Sandberg, Anders (2013), 'Feasibility of whole brain emulation', in Vincent C. Müller (ed.), *Theory and Philosophy of Artificial Intelligence* (SAPERE; Berlin: Springer), 251-64.

Feasibility of Whole Brain Emulation

Anders Sandberg

Future of Humanity Institute

Faculty of Philosophy

Oxford University

Anders.sandberg@philosophy.ox.ac.uk

Introduction

Whole brain emulation (WBE) is the possible future one-to-one modeling of the function of the entire (human) brain. The basic idea is to take a particular brain, scan its structure in detail, and construct a software model of it that is so faithful to the original that, when run on appropriate hardware, it will behave in essentially the same way as the original brain. This would achieve software-based intelligence by copying biological intelligence (without necessarily understanding it).

WBE has been a staple of science fiction and philosophical thought experiments for a long time, from the early futurist visions of (Bernal, 1929) to (Parfit, 1984)(Chalmers, 1995)(Searle, 1980). While the philosophical literature has explored the possibility as a tool for elucidating theories of identity and mind, it has not overly concerned itself with the issue of whether it could actually be achieved technologically.

The first attempt at a technical analysis of brain emulation was a report (Merkle, 1989) reviewing automated analysis and reconstruction methods for brains. It predicted that "a complete analysis of the cellular connectivity of a structure as large as the human brain is only a few decades away". The first popularization of a technical description of a possible mind emulation scenario can be found in *(Moravec, 1988)*, where the author describes the gradual neuron-by-neuron replacement of a (conscious) brain with software. Since then a number of reports have attempted to analyse the technical requirements and constraints of WBE (e.g. (Sandberg & Bostrom, 2008) and (Parker, Friesz, & Pakdaman, 2006)), several projects aimed at large scale scanning and reconstruction have been started (Blue Brain, the Human Connectome Project, brainpreservation.org) and there is also a renewed philosophical interest in the possible impact of software intelligence (Chalmers, The Singularity: A Philosophical Analysis, 2010)

WBE is interesting for several reasons:

- It is the logical endpoint of computational neuroscience's attempts to accurately model neurons and brain systems, and the emergent dynamics that occur in such models. Neuroinformatics, like other areas of bioinformatics, aims at documenting maps as complete as possible of biological systems at different levels of resolution. WBE would be a combination of an accurate map and sufficiently accurate modeling.
- WBE might produce useful data or inspiration for AI even if the full aim is never realized.
- WBE might lead to AI and possible superintelligence through mental enhancement (Chalmers, The Singularity: A Philosophical Analysis, 2010).
- Attempts at brain emulation would itself be a test of ideas in the philosophy of mind and philosophy of identity (Shores, 2011).
- The impact of successful WBE could be immense. Given that human capital is a main driver of economic growth, copyable human capital (in the sense of systems able to perform the same tasks as a human) implies extremely fast economic growth, and would have profound societal and ethical consequences (Hanson, 1994) (Hanson, Economics of the singularity, 2008). Even low probability events of such magnitude merit investigation, especially if early coordination is necessary to avoid disastrous outcomes.

WBE represents a formidable engineering and research problem, yet one that appears to have a well-defined goal and could, it would seem, be achieved by extrapolations of current technology. This is unlike many other approaches to artificial intelligence where we do not have any clear metric of how far we are from success.

Arguments of incredulity are not sufficient to disprove WBE – the complexity of the brain might be high, but there are many of examples where people have scanned or simulated very complex systems (genomes, proteins, integrated circuits) that would have appeared infeasible just a few years earlier. We cannot trust intuitions formed in scientific and technological environments different from the environment where the eventual development will take place.

Still, the existence of simple prototypes today does not constitute a proof of eventual success; the way to avoid the “first step fallacy” (Dreyfus, 1992) is to look at the constraints of the process and preconditions that might imply its eventual infeasibility. This paper will explore the feasibility of WBE, investigating what preconditions - philosophical, scientific and technological - are necessary for various degrees of success and the extent they can be estimated given our current state of knowledge.

