Chapter 2 The Nature of Mind and Mind-Body Problem
Stan Franklin

[....Earlier portions of chapter omitted....]

But for now, let's check out Sloman's view of free will as scalar rather than Boolean.[Note 12]

Section: Free Will à la Sloman
I'm rather taken with Sloman's notions about free will (1988). But how do I justify bringing up the freewill issue in a chapter entitled "The Nature of Mind and the Mind-Body Problem"? Well, we've just talked about an agent creating information according to its needs and goals. What if there are goal conflicts? Does our agent exercise free will? That's one connection. Recall that the central issue, always, for a mind is how to do the right thing. Does a mind exercise free will in deciding what is the right thing to do? That's second connection.

But all this is rationalization. Sloman disposes of the freewill problem by showing it to be a pseudo-problem. He refocuses our attention away from the question "Does this agent have free will or not?" by exposing us to all sorts of degrees of free will. I want to convince you that it will be useful to think of all sorts of degrees of having mind, of all sorts of degrees of having consciousness. I hope Sloman's approach to free will serve as a formal analogy.

Sloman maintains that the basic assumption behind much of the discussion of free will is the assertion that "(A) there is a well-defined distinction between systems whose choices are free and those which are not." However, he says,

> ...if you start examining possible designs for intelligent systems IN GREAT DETAIL you find that there is no one such distinction. Instead there are many "lesser" distinctions corresponding to design decisions that a robot engineer might or might not take--and in many cases it is likely that biological evolution tried ... alternatives.

The deep technical question, he says, that lurks behind the discussion of free will is "What kinds of designs are possible for agents and what are the implications of different designs as regards the determinants of their actions?"

What does Sloman mean by "agents"? He speaks of a "behaving system with something like motives." An agent, in this sense of the word,[Note 13] operates autonomously in its environment, both perceiving the environment and acting upon it. What follows is a representative sample of Sloman's many design distinctions, taken mostly verbatim.

Design distinction 1
Compare an agent that can simultaneously store and compare different motives as opposed to an agent that has only one motive at a time I would say that the first exercises more free will than the second.

Design distinction 2
Compare Artificial Minds agents all of whose motives are generated by a single top level goal (e.g., "win this game"), with agents with several independent sources of motivation (e.g., thirst, sex, curiosity, political ambition, aesthetic preferences, etc.). If you're designing an autonomous agent, say a Mars explorer, here's a design decision you have to make. Will you design in only one top-level goal, or will you create several independent ones? If you choose the second, I'd say you must build in more free will. This is going to get tiresome after a while because there are lots of these design distinctions. But that's Sloman's point! I have to show you lots of them, or you may miss it. Just skip ahead when you feel convinced.

Design distinction 3
Contrast an agent whose development includes modification of its motive generators in the light of experience, with an agent whose generators and comparators are fixed for life (presumably the case for many animals).

Design distinction 4
Contrast an agent whose generators change under the influence of genetically determined factors (e.g., puberty), as opposed to an agent for whom they can change only in the light of interactions with the environment and inferences drawn therefrom. In this case, I couldn't say which one exercises more free will. It's not much of an issue any more. The issue dissolves as we focus on whether to design in a certain decision-making property. And I think the issues of what has mind and what doesn't, or what's conscious and what's not, are going to dissolve in the same way when we get down to designing mechanisms of minds.

Design distinction 5
Contrast an agent whose motive generators and comparators are themselves accessible to explicit internal scrutiny, analysis and change, with an agent for which all the changes in motive generators and comparators are merely uncontrolled side effects of other processes, such as addictions, habituations, and so on. In the first case we have an agent that not only can change its motive generators but also can change them consciously. That seems like quite a lot of free will.

Design distinction 6
Contrast an agent pre-programmed to have motive generators and comparators change under the influence of likes and dislikes, or approval and disapproval, of other agents, and an agent that is only influenced by how things affect it. The agent has some social awareness. There's much more to designing agents than just the pseudo issue of free will.

Design distinction 7
Compare agents that are able to extend the formalisms they use for thinking about the environment and their methods of dealing with it (like human beings), and agents that are not. (most other animals?) Agents that can think about their paradigms and change them would seem to have a lot more free will.

Design distinction 8
Compare agents that are able to assess the merits of different inconsistent motives (desires, wishes, ideals, etc.) and then decide which (if any) to act on, with (b) agents that are always controlled by the most recently generated motive (like very young children? Some animals?).

Design distinction 9
Compare agents with a monolithic hierarchical computational architecture where subprocesses cannot acquire any motives (goals) except via their "superiors," with only one top-level executive process generating all the goals driving lower-level systems, with agents where individual subsystems can generate independent goals. In case (b) we can distinguish many subcases, such as (b.1) the system is hierarchical and subsystems can pursue their independent goals if they don't conflict with the goals of their superiors (b.2) there are procedures whereby subsystems can (sometimes?) override their superiors. [e.g., reflexes?]

Design distinction 10
Compare a system in which all the decisions among competing goals and sub-goals are taken on some kind of "democratic" voting basis or a numerical summation or comparison of some kind (a kind of vector addition, perhaps), with a system in which conflicts are resolved on the basis of qualitative rules, which are themselves partly there from birth and partly the product of a complex high-level learning system.

