A Subconscious Reading Machine

Instead, let's talk about tennis. A tennis player, in running to hit the ball, has many activities taking place. The legs and lower body to move to the right location, the neck and eye muscles to fixate on the ball trajectory, the arm to be moved into a position where it will have a power stroke, the brain sending pulses to the muscles, and receiving kinaesthetic pulses, so the muscles perform their task, irrespective of energy level, the eyes also registering the follow through and other clues from the opponent as to the ball's likely path, the brain making a choice of shot – hit a backhand or run around to hit a forehand, detailed shot selection – topspin or slice. All of this is done in a few seconds, the ball clears the net, spins down and clips the line – accuracy at one part in a thousand. It would be impossible to do these things consciously, or even talk about any of them in any detail, as each relies on so many other things occurring simultaneously.


Reading text is much like playing tennis – there are many things happening at once, they are just not as visible as when playing tennis. Reading requires the use of a subconscious machine, one that has been built over time, and which incorporates all of deduction, modus Tollens, induction, abduction. The child doesn't need to learn these formally – they come for free with a central nervous system with memory – what the child does have to learn are their limits, and by the time they can use a language fluently, the various methods of logical analysis have been heavily overlaid with exceptions, inducing a certain wariness in accepting the obvious. The child has also become expert in the (subconscious) handling of sets and groups – "Jack and Jill went up the hill" or sets where the head of the set exists only as a cypher – and other types of reasoning – temporal, constraint reasoning.

The reader is using words on a page to build a semantic structure in their head. The words hold concepts – snow tyre, tennis ball, multiple-family group adaptive behaviour treatment guidance – and relations between and among concepts. The words may cause concepts to be fleshed out – nuclear power station – or introduce concepts unknown before. An object in the text may link to something many words before, or many words after. Logical threads are being spun, both propositional and existential, as well as relational. The reader is using their vocabulary, in the tens of thousands of words, and their store of concepts, also in the tens of thousands. Many words and concepts will have multiple meanings, which may require local disambiguation – "His back hurts" or "He's back" or "he took a back seat" – or disambiguation from what has been read before in the text, or the context in which the text is being read, or disambiguation must wait, either for later in the text, from later texts, or when the structure is actually put to use. The machine pathways accommodate complex messages, with often there being multiple alternative structures from which to choose.

Where the reading machine differs from the tennis playing machine is that what is read can change the behaviour of the reading machine immediately – an opponent can introduce a new tennis shot and the machine will take a while to become accustomed to it – text can introduce new imperative structure, and it will be read only once.
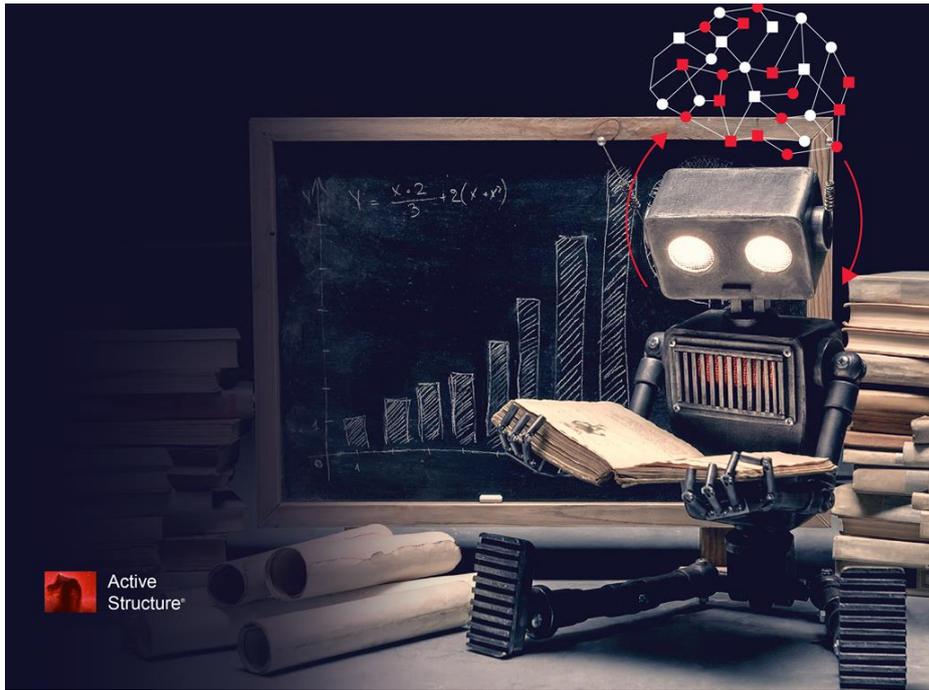
Just as with the subconscious tennis playing machine, asking the reader about what they are doing as they read will lead nowhere – the machine was built subconsciously over twenty years and hundreds of thousands of trials, and is too detailed and too heavily interconnected to be brought to a conscious level.

What does this tell us if we want to automate the reading of text. We need to study the process enough so we become familiar with all the moving parts. Some tips:

- Don't use computer languages as a lead-in – they work completely differently.
- Don't use current programmatic assumptions – relations are objects too, and can have a single point of attachment to what is a complicated structure (or may have many points of linkage).
- Don't throw away what you don't understand – take the time to understand what it is contributing.
- Aim for a high reliability from the automated reading machine – better than for a person.
- Error detection is vital – "cardiopulmonary bypass" is not "artery bypass", even though there is a pulmonary artery – otherwise it looks like a stupid machine.

Unsurprisingly, the subconscious driving machine shares some characteristics of both machines – strongly stereotyped actions – turning the steering wheel in response to a curve in the road - coupled with an executive layer which handles unpredictable events – "Roadwork – New work – no lines marked next 5 km", or being directed around an accident by the police. It also highlights the reliance on trust (which is present in some degree in reading text), but in this application there will be strong antipathy, perhaps leading to fake roadwork which directs the hapless autonomous vehicle over a cliff. Is someone in a blue cap near a flashing light to be trusted? Will this require a huge expansion of the executive layer from the stub currently proposed?

Orion Design Note