Perceptual time, perceptual reality, and general intelligence

Leslie S. Smith

Computing Science and Mathematics, University of Stirling
Stirling FK9 4LA, UK
`l.s.smith@cs.stir.ac.uk`

Abstract. Perceptual time is a critical aspect of how humans (and probably animals too) perceive the world. It underlies general intelligence, particularly where that general intelligence is about interacting with the world on an everyday basis. We discuss what is meant by the perceptual instant, and how this may be important for (artificial and real) general intelligence. Lastly, we briefly discuss how perceptual time might be included in an artificial system which might display general intelligence.

1 Introduction

We believe that the nature of general intelligence is strongly intertwined with the nature of perception. This means that artificial general intelligence (AGI) is more than an abstract concept, but needs an understanding of the nature of perceptual reality in order to develop. This is an aspect of the issue of embodiment: like many (described in `http://en.wikipedia.org/wiki/Embodied_cognition`, and reviewed in [3] and [5]): we believe (but cannot prove) that embodiment is critical for AGI. What this paper considers is one aspect of the issue of what embodiment means in terms of the (internal) environment in which general intelligence (and hence also AGI) operates.

This aspect is perceptual time. This has been largely ignored in this context, yet seems critical for (real) behaviour and hence (real) general intelligence. In this context, *real behaviour* implies both real-world and real-time behaviour. In addition understanding perceptual time may shed some light on the differences between state-based systems and actual neural systems.

2 Perceptual time and perceptual reality

One of the most difficult (and most unfashionable) questions in philosophical neuroscience is that of the nature of the neural construction of reality. That it has a neural construction appears to be generally agreed, yet avoiding the homunculus issue seems difficult, unless the neural activity directly gives rise to the first person experience of being[1]. One important aspect of this question is that of the nature of the neural construction of perceptual time. It is clear that perceptual reality differs from physical reality, and equally that perceptual time is also different from physical time (indeed, this goes back a long way, perhaps to St Augustine and certainly to von Baer [9]. We argue that the nature of the difference between physical reality and perceptual reality (including physical time and perceptual time) is critical for the nature of perception, and, further, for the nature of our everyday interaction with the world. We believe that understanding certain perceptual aspects of

---

[1] This underlies a central issue in artificial "first person" systems: can non-neural (electrical, for example) activity give rise to "first person" (machine?) experience?

time may help to elucidate the differences between current computational approaches and natural generally intelligent systems.

It seems reasonable to accept that the nature of perceptual reality differs between different animals, and from human to human, and indeed, over time for a single person. Given that the location of the generation of perceptual reality (including perceptual time) is in the brain, it seems reasonable to posit that perceptual reality (including perceptual time) for humans has a neural basis. But what is the nature of this neural basis? And how should we look for it? Further, if we can identify and even understand the nature of the neural basis in humans or animals, what are the implications for *artificial* general intelligence? What aspects of it can be re-created in non-neural (e.g. electronic) systems?

Related questions arise for general intelligence. Firstly should one be seeking a specific physical location of the general intelligence, or should we consider this as, in some sense, emerging from the whole brain? Secondly, whatever the answer to the previous question, how should we seek to understand how it operates (investigating how some part of the brain is supporting general intelligence, or alternatively investigating the way in which general intelligence arises from the whole brain)? That is, what is the nature of the relationship between these physical aspects and actual intelligence? We may look for specific structures. Taylor (for example [13]) has suggested that it is to be found in the parietal lobes, although in later papers (e.g. [15], [14]) he moves away from specific locations. Alternatively, we may consider the overall nature of the brain, whether that be its constitution as a very large number of highly interconnected neurons, or in the nature of ionic and neurotransmitter (etc.) interactions within the brain. Finding these physical underpinnings is difficult enough, yet connecting them on to the nature of general intelligence (or awareness, or consciousness or whatever) seems even more difficult.