Simulations and emulations

Simulations are processes that mimic the relevant features of target processes (Hartmann, 1996). A computer simulation is an attempt to model a particular system by creating a software representation that represents objects, relations and dynamics of the system in such a way that relations between objects in the simulation map onto relations between equivalence classes of objects in the original system.

Simulations can be of different levels of resolution. For the current paper we will focus on simulations that attempts to achieve full functional equivalence – all relevant behavioral properties and internal causal links of the original system are replicated. Exactly what this requirement entails depends on both the success criterion used by human researchers (the goal aimed for) and the corresponding necessary resolution scale in the brain.

A key issue in simulation science is validation, testing that the real and simulated systems correspond to each other. There are three types of validity (Zeigler, 1985) (Zeigler, Praehofer, & Kim, 2000):

- Replicative validity: the simulation matches already observed data from the real system (retrodiction)
- Predictive validity: the simulation matches data before they are acquired from the real system (prediction)
- Structural validity: the simulation "not only reproduces the observed real system behavior, but truly reflects the way in which the real system operates to produce this behavior." (Zeigler, 1985) In this case the map between the real system and the simulation is a homomorphism: all relationships between elements in the real system have corresponding relationships in the simulated system.

Given that scanning methods of brains are very likely to be destructive (Appendix E of (Sandberg & Bostrom, 2008)) predictive validity in the simple sense might be impossible. Conversely, by the definition of WBE structural validity is a necessary condition for success. However, this is not directly observable: we cannot know that all parts are included, merely that the replicative validity is good.

Emulations

In software engineering the term emulator is used for hardware and/or software that duplicate the functions of a computer system in another computer system. Typically the focus is on exact reproduction of external behavior rather than the exact internal structure. Internal states need to change as a function of inputs, producing outputs compatible with the modeled system but the states are not necessarily corresponding to any components of the system.

Many emulation are used to run software from older computers on newer computers; here the emulation of the old hardware and operating system underlying the software layer allows the execution of the software to be simulated in a one-to-one manner. Emulation in this sense is something enabling an accurate or one-to-one simulation by providing a sufficiently accurate interface that imitates low-level functions that are not relevant to the simulation[1].

An impressive example of such emulation is the reconstruction and emulation of the MOS 6502 processor by the Visual6502.org project. Unlike normal emulator construction based on implementing the description in chip specification documents this project scanned and interpreted a physical instance of the processor. Working from a single chip they exposed the silicon die, photographed its surface and substrate at high resolution, generated polygon models of the individual components, used the known rules for how they intersect to form circuits to automatically deduce the circuit diagram and hence produce a transistor-level simulation of the chip.

The reason for the physical scanning was that available design information tends to be incomplete or incorrect and manual reconstruction from the actual chip is not feasible for complex chips (Visual Transistor-level Simulation of the 6502 CPU, 2011). In this regard the project has many similarities to a hypothetical WBE project, although of course it was helped by the fact that the chip has a well-defined structure, perfectly understood components and merely 3,510 transistors.

This simulation is capable of running any programs the original processor could, not just emulating the response to instructions but the actual logic. It leaves out resistance and capacitance, has no propagation delays and makes use of some simple heuristics to handle analog behavior of transistors (James, Silverman, & Silverman, 2010). There is no need to perform an electrical simulation of the components (or hardware in the second case) since the digital nature of the system allows a sharp abstraction boundary where higher level layers do not depend on the details of lower levels. As we will see, the issue of whether sharp abstraction boundaries exist in the brain is of key importance for the feasibility of brain emulation.

---

[1] It should be noted that the use of emulation here diverges slightly from the usage in (Sandberg & Bostrom, 2008); there it denotes structurally valid simulation, while here it denotes a platform that enables a structurally valid simulation.

