skin tones

These happen to be three details that robotic vision is good at detecting. Interestingly enough robots see skin tones as skin tones with color not at all important.

The weight it gives to each of these depends on how the robot is feeling ☺ at any moment. For example, if it is lonely, it looks for skin tones; if it is bored, it looks for bright colors. You can imagine that loneliness can be simply a measure of how long since the last human appeared.

In addition, the robot had an **emotional state** built onf three variables – valence (a measure of happiness), arousal (how tired versus how stimulated), stance(how open to new stimuli).

Example: something coming near Kismet's face, it will arouse Kismet

Arousal: if stance and valence are neutral, Kismet will be surprised.

If valence were negative, however, then

If stance open might produce fear

If stance closed, might produce anger

And finally, the robot could understand human speech in the way a baby does. That is, Kismet does not actually understand any words. What it can detect is prosody, the variations in pitch in a person's voice. Mothers across all cultures give four signals to infants via prosody – approval, prohibition, attention-getting and soothing, and that's what babies and Kismet can detect.

The net result of all of this is a robot with which human being are naturally drawn to talk.