Because of the difficulty of this problem, we restrict ourselves to considering perceptual time: we believe that there may be lessons from this area for the study of (real) general and artificial general intelligence, and perhaps of the difference between Turing (state-based) machines and neural systems. Further, time is a central issue, because our every sensation, our every action, and that of all living creatures is bound up in time, both physical (external) and perceptual (internal).

3 Time, events, and the perceptual instant

Time, from a physicists viewpoint, is considered as a spatial dimension though which we travel. Each instant is a point, and the points are continuous (or perhaps divided from each other by a very small amount: 5.39×10^{-44} seconds: see `http://en.wikipedia.org/wiki/Planck_time`), and form a 1 dimensional line. Yet our experience of time is very different from this: we experience it as episodes, each with some duration, possibly overlapping, but retaining their order, perhaps inherited from the (underlying) physical time.

One view of perceptual time is of events, each occurring at some "point" in time. Events have been discussed in many contexts, ranging from events in Milner's calculus of communicating systems [8], to synchronising communication in Hoare's Communicating Sequential Processes [6], and the various systems developed from them, such as Σ-algebras, to more generalised views of events, as discussed in the chapters of a recent conference book (From Event-Driven Business Process Management to Ubiquitous Complex Event Processing (EDBPM 2010): see [2]). There are many candidates for events in neural systems, from the release of a neurotransmitter vesicle, to the arrival of a single spike at a synapse, to the initiation of a movement.

Perceptual events are always over some length of time: nothing can happen within a physicist's point of time. As Dunne noted in 1925 "attention is never really confined to a mathematical instant. It covers a slightly larger period." [4] chapter 22.

The duration of the present instant (called "the minimum duration of the conscious present" by Schaltenbrand [11], or the specious present by Clay, renamed the mental present by Whitrow [17]) seems to have two rather different interpretations. On the one hand, there is a lower bound below which the present seems not to be divisible: this is set to about 40ms by von Baer [10], though it is possible to distinguish events closer than this if presented auditorially [11]: however, it is the case that continually presented auditory pulses fuse into a tone at about 18 presentations/second, and continually presented (similar) images fuse into apparent movement at about 20 Hz, suggesting some cross-modal integration time of 40 to 50ms [10]. Such a time period appears to correlate well with certain neural oscillations, such as those found in local field potentials, and may relate to temporal and cross-modal integration (see section 6.1).

Clearly, such perceptual instants are not coded by a purely spatial neural representation, but by one that extends over time.

Yet although we can perceive time as a sequence of events, our perception of time is not as a sequence of such instants. Poppel discusses a longer division of time, particularly in the context of pre-semantic temporal integration, and this he estimates at about 3 seconds [9][10][19]. This longer time period seems to be integrated at a higher level. This longer temporal integration period is associated with conscious perception [9].

4 Time and state

The perceptual views of an instant discussed above differ from a simple temporal ordering of events because the instants extend across time. This means that it is no longer possible to take a "snapshot" of the system. If events occur as points in time, then assuming no further interactions with the system's environment, this snapshot of state determines what will happen in the future to the system. Synchronous logic systems have this property (and very useful it is as well, enabling deterministic computation to be carried out). When events (and percepts are internal events) are spread across time, it is no longer clear what might be meant by such a snapshot.

In general, asynchronous logic systems do not generally have this snapshot property (and a great deal of work is often carried out to ensure that real computer systems which have asynchronous components behave like their synchronous counterparts). There is a realisation that asynchronous operation can bring its own advantages, but the mechanisms of taming this power have not yet, in general been found [1]. Can we use a synthetic version of a perceptual view of time to achieve this?

Real neural systems are highly asynchronous. They do not really have a usable instantaneous state: if, as seems likely, the spikes emitted by neurons are critically important, it is the pattern of spikes (over time, and over the set of neurons) that have been emitted that matters, so that any equivalent of state would need to consider the spikes over some period of time. But over how long? And should other matters (concentrations of different neurotransmitters and neuromodulators, depolarisation of patches of dendrite, for example), also be taken into account? It becomes impossible to know where to stop: as Hong [7] notes, even single interactions between molecules are stochastic because of the rapid thermal movement of the active areas of the interacting molecules.