Philosophical feasibility

Philosophy of mind: physicalism, functionalism

WBE assumes that everything that matters in brains supervenes on the physical. The major difference to AI is that WBE does not only require physicalism, but that all relevant properties are in principle observable. If mental supervenience requires properties that can never be observed for some reason, then WBE would not be feasible while strong AI might still be achievable. The functional relations between the properties might be unobservable, preventing the construction of brain emulations in general, or individual properties of importance might be unobservable so that while emulations are possible gaining the necessary data to make an emulation of a particular brain will remain out of reach.

WBE makes roughly similar assumptions as strong AI about the philosophy of mind when it comes to the machine implementation of intelligent behavior, at least in the wider sense of the term "strong AI" as systems that act like they have minds rather than the more precise original sense in (Searle, 1980) – some success criteria for WBE do not require a mind emulation, merely appropriate behavior.

WBE is also committed to functionalism, since it assumes that by copying the functional relationships of a brain the relevant properties are copied or will emerge from their execution. A successful WBE project implies multiple realizability since the software could be copied to multiple hardware platforms.

As noted by (Shores, 2011), WBE might act as a test for theories of downward causation or holism of minds: while WBE assumes an emergent mind, it assumes a particular form of emergence from simple components that might not be compatible with other theories of emergence.

Can meaningful degrees of success be defined and observed?

The degree to which simulations are judged successful in science depends on how well a simulation achieves the desired function of the simulation in the scientific process. This does not have to correspond to a close match of behavior if the goal is to inspire experiments, or act as pedagogical or heuristic tools. Simulations used as substitute for experiments on the other hand will be judged as more successful the closer their results match their counterpart real experiments, at least along the dimensions the experiment aims to measure.

However, WBE can aim at something different from improving scientific understanding. It can also be an engineering goal, where it is the usable result that matters. A working simulation of the human mind that does not help lead to an understanding of how intelligent behavior is generated may be scientifically useless, but could still hold great practical and philosophical value.

The development of WBE would entail a sequence of generating simulations based on theory and measured data, comparing them with reality, building revised simulations, and so on. A somewhat unusual aspect is that it also includes constructing technological tools for automatically converting biological inputs into simulation: the project includes not just the normal practice of simulation but a partial automation of it. It is not implausible that attempts to automate aspects of validation and verification would also be included, producing a semi-automated simulation building pipeline.

It is possible to distinguish several potential success criteria for WBE:

1. "Functional brain emulation": The emulation simulates the objects derived from brain scanning with enough accuracy to produce (at least) a substantial range of species-typical basic emergent activity of the same kind as a brain (e.g. a slow wave sleep state or an awake state). It exhibits generically correct causal micro-dynamics but not functionally unified into meaningful behavior.
2. "Species generic brain emulation": The emulation produces the full range of species-typical emergent behavior and learning capacity, but does not have any behavior linked to the individual brain(s) used for scanning.
3. "Individual brain emulation": The emulation produces emergent activity characteristic of that of one particular (fully functioning) brain. It is more similar to the activity of the original brain than to the activity of any other brain.
4. "Social role-fit emulation"/"Person emulation": The emulation is able to fill and be accepted into some particular social role, for example to perform all the tasks required for some normally human job.
5. "Mind emulation": The emulation produces subjective mental states (qualia, phenomenal experience) of the same kind that would have been produced by the particular brain being emulated.
6. "Personal identity emulation": The emulation is correctly described as a continuation of the original mind; either as numerically the same person, or as a surviving continuer thereof. The emulation is an object of prudentially rational self-concern for the brain to be emulated.

Of these success criteria only 1-4 are directly observable. Criterion 4 is a borderline case since it depends on interaction with others. The emulation should be able to pass a personalized Turing test: outsiders familiar with the emulated person would be unable to detect whether responses came from the original person or emulation.

An emulation that exhibits these individual traits might still fail at being a mind emulation (it lacks mental properties) or person emulation (it lacks necessary aspects of personal continuity). However, success criteria 5-6 does not appear to be directly observable and to what extent they might be entailed by the criteria for 3 and 4 depends on what theory of mind and identity is adopted (Chalmers, The Singularity: A Philosophical Analysis, 2010).

