**PHI Project Related Thought Snippets**

[Mind/Memory Hierarchy](#MindMem)

[Concepts](#Concepts)

[Noise](#Noise)

Mind/Memory Hierarchy:

It seems reasonable that:

* Big concepts are built of smaller ones.
* Information is not endlessly repeated so the same datum/concept is not duplicated in every concept/memory that refers to it. In other words, concepts are referenced with something like pointers.

However, this leads to several consequences:

* If a concept changes, then anything referencing it will also change.
* If the change is small, perhaps this does not have great consequences.
* If , however, the change in the concept is big, or if it becomes invalid, then the concepts/memories that reference it either become incorrect or end up having what is in effect a dangling pointer – ie a reference to a no longer existent concept.
* We can think of the concept hierarchy as a building where “foundation concepts” form the base and new concepts are built on top. Modifying the base can cause large *quakes* in the whole structure.

Is this reflected in what we know about memory development?

When the system has very little experience, there isn’t much that already exists and can be properly considered to be foundational data. This means that structures built are quite likely to be built incorrectly (strange/useless concepts, weird/crazy causality, overly convoluted organizations, very little generalization, etc.) and that these “uneducated” systems will inevitably be subjected to deep revisions as time goes by and “refereeing data” invalidates previous concepts.

This is inevitable:

1. A core of experience **must** be acquired in order to properly create/organize concepts correctly/usefully. In the beginning, there will not be enough history to do so and things will very likely be messy.
2. Since concepts are continually being built on other concepts and memories are continually being built referring to whatever concepts we currently have on hand, a structure with shaky foundations **will** be built at first.

Concrete Consequences:

Since the less experience a system has, the more malleable are its foundations, then “younger” systems will not be able to “remember” as far back as a more experienced system. The concepts that form memories in young systems will get invalidated much more often and these memories will cease to make as much sense. When the concept changes are small, the invalidated concepts might cause only a fuzzy memory to remain. When the change is big, the memory might be sufficiently damaged so as to make it *terminally invalidated*. These probably get discarded or simply wither away through some auxiliary mechanism.

This also provides a more palatable alternative to the idea that we have hardwired phases of development such as strict hardwired learning windows. Instead:

1. We would explain the beginning of a learning window as the (fuzzy) point in time where a sufficiently reliable foundation exists on which to build a particular type of concept. Before this point, either the necessary base concepts do not exist or are in flux, preventing the formation of a solid and persistent new concept.
2. Similarly, the end of a learning phase is not “genetically preprogrammed” but is instead a consequence of the foundations firming up as its concepts become more established and entrenched through verification. Presumably a mechanism exists where a heavily verified and/or relied upon concept is not easily changed as something that is new and/or untried.

Anecdotal Evidence:

1. By observing young children, 8 year-olds, 12-year olds, etc. It is readily apparent that they have \*very\* good memories. There is nothing wrong with their ability to form memories that are precise and robust. Yet, we know that the great majority of these memories will be lost as they grow older. None of us remember much from those years and when we do, not with any great clarity. This explanation proposes that what happens is that most memories become *terminally invalidated* and subsequently lost. Memories of deep emotionally significance, for example, might, on the other hand, still retain enough “pointers” to valid base foundational concepts (the emotions themselves) so as to survive – but with likely great lack of clarity since other components of the memory will probably have been invalidated.
2. In contrast, with older people, we do not see this widespread “wiping out” of whole periods of memory. It is quite common for even very old people to say “they remember that as if it was yesterday” even when referring to events that occurred many decades ago. This explanation proposes that for older people, very large and time tested hierarchies exist at the time of memory formation and that they consequently do not require deep revisions afterwards. Thus the memory is preserved with great clarity.
3. It should be said that this doesn’t mean that older, more experienced people will therefore remember everything. Instead, what is being said is that if other mechanisms mark a memory as worth remembering, it will last longer and retain greater clarity for these more experienced people due to the stability of their mental hierarchy. (We are of course referring to healthy people – at advanced ages, senility may cause the system to malfunction in a variety of ways – in both memory formation and retrieval.)

Consequences to Phi

1. To get the ball rolling, there should probably be some hardwired foundational concepts.  
   1. Simple feelings and emotions seem intuitive: pain, hunger, fear, etc.
   2. Maybe all *qualia* are foundational concepts. However, when we think of color for example, it seems clear that color qualia are not fully hardwired but instead develop over time (but probably very early). Consider that an animal that is born with a mutation that endows it with a new photo-pigment would develop qualia to represent the new colors. Since this happens often throughout evolution with success – it seems that the qualia cannot be fully hardwired because it would be extremely unlike to have both the mutation for the photo-pigment and a suitable mutation to the hardwiring occurring at the same time.
2. Suitable mathematical mechanisms must be devised to:  
   1. Form concepts.
   2. Evaluate the trustworthiness (though validation) and authority (through being relied upon) of concepts
   3. Controlling the ease of modification/invalidation of concepts. (Can concepts be modified or only invalidated and replaced? Or both?)
   4. Recovering from such invalidations both in the concept hierarchy and in memories.

­

Concepts:

1. It seems intuitive that concepts must be some sort of (exclusive?) categorization of the state and that any concept cannot be always active.
2. If two concepts are indistinguishable, they are logically the same concept
3. If a concept is always active, then it is not useful to base a decision. So concepts must be active only part of the time.
4. Similarly, if a concept is never active, again it is useless.

Concepts and info theory:

1. It follows from the above that with concepts, some sort of information entropy argument should apply. Ideally there should be an equal distribution of the activity of concepts or something.
2. Also, since concepts are categorizations of state, we are often interested in cause and effect between them.
3. However, some concepts have very deterministic consequences. It would be counter productive

Noise

With regards to states and concepts, we are always most interested in transitions (changes) between them. The problem is that noise can cause random transitions that cannot be solved by simple discretization.

1. Even with states being categories that are discretized far apart, if the true state of the system is on the border between categories, noise can make them appear to ping-pong back and forth.
2. A mechanism is required to prevent this ping-ponging such as:
   1. Some type of Hysteresis
   2. Some sort of momentum mechanism that keeps the state moving into the new symbol and thus away from the border
   3. Having a “no-man’s land” between states and having a mechanism that generates change events only if the transition is part of a sequence of the form. If the no-man’s land is wide enough, noise will not be able to fully traverse it on its own to enter a new symbol and thus will eventually reenter the same symbol – causing the change to be ignored.

States & Change

1. Changes are the fundamental entities from which any intelligent system is built since, ultimately, only changes can be used to trigger decisions and thus actions (physical or mental) that the system predicts will be beneficial. This follows from the logic that:
   1. If an action is to occur, it must perforce have a cause or decision preceding it to trigger it to begin.
   2. This trigger must occur at a specific moment to have beneficial results.
   3. A “moment” can only be identified if it is different from the moment immediately preceding it in some way.
   4. Thus the change that allows identifying the beginning of a specific moment is THE fundamental entity from which intelligent systems are built.
2. States on the other hand, are related to changes because we can say that a transition from one state to another is what constitutes the change.