This suggests that the lack of an identifiable instantaneous state in neural systems illustrates a specific difference between computer and neural systems. One might argue that computer systems can model anything, including systems which have this absence of instantaneous state, and asynchronous nondeterminism, and while that may be true, it would require a very large amount of electronic circuitry to model even a single neuron to any degree of accuracy.

5 Time and context

Context has long been known to be vital for interpreting data. Context may be spatial, temporal, or both. In a computer program, context is (generally) implemented using the internal state (values of variables) within a program, so that the interpretation of some particular datum will depend on the explicitly adjusted values that make up this state. In non-algorithmic modes of computation, (such as those of neural networks and reinforcement learning), context is made up from the values of the different elements in the system. For example, in trained neural networks, the eventual interpretation of some input will depend on the dataset used to train the system (as well as the actual learning rule and architecture of the system). Thus a particular data element is interpreted in the context of the training set.

Context arises at many levels in both real and artificial neural systems. For example, in an integrate-and-fire neuron (which fires when its activity reaches a certain threshold level, after which the activity is reset), there is an activity level context which will determine whether some particular input results in the neuron reaching the firing threshold. In a pure integrate-and-fire neuron, this is simply the sum of all the inputs received since the last time the neuron fired. In a leaky integrate-and -fire (LIF) neuron, the activity has a time constant over which it leaks away, so that the current activity is a function particularly of inputs that have been received recently. In a similar way, Temporal Difference (TD) systems and reinforcement learning systems which gradually (and geometrically) discount recent events and changes also have a temporal context which values more recent inputs more highly than less recent ones.

In real neural systems, the neuronal membrane is leaky, but is not a point-like entity as it is in LIF neurones. Thus there is both a local temporal context, and local spatial context. Further, the strength of this context can be amplified for example through the way in which NMDA synapses work (because the local depolarisation level affects the presence or absence of Mg+ ions that permit these channels to open). At a slightly larger scale, the retina uses the context of both spatially and temporally neighbouring retinal neurones (through the action of the inner and outer plexiform layers) to determine its output, and this is partly responsible for our ability to operate in very variable light levels. Blackboard-based AI systems use the blackboard itself as context: in this case, the particular temporal and spatial (and higher-level) contextual effects are explicit, rather than implicit.

All of this is shows that AI (and other) systems already consider the effect of time, generally implementing its effects through the modulating effects caused by changes made by earlier events. These may be at many different levels: in explicit systems (like blackboard based systems) this is entirely up to the programmer. In implicit systems, it will depend on the different time constants within the system: there are often many of these, ranging from those of individual neuronal patches of membrane, to much slower effects resulting from gradual alteration in weights, such as might occur through STDP or back-propagated delta rule weight alteration. In section 3, we are arguing that there is a specific set of temporal contextual constraints at work in neural systems (and that these may well differ for different animals, and indeed, different values may be appropriate for different tasks). The temporal context applied in the systems discussed in this section is one way of achieving the same effect: we suggest that more careful consideration of the neural approach to perceptual time might lead to better, and perhaps more effective, temporal contextual modulation.

6 Why perceptual time matters for artificial general intelligence

Time, in terms of ordering of inputs and outputs, has always been included in AI and AGI systems. It is true that systems for interpreting or classifying static images can ignore time: but clever though these may be, they are not intelligent systems. Only simple pattern discrimination systems such as back-propagated delta networks or radial basis function networks consider patterns one by one, without reference to their ordering. Further, each individual pattern is presented all at once. But even in these cases, when training is taking place, the order of presentation may matter, as the internal parameters are gradually altered in a non-linear way as a result of each patterns being presented (unless specific care is taken to avoid this, as occurs in so-called batch-processing weight update).