Success criterion 5 assumes multiple realizability (that the same mental property, state, or event can be implemented by different physical properties, states, and events). Sufficient apparent success with WBE would provide persuasive evidence for multiple realizability. Generally, emulation up to and including level 4 does not appear to depend on any strong metaphysical assumptions.

Chaos

An issue is whether simulations of chaotic systems are meaningful. Given that the brain almost certainly contains chaotic dynamics (since even a three neuron system can become chaotic (Li, Yu, & Liao, 2001)), the state of a simulation will diverge from the state of the original quickly and the predictive validity of the simulation appears low.

However, what matters is the dynamics and causal structure, not the exact dynamic state. Brains or minds in a slightly different activity states are still recognized as the same brains or minds, even though their contents might differ. There exists a significant amount of noise in the brain but it does not prevent meaningful brain states from evolving despite the indeterminacy of their dynamics. The structural validity demand on WBE does not demand identical output of the simulation and the modeled brain, merely output that is compatible with the output that would have been given by the brain if it had been in a similar internal state.

While predictive validity is important for many scientific models it has not the same weight in engineering, where a predictable behavior is more important. For a full WBE long-term divergence is also expected: if learning processes and different experiences doesn't cause the system to change in character like a real brain would change, it would not be a successful WBE.

Non-organicism

A key assumption, characteristic of the WBE approach to AI, is non-organicism: total understanding of the brain is not needed, just understanding of the component parts and their functional interactions. In normal science top-level understanding is seen as the goal, with detail understanding merely a step towards it. This is why WBE evades many of the standard issues of explanation in the philosophy of simulation: it does not attempt to explain the brain or mind, just copy them.

Can a system be copied without understanding its purpose? It does not seem implausible that a person with no understanding of carpentry (or a mindless robot) could follow sufficiently detailed IKEA instructions to build a piece of furniture.

A better understanding of the high level aspects would enable them to perform better, but it is not necessary. What is required is the appropriate low-level actions that builds the system.

A simple example of how understanding may not be required for creating complex simulations is software compilers. Compiler programs do not understand software and merely perform syntactic operations that transform human-readable source code into machine executable machine code. Similarly a WBE pipeline might without any understanding mechanically convert a physical system (a brain) into a software system (a simulation).

Constructing the WBE pipeline might embody a sophisticated understanding of the brain: requisite scan resolution and modalities, how components work, how to test and validate the system. The claim of WBE is that this understanding does not have to extend to the meaning of neural systems, merely their internal function. One could imagine that the team that reconstructed the 6502 processor were given an unknown chip to reverse engineer: their method would have a good chance of succeeding, although they would have a hard time testing the validity of their reconstruction. This also shows a key challenge for WBE: even for fairly modest success criteria, it can be hard to validate against a system whose function is unknown.

Holist theories of mind suggest that everything that is going on will be a function of what all the other components are doing, with little hierarchy (Thompson, Varela, & Rosch, 1991). While the holistic system might emerge if the parts are in place and in the right states, the holistic view might be an argument against non-organicism: there is no way of separating the levels, and the required understanding to create the emulation will be distributed across them. This links with the scale separation issue below.

Scientific issues

This group of issues deals with the actual physical properties of the brain and the possibility of humans inferring enough information about them to achieve WBE. It also includes the methodological question of how a WBE research program could be implemented so as to approach a successful emulation over time.

Level of understanding

A key issue is what level of detail of understanding the brain is needed. This is closely tied to size scales: a higher level of detail typically requires gaining neuroscientific information on smaller scales, requiring new modalities of measurement. High resolution scanning also produces more information, requiring more storage and processing. Abstract models on the other hand require more complex functional understanding of the systems, but less data. The fundamental approach of WBE is that it trades high-level understanding for brute force requirements.

At present there is no consensus on what level of understanding would be needed to achieve WBE. An informal poll among researchers suggested that the electrophysiological level (cellular compartments) is most popular, but this merely represents an opinion (Sandberg & Bostrom, 2008). Scale separation might represent a principled way of reaching a consensus.