Here we have the distinction between connectionist systems (a) and symbolic AI systems (b). This distinction will occupy us during a later stop on our tour. Surely you've gotten the point by now and we can stop with ten examples, although Sloman did not.

It's a strange kind of argument. Sloman argues against free will not directly but by pointing out that free will is based on the assumption of a sharp distinction. He then says that if you look closely enough, you don't find this sharp distinction. The whole idea is to point out that free will is really a nonissue, that these specific design distinctions are the important issues. He's essentially taking the engineer's point of view rather than the philosopher's, even though he is a philosopher. When we explore the fascinating space of possible designs for agents, the question of which of the various systems has free will becomes less interesting.[Note 14] Design decisions are much more fascinating.

Degrees of Mind
As we begin to make specific design distinctions concerning aspects of mind other than control, the question of mind attribution should dissolve as the freewill question did. Here are a few such distinctions couched in the style of Sloman. Note that all of them could and should be refined into finer distinctions, and that some of them may well turn out to be spurious.

Design distinction S1
Compare an agent with several sense modalities, with (b) an agent with only one sense (e.g., a thermostat).

Design distinction S2
Compare an agent only one of whose senses can be brought to bear on any given situation (e.g., a bacterium [?])., with an agent who can fuse multiple senses on a single object, event or situation.

Design distinction M1
Compare an agent with memory of past events, with an agent without memory. Does a thermostat have memory? It certainly occupies one of two states at a given time. One might claim that it remembers what state it's in. Allowing memory in such a broad sense makes it difficult to imagine any agent without memory. This distinction should be taken to mean memory in some representational sense, perhaps the ability to re-create a representation of the event.

Design distinction M2
Compare an agent with short- and long-term memory, with an agent with only short-term memory. Do insects have only short-term memory for events, or can some of them recall the distant past?

Design distinction M3
Compare an agent that can add to its long-term memory (learn?), with an agent that cannot (e.g., some brain-damaged humans). I can imagine wanting to design an agent that remembers during its development period but not thereafter. An analogous situation is a human child's ability to learn a new language easily until a certain age and with more difficulty thereafter.

Design distinction M4
Compare an agent that can store sensory information from all its modalities, with an agent that can store sensory information from only some of them. I think some of Rodney Brooks's robots satisfy (b), at least for brief periods (1990c). His work will occupy a stop on our tour.

Design distinction T1
Compare an agent that can operate only in real time, with an agent that can plan. This is a graphic example of a coarse distinction that could be refined, probably through several levels. The ability to plan can range from a dog walking around an obstacle to get to food on the other side to a complex chess stratagem involving ten or more moves of each player.

Design distinction T2
Compare an agent that can ''visualize" in all sensory modalities, with an agent that can "visualize" in none, with an agent that can visualize in some but not others. I, for instance, can easily conjure up a mental image of my living room, but I cannot "hear" a melody in my head or imagine the smell of popcorn. Note that this is not simply a lack of memory for music or smells. I can readily recognize familiar melodies on hearing them, and familiar odors as well.

Design distinction T3
Compare an agent that can create mental models of its environment, with an agent that cannot. The utility of mental models is a hotly debated issue just now, in the guise of "to represent or not." We'll visit this topic toward the end of our tour. Also, we have here another obvious candidate for refinement. There are surely many degrees to which an agent could create mental models, mediated by sensory modalities, type and capacity of available memory, so on. Could we say that agents with some of these abilities display a greater degree of mind than the counterpart with which they were compared? I think so. Surely it takes more mind to plan than to act, if only because planning--for an agent, not a planner-- presumes acting. Thinking of mind as a matter of degree may well make the task of synthesizing mind seem less formidable.[Note 15]

We can start with simple minds and work our way up. I expect a real understanding of the nature of mind to arise only as we explore and design mechanisms of mind. Let me remind you of a quote you've already seen in Chapter 1 by way of driving home the point: "If we really understand a

system we will be able to build it. Conversely, we can be sure that we do not fully understand the system until we have synthesized and demonstrated a working model" (Mead 1989, p. 8). The next stop on the tour is zoology, the animal kingdom. If we're to find minds and their mechanisms in other than humans, surely that's the first place to look.

End Notes from Franklin Chapter 2 Relevant to discussion of Sloman's paper
(From note 12)

[12] Computer scientists speak of a Boolean variable, or just Boolean, as one that assumes only two values, usually 0 and 1. A scalar variable, or scalar, assumes one of some finite set of values.

[13] Minsky, in his Society of Mind (1985), uses "agent" to refer to a mental process with a certain amount of autonomy and some particular competence.

[14] Not everyone will agree. Bob Sweeney comments as follows: "Free will has as much to do with the issues of accountability, ethics and 'sin'. These cannot be glossed over by reductionist arguments about design features."

[15] I'm indebted to Dan Jones for this observation. *(Franklin's words)*
_____

Sloman, Aaron, 1988
"Disposing of the free will issue"
Posted to internet Connectionist List, June 19 1988.
Expanded version available online
http://www.cs.bham.ac.uk/research/projects/cogaff/sloman-freewill-1988.html

Additional notes from Franklin's Chapter 2:
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/FranklinSlomanFreewill-extra-notes.html