Taking general intelligence to be some mixture of common sense behaviour in a known or unknown environment and maintaining an organisms's overall goals under the vagaries of an unpredictable environment, it is clear that time plays a critical part. As discussed in section 5, this is not new, but how might the ideas on the perceptual instant in section 3, and on the effect of time on state-based machines (section 4), impact on the design of generally intelligent systems?

6.1 The perceptual instant

We consider the perceptual instant first: as noted in section 3, there seem to be two gradations of perceptual instant, one being around 40 to 50ms, and the other considerably longer at about 3 seconds. It seems possible that the faster of these relates to the way in which local oscillations occur in neural columns, and this may well be critical for cross-modal integration of senses. The timescale is within the range of beta oscillations (15 to 30Hz: i.e. 33 to 66ms period), and there are suggestions that these and gamma oscillations (30-80Hz: 12.5 to 33ms period) may be implicated in sensory integration [16]. These oscillations are strongly tied to the architecture of the cortical column, and specific mechanism related to the interplay between excitatory and inhibitory neurones have been suggested to underlie this behaviour [18]. Further, these oscillations have been suggested to be critical for encoding relations and binding different aspects of percepts [12].

It thus seems likely that this fast perceptual instant is closely linked to the columnar architecture of the cortex, and how it fuses the different aspects of sensory perception. Thus it is likely to be critical in human perception and perhaps human general intelligence. However, this does not necessarily imply that it is important for *artificial* general intelligence. Yet there are undoubtedly links between the nature of our perceiving organs, the coding of these percepts as they are converted from the actual transducer, through the brainstem, to the cortex, and the timescale of the sensory integration. These strongly colour our perception of our environment. One result that this has, is that the key percepts that humans use, the percepts that drive our interaction with our environment, take place (at least at one level) over this timescale. One may argue that this may be either a cause or an effect: for example, visual and auditory effects from a remote stimulus (like someone hitting a nail with a hammer) 10 metres away arrive about 30ms apart. Processing at this timescale influences what we consider to be general intelligence, at least in terms of the percepts that we expect to contribute to it.

Although there are slow oscillations within the cortex, they do not appear to be correlated with particular neuronal structures (beyond the cortex), or behaviour (beyond REM sleep): see `http://www.scholarpedia.org/article/Thalamocortical_oscillations`. Perhaps this is not surprising, as the longer perceptual instant seems to be more like a travelling window, gathering together a number of shorter perceptual instants, than a three-second tiling of physical time. As noted in section 5, many models of neural systems do take time into account. However it tends to be a travelling exponentially decaying mechanism that is used, rather than a more even one, illustrated figure 1a and 1b . What lines c and d in figure 1 suggest, is that up to a certain time into the past, recent events may be treated equally in terms of their contribution. Line c suggests a sudden change after a particular length of time, which is perhaps inappropriate, but line d suggests that there could be a decreasing contribution for some longer time. Such an integration interval does seem to coincide with a common-sense view of the world, where events that occurred in the last few seconds do contribute equally to the current state of the world, with events that happened a little longer ago having a smaller (or perhaps already acknowledged) effect. (Again one can argue that this might be either cause or effect: the entities that matter to us in the world are grouped into this timescale, and therefore our time percepts work in this way:

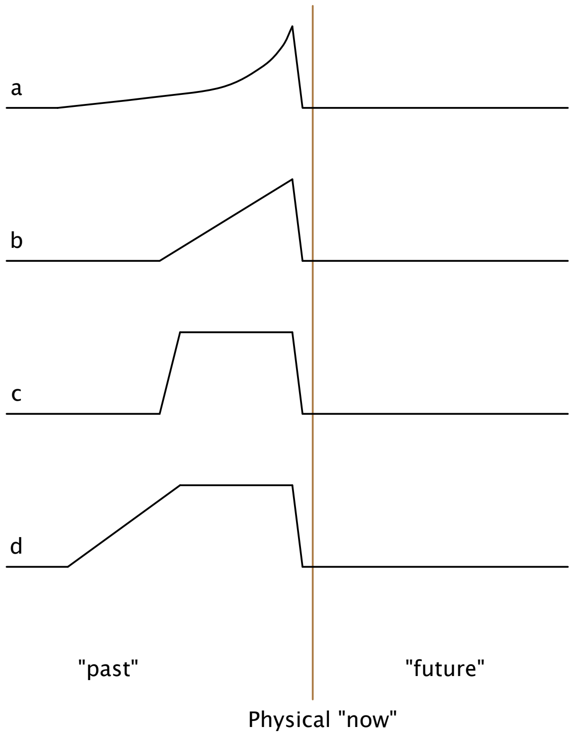

Fig. 1. Schematic of time course of integration of information. Lines *a* and *b* show the integration that is implied by reinforcement or temporal difference learning, where the "current" time is the most important, and there is an exponential (*a*) or linear (*b*) drop off in the importance applied to previous information. Lines *c* and *d* show integration of recent time equally over a period, with a sudden drop-off in *c*, and a more gradual drop-off in *d*. We note that there is always a short delay between the current time and any information being used.

or our time percepts work in this way, and this results in the entities that matter to us in the world being grouped at this timescale.)

6.2 Implications for implementations

Given that the nature of perceptual time is an important aspect of human general intelligence, we are now interested in how to build a system whose percepts and reactions to percepts bear a resemblance to the timings used in human intelligence.

In section 5 we noted that temporal coding (which is certainly required for the types of perceptual time that we are discussing) means that there is no clear-cut notion of instantaneous state. What implications does this have for simulating such systems on standard (state-based) computers, or for implementing artificial general intelligence which includes perceptual time? It means that whatever representations are actually used, they must be representable as a state vector. Thus the simulation necessarily loses some accuracy, and quite possibly represents entities using different mechanisms. Whilst it is possible that these result in deep-seated differences between the capabilities of neural and computational systems[2], we believe that it should be possible to create a system programatically that can emulate perceptual time.

What might such a system consist of?

Such a system would be essentially asynchronous (though if implemented on a digital computer, it would be implemented on synchronous logic). It would have a number of parallel processing entities, processing different modalities. Note that each modality might be associated with a particular sense, but that there might well me multiple modalities per sense (for example, one might choose to process the *where* and the *what* information both in auditory and visual sensory systems separately). These would be integrated over a 30 to 50 ms timescale. This would model (in a functional sense) the cortical columns beta band oscillation, but would almost certainly not be implemented in the same way. (Of course, this implies a better understanding of the nature of the processing in these cortical columns, beyond that in [16] and [18]). This would enable a machine-based representation that matched the shorter version of the perceptual instant. It would aim to group together processed sensory information in pieces that represent events in the environment that take place over these timescales. (Of course, for a completely different, perhaps virtual, environment, the timescales might also be completely different.)

It is less clear how one should implement the three second long perceptual instant (and we note that it might not be a three second long perceptual instant in a different type of environment). We have no clue as to what the neural representation might be. One possibility would be to consider a short-term blackboard-like store which is refreshed from the short-term perceptual instant, but which loses information after about three seconds. This would be used to mirror the moment-to-moment awareness of the world in which we normally live. More likely, it would integrate the shorter perceptual instants over a dynamically varying period, reflecting the changing circumstances of the AGI system.

Adaptation to the environment would use something like Reinforcement Learning or Temporal Difference Learning, but with a discount function more like that in figure 1c or 1d. Action choice would take place at a number of temporal levels, one corresponding to the fast integration (30 to 50ms), one to the slower but immediate temporal percept, at about 3 seconds, and one related to slow considered planning over a longer timescale. (Again we note that these timescales relate to human-level interaction with the real environment, but might be quite different in other environments.) One might consider the fastest of these to be like reactive actions, the middle one to be more like the immediate actions that humans take, and the slowest one to relate to the fulfilment of longer term plans and goals.