Finding biological modalities

Analysing the potential of the WBE project also involves estimating the number and complexity of biological modalities that need to be modeled. Some issues such as whether dynamical state, the spinal cord, volume transmission or glia cells need to be included can already be estimated with some precision and does not pose any insurmountable simulation problems (Sandberg & Bostrom, 2008). Known unknowns such as the number of neuron types, neurotransmitters or relevant metabolites can be bounded. While estimating what remains to be discovered in a finite domain is surprisingly problematic (compare with attempts at estimating the number of species on Earth (Bebber, Marriott, Gaston, Harris, & Scotland, 2007)) the boundedness of the number of possible entities means that the complexity of the simulation is not as strongly affected by new discoveries as it would be by requirements of finer resolution.

The interesting challenge is issues of assessing unknown unknowns, such as whether there exist entirely new forms of interactions in the brain. This is truly unpredictable, even by analyzing past discoveries. The only way to be certain all relevant processes have been included in a simulation is successful brain emulation. Conversely, failure of WBE attempts can give information about missed modalities, especially if they are done in close conjunction with in vivo studies.

Computability

WBE assumes that brain activity in large is Turing-computable. Should important functions be uncomputable WBE becomes infeasible (at least on conventional computer architectures: it might work on unconventional hardware). At present there is no convincing empirical evidence for uncomputability in the brain, although there is no shortage of claims for it.

A related challenge is component tractability: can the simplest components simulated be understood and measured? For example, if the quantum-mind proposals of (Penrose, 1989) were true, the relevant components might be quantum states that cannot be measured even in principle, even if their dynamics were known and implementable on suitable quantum computer hardware.

A less exotic form of component tractability problem might be the need for analog signals or the right kind of randomness. While we have argued that these are unlikely to matter due to noise constraints (Sandberg & Bostrom, 2008), others have responded that they might hold an important role in the mind (Shores, 2011).

Scale Separation

In order for simulations on a particular scale to be valid, states and interactions on smaller scales must be encapsulated within the states and interactions of the emulation. Otherwise microscale events would produce macroscale outcomes that are not captured by the dynamics of the simulation.

In some physical systems scale separation occurs: there exists a level where interactions on shorter length and time scales average out, producing macroscale dynamics uncoupled from the dynamics on smaller scales (Hillerbrand, 2007) A typical example is the statistical mechanics of gases, where the exact molecular interactions do not matter for deriving equations of state describing the macroscale behavior of the system. Another example is the scale separation between electric currents and logic operations in a computer, which enables emulation such as the earlier mentioned the 6502 emulation. Unfortunately not all systems show scale separation. Turbulent flows for example show correlations between size scales that make them interdependent: a simulation leaving out events at a fine resolution will produce nonphysical behavior (Bec, Cencini, Hillerbrand, & Turitsyn, 2008).

Scale Separation is a key challenge for the WBE project. Does there exist a scale in the brain where finer interactions average out, or are each scale strongly linked to larger and smaller scales? If no such scale separation exists, then the feasibility of WBE is much in doubt: no method with a finite size cut-off would achieve emulation. Biologically interesting simulations might still be possible, but they would be local to particular scales and phenomena. The existence of scale separation is a fundamental requirement of WBE, a practical problem for finding the optimal resolution of the model, as well as an intriguing scientific problem.

In neural modeling it is common to separate the “mnemonic equations” (permanent or quasi-permanent changes in neural activities, such as memory) from the “neuronic equations” (the instantaneous behavior of the system) decoupled (Caianello, 1961) because they typically occurs on different time scales and hence are assumed to be largely decoupled. While the scale separation between different levels of the nervous system does not have as radical separation as in statistical physics, the different levels of the hierarchy – neural fields, neuron populations, individual neurons, ion channels – are often separated by one or two orders of magnitude, and may hence be amenable to statistical treatments that average small and fast scales, possibly introducing random noise from the coarsegraining of microdynamics (Berglund, 2011).

Identifying scale separation

One way of identifying scale separation is to analyze the capacity for error correction, where processes either dissipatively dampen deviations or they do not have any effect on other systems. In gases macrostates are treated as identical, in digital circuits small deviations in voltage are still treated as one or zero. In neurons small differences in membrane potential have no effect on the all-or-nothing action potential generated or the postsynaptic potentials; at most they can act by influencing the exact timing of the signal generated.

Given that brains evolved to function in a noisy environment where external (e.g. environmental conditions, microtraumas, changing nutrient states, parasites etc.) and internal disturbances (e.g. developmental noise, thermal noise, chemical noise) are common, various forms of error correction and robustness should be expected. Brains sensitive to microscale properties for their functioning would exhibit erratic and non-adaptive behavior. If the differences introduced by simulation are smaller than the normal noise level (and of correctable type) then it is likely that scale separation would occur.

A model of a dynamical system might deviate from the original system due to uncertainty in initial conditions, parameter uncertainty and model uncertainty. Typically the measure of points in parameter space where the dynamics shifts qualitatively is small, and for a biological system one should also assume that minor changes in structure do not cause catastrophic deviations: they would tend to evolve towards structural stability. Hence the qualitative properties of the system have a finite tolerance, and a simulation within this tolerance would produce similar behavior.

Empirical bounds on scale separation in the brain

Microstimulation of individual neurons can influence sensory decisions (Houweling & Brecht, 2008). In their experiment rats were trained to behaviorally respond to microstimulation of single neurons, showing that scale separation doesn't occur between the single neuron firing level and the behavioral level. However, the experiment only succeeded for 5% of the trials and often just induced weak and slow biases. It is not clear whether the experiment could succeed with single synapse stimulation.

The noise level in the nervous system is fairly high, with spike-timing variability reaching milliseconds due to ion channel noise. Perceptual thresholds and motor precision are noise limited. Various noise management solutions such as redundant codes, averaging and bias have evolved (Faisal, Selen, & Wolpert, 2008).

In synapses the presynaptic transient influx of calcium ions as a response to an action potential corresponds to just 13,000 ions (Koch, 1999) (p. 458), and on the postsynaptic side just 250 ions (Koch, 1999)(p. 302). These numbers are so small that numeric noise begins to be significant, and the chemical dynamics can no longer be described as average concentrations. However, biological systems can resist the discretization noise through error correction mechanisms that lead to discrete attractor dynamics, in line with the evidence that synaptic plasticity involve discrete changes rather than graded response (Ajay & Bhalla, 2006) (Bhalla, 2004)(Elliott, 2011).

It is hence not implausible that there exist sufficient scale separation on the synaptic and neuronal level: information is transmitted in a discrete code (with a possible exception of timing) between discrete entities. At finer resolution thermal and chemical noise will be significant, suggesting that evolution would have promoted error correction and hence scale separation.

Brain-centeredness

A brain emulation would need to include at least some body and environment simulation. Bodily states are necessary for perception and action, since the brain's interaction with the environment is mediated by a body transducing between neural signals and sensory and motor signals. Bodily states also influence brain states directly and can contribute content (e.g. feelings of hunger triggered by hormones and stomach contractions). Hence some aspects of the body need to be part of the emulation framework. By the same token some environment for the body will have to be included.

The level of brain-centeredness of WBE can get away with is uncertain. Some of the more extreme interpretations of the extended mind hypothesis seem to require emulating not just a brain but a whole social and physical environment (or linking the emulation through a robotic body with the physical world). On the other hand, people with serious disabilities still exhibit minds and selves despite strongly constrained bodies.

The science and technology needed for accurate body models is likely to arrive well before WBE itself, especially since many of the physiological simulation and measurement methods may be necessary for developing WBE. Medical needs and entertainment (VR, realistic games) are likely to push realistic limits.

Technological issues

This group of issues deals with the technological feasibility of scanning brains and emulating them.