7 Conclusions

It is difficult to imagine an artificial general intelligence operating in a real environment unless it can process events and percepts in time in a way which at least bears some relation to how events and percepts are processed in time by real intelligent systems. Most of the activities which animals (and presumably artificially generally intelligent entities) perform take place over time, whether that be opening a door, navigating a route, telling a story, playing music or any other activity. We have looked at what appear to be the basics of human time perception, and tried to show how these might be transferred to machines. A great deal more work needs to be done to actually implement such a system, and we have tried to show what might be initially required.

---

[2] See `http://www.cs.stir.ac.uk/~lss/recentpapers/lss_edinburgh_oct2007.pdf` for more discussion of this.

References

1. Buesing, L., Bill, J., Nessler, B., Maass, W.: Neural Dynamics as Sampling: A Model for Stochastic Computation in Recurrent Networks of Spiking Neurons. PLoS Computational Biology **7**(11) (November 2011) e1002211
2. Cezon, M., Wolfthal, Y., eds.: Towards a Service-Based Internet. Lecture Notes in Computer Science. Springer Verlag (July 2010)
3. Chrisley, R.: Embodied artificial intelligence. Artificial intelligence **149**(1) (September 2003) 131–150
4. Dunne, J.: An Experiment with Time. Black (January 1929)
5. Goertzel, B.: Patterns, hypergraphs and embodied general intelligence. Neural Networks, 2006. IJCNN'06. International Joint Conference on (2006) 451–458
6. Hoare, C.A.R.: Communicating sequential processes. Springer-Verlag New York, Inc (January 2002)
7. Hong, F.T.: A multi-disciplinary survey of biocomputing: Part 1: molecular and cellular aspects. Imperial College Press (January 2005) 1–129
8. Milner, R.: Calculus of Communicating Systems. Springer Verlag (July 1980)
9. Poppel, E.: Lost in time: a historical frame, elementary processing units and the 3-second window. Acta neurobiologiae experimentalis **64**(3) (2004) 295–302
10. Poppel, E.: Pre-semantically defined temporal windows for cognitive processing. Philosophical Transactions of the Royal Society B: Biological Sciences **364**(1525) (May 2009) 1887–1896
11. Schaltenbrand, G.: Consciousness and time. Annals of the New York Academy of Sciences **138**(2) (February 1967)
12. Singer, W.: Neocortical rhythms: an overview. In von der Malsburg, C., Phillips, W.A., Singer, W., eds.: Dynamic co-ordination in the brain. MIT Press: Strungmann Forum Reports (January 2010) 159–168
13. Taylor, J.G.: The Central Role of the Parietal Lobes in Consciousness. Consciousness and Cognition **10**(3) (September 2001) 379–417
14. Taylor, J.G.: On the neurodynamics of the creation of consciousness. Cognitive Neurodynamics **1**(2) (December 2006) 97–118
15. Taylor, J.G.: CODAM: A neural network model of consciousness. Neural Networks **20**(9) (November 2007) 983–992
16. Wang, X.J.: Neurophysiological and Computational Principles of Cortical Rhythms in Cognition. Physiological Reviews **90**(3) (July 2010) 1195–1268
17. Whitrow, G.J.: The natural philosophy of time. Thomas Nelson and Sons (October 1961)
18. Whittington, M.A., Kopell, N.J., Traub, R.D.: What are the local circuit design features concerned with coordinating rhthyms? In von der Malsburg, C., Phillips, W.A., Singer, W., eds.: Dynamic co-ordination in the brain. MIT Press: Strungmann Forum Reports (October 2010) 115–132
19. Wittmann, M.: The inner experience of time. Philosophical Transactions of the Royal Society B: Biological Sciences **364**(1525) (May 2009) 1955–1967