Simulation tractability

The challenge of simulation tractability is whether simulation at the level set by scale separation can be done on a realizable computer. This might be fundamental (if the brain components are doing uncomputable operations) or practical (there will not be enough computing power available in the future to achieve meaningful WBE). As argued above, no uncomputable operations have so far been observed to play a biological role. However, at present we are certainly unable to muster the computer power required for WBE: the real feasibility question is if and when such computer power becomes available.

One way of approaching this problem is to estimate available future computing power and compare it to estimates of brain emulation requirements (c.f. (Sandberg & Bostrom, 2008) p. 79-81). This produces a lower bound on when the technology might be available, since it is possible that the necessary interest, science or funding has not arrived at the time. While this might be of limited use for arguing in favor of the eventual feasibility of brain emulation, it allows bounds on earliest arrival times that might be relevant for risk or policy considerations.

Scanning tractability

A related issue is whether scanning methods for the necessary level of detail are realisable (or ethically acceptable).

Technologically there currently exist methods of imaging volumes of neural tissue at resolutions enough to discern the finest fibers (Hayworth, Kasthuri, Schalek, & Lichtman, 2006) and detecting chemical content at slightly lower resolution (Micheva & Smith, 2007). The main limitation is that the scan volume is very limited. Arguments for the feasibility of scaling this up to mouse-brain sizein the very near future have been made (http://www.brainpreservation.org/). If the required resolution is finer, for example involving molecular complexes or the exact genetic state of each cell, then the realisability becomes more uncertain.

Scanning brains to produce emulations will likely be a destructive process, and the research needed to bring brain emulation to a success criterion will most certainly involve running software that might have phenomenological states under conditions that are aversive. There might hence exist a hindrance due to research ethics to enabling brain emulation: the necessary experiments might be technologically possible, but would be unethical to perform because they involve excessive risk of suffering. However, ethical unfeasibility does not seem likely to prevent practical exploitation if the rewards are high enough.

Conclusions

WBE is a deeply challenging and long-term prospect. Given current neuroscientific and technological knowledge there doesn't seem to exist any fundamental obstacles, merely a large amount of engineering and research. Yet, extrapolations of technology and neuroscience are untrustworthy, especially given the possibility of foundational objections. While there doesn't seem to exist any convincing knock-down arguments within the philosophy of mind against WBE, part of the reason may be that the overall success criteria are relatively floating.

A problematic issue for the feasibility of WBE appears to be to bridge the high aims of structural validity with the limitation to just replicative validity. Development of new methodologies of testing and quality assurance are likely necessary.

In the near future the scale separation issue might provide a fruitful empirical way of testing the feasibility of WBE, with relevant implications in philosophy of mind and neuroscience. Attempts at achieving WBE may yield fruitful information about the way complex behavior and perhaps minds emerge from neural systems. This includes the roles of noise and analog signals, the interaction between systems on different scales, the epistemology of neuroscience and (in the case of a convincing success) evidence for or against some theories of mind.

References

Visual Transistor-level Simulation of the 6502 CPU. (2011, April 6). Retrieved 12 2, 2011, from http://www.visual6502.org

Ajay, S. M., & Bhalla, U. S. (2006). Synaptic plasticity in vitro and in silico: Insights into an intracellular signaling maze. *Physiology, 21*, 289-296.

Bebber, D. P., Marriott, F. H., Gaston, K. J., Harris, S. A., & Scotland, R. W. (2007). Predicting unknown species numbers using discovery curves. *Proc. R. Soc. B, 274*(1618), 1651-1658.

Bec, J., Cencini, M., Hillerbrand, R., & Turitsyn, K. (2008). Stochastic suspensions of heavy particles. *Physica D, 237*, 2037-2050.

Berglund, N. (2011). Stochastic dynamical systems in neuroscience. *Stochastic dynamical systems in neuroscience, Oberwolfach 2011.*

Bernal, J. (1929). *The World, the Flesh, the Devil: An Enquiry into the Future of the Three Enemies of the Rational Soul.*

Bhalla, U. S. (2004). Signaling in small subcellular volumes. II. Stochastic and diffusion effects on synaptic network properties. *Biophysical Journal, 87*, 745-753.

Caianello, E. (1961). Outline of a theory of thought processes and thinking machines. *Hournal of Theoretical Biology, 2*, 204-235.

Chalmers, D. (1995). Absent qualia, fading qualia, dancing qualia. In T. Metzinger (Ed.), *Conscious experience*. Imprint Academic.

Chalmers, D. (2010). The Singularity: A Philosophical Analysis.

Dreyfus, H. L. (1992). *What computers still can't do: a critique of artificial reason*. MIT Press.

Elliott, T. (2011). Stability Against Fluctuations: Scaling, Bifurcations, and Spontaneous Symmetry Breaking in Stochastic Models of Synaptic Plasticity. *Neural Computation, 23*(3), 674-734.

Faisal, A., Selen, L. P., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature reviews neuroscience, 9*, 292-303.

Hanson, R. (1994). If uploads come first: The crack of a future dawn. *Extropy, 6*.

Hanson, R. (2008). Economics of the singularity. *IEEE Spectrum*, 37-42.

Hartmann, S. (1996). The world as a process: simulations in the natural and social sciences. In S. Hegselmann (Ed.), *Simulation and modelling in the social sciences from the philosophy of science point of view* (pp. 77-100). Dordrecht: Kluwer.

Hayworth, K. J., Kasthuri, N., Schalek, R., & Lichtman, J. W. (2006). Automating the Collection of Ultrathin Serial Sections for Large Volume TEM Reconstructions. *Microscopy and Microanalysis, 12*, 86-87.

Hillerbrand, R. (2007). Scale separation as a condition for quantitative modelling. Why mathematics works for some problems and fails for others. *Models and Simulations 2*. Tilburg.

Houweling, A., & Brecht, M. (2008). Behavioural report of single neuron stimulation in somatosensory cortex. *Nature, 451*(7174), 65-8.

James, G., Silverman, B., & Silverman, B. (2010). Visualizing a classic CPU in action: the 6502. *Proceedings of SIGGRAPH Talks 2010*. New York: ACM.

Koch, C. (1999). *Biophysics of Computation: Information Processing in Single Neurons*. New York: Oxford University Press.

Li, C., Yu, J., & Liao, X. (2001). Chaos in a three-neuron hysteresis hopfield-type neural networks. *Physics Letters A, 285*, 368-372.

Merkle, R. (1989). *Large scale analysis of neural structures*. Palo Alto, California: Xerox Palo Alto Research Center.

Micheva, K. D., & Smith, S. J. (2007). Array tomography: A new tool for imaging the molecular architecture and ultrastructure of neural circuits. *Meuron, 55*, 25-36.

Moravec, H. (1988). *Mind children: The future of robot and human intelligence*. Harvard University Press.

Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press.

Parker, A. C., Friesz, A. K., & Pakdaman, A. (2006). Towards a Nanoscale Artificial Cortex. *The 2006 International Conference on Computing in Nanotechnology (CNAN'06), June 26-29, 2006*. Las Vegas, USA.

Penrose, R. (1989). *The emperor's new mind*. New York: Oxford University Press.

Sandberg, A., & Bostrom, N. (2008). *Whole brain emulation: a roadmap*. Oxford: Future of Humanity Institute, Oxford University.

Searle, J. (1980). Minds, Brains and Programs. *Behavioral and Brain Sciences, 3*(3), 417-457.

Shores, C. (2011). Misbehaving Machines: The Emulated Brains of Transhumanist Dreams. *Journal of Evolution and Technology, 22*(1), 10-22.

Thompson, E., Varela, F., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, Massachusetts: MIT Press.

Zeigler, B. (1985). *Theory of Modelling and Simulation*. Malabar: Krieger.

Zeigler, P. B., Praehofer, H., & Kim, T. (2000). *Theory of modeling and simulation: integrating discrete event and continuous complex dynamics systems*. Academic Press.