Artificial Reward and Punishment

Grounding Artificial Intelligence through motivated learning inspired by biology and the inherent consequences for the Philosophy of Artificial Intelligence.

Master thesis History and Philosophy of Science,
Utrecht University
Nathan Perdijk
0473170
52.5 ECTS
Supervisor: J.J.C. Meyer
Second reviewer: J.H. van Lith

Artificial Reward and Punishment

Grounding Artificial Intelligence through motivated learning inspired by biology and the inherent consequences for the Philosophy of Artificial Intelligence.

By: Nathan R.B. Perdijk

Cover illustration

"Punishment and Reward"
Custom image provided by Izak Perdijk
http://pinkoliphant.deviantart.com/

Table of Contents

Introduction

Ever since the dawn of computer science, leading computing experts have wondered if it is possible to build a machine intelligence equal to humans. Computation experts and science fiction writers alike have since speculated what this "Artificial Intelligence" (AI) would look like and how a logical machine, free of our base emotions, would operate. Some even contemplated whether a far enough advanced machine intelligence, achieved by a high computations per second limit and sophisticated algorithms, would grant machines self-awareness or emotions, enabling and motivating the computer to undertake actions other than those it was strictly programmed to do. Ever since then, talented programmers have set out to build logically operating systems to mimic human intelligence and philosophers of Artificial Intelligence have wondered what the status of these reasoning machines should be. In their review of the computer science literature, most philosophers have paid particular attention to those machines equipped with software that reasoned based on logic, while largely disregarding the learning computer programs.

There is a strange dissonance here. In the development of these systems, a large focus has been placed on replicating the surface reasoning tasks that humans are famously capable off. Whether it concerns playing chess (Deep Blue),[1,2] *Jeopardy* (Watson),[3] or solving a puzzle to find out who the murderer was given a particular list of statements (General Problem Solver),[4] or even carrying out a conversation such as Eliza or Parry,[5] attempts to create intelligence, and philosophical arguments surrounding those attempts, have often focussed on giving it tasks we strongly associate with logical reasoning or the straightforward computation of all possible outcomes. Due to a wide variety of reasons, human cognitive reasoning has been taken as the part of intelligence that required duplication for a functional and intelligent AI, with little or no attention to the foundation of mental faculties that human intellect has been built on and, amongst philosophers of AI at least, a general disregard for any learning capacity. This is, in my view, wrong. If we truly wish to create a *real* Artificial Intelligence comparable to ours, we cannot

---

[1] Fine, J. (1997). Deep Blue wins in final game of match; Chess computer beats Kasparov, world's best human player. http://faculty.georgetown.edu/bassr/511/projects/letham/final/chess.htm. MSNBC (retrieved 6 June 2014).

[2] IBM 100 (2011). Icons of Progress; Deep Blue Overview. http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/. *IBM* (retrieved 16 May 2014).

[3] Jackson, J. (16 February 2011). IBM Watson vanquishes human Jeopardy foes. http://www.pcworld.com/article/219893/ibm_watson_vanquishes_human_jeopardy_foes.html. *PCWorld* (retrieved 15 May 2014).

[4] Copeland, B.J. (1993), *Artificial Intelligence: A philosophical introduction* (Oxford, 1993). 24-26.

[5] Copeland, B.J. (1993), *Artificial Intelligence: A philosophical introduction* (Oxford, 1993). 12-15.

disregard the very foundations upon which our own intelligence has been constructed, nor can we leave the capability to learn out of the philosophical debate. If a true AI is to be created, one recognisable as intelligent and capable of a wide variety of tasks, it is important that the research involved pays much closer attention to the more natural, biological foundations upon which human intelligence is built and it should therefore include not just out reasoning capabilities, but our learning abilities and an eye for our motivations as well. In short, those feelings and emotions that were sometimes supposed to arise from a high enough reasoning intelligence, should instead be included from the get go.

In the rise of organic intelligence, the brain structures that govern emotions developed first,[6] before the cerebral cortex which handles rigorous reasoning, logic, and which is also responsible for supporting the arithmetic that is the primary *raison d'être* for digital computers. Not only are those ancient, more primitive brain structures still present in the human brain, they also still have a major impact on our behaviour.[7,8,9,10,11] Starting at the cortical, analytical functionality of the brain disregards the entire foundation that our logical reasoning is built on. This is one of the reasons that AI's, especially those under the review of philosophers, often seem to lack common sense.[12] Scientific evidence suggests that humans do not so much arrive at their conclusions through applying rigorous logic, but instead apply a whole range of intuition-driven processes that involve past experiences and even current moods. These processes are based on motivational learning that ties external events to internal consequences. It is this web of connections that also supports logical reasoning, but logical reasoning is often not involved in decision-making. Of course, when you ask a human why they have made a particular decision, they will often invoke logical explanations as a justification for their decisions, even when logic has had little to do with it.[13] This sparks the question:

---

[6] Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions* (New York, 1998).

[7] Decety, J. & Svetlova, M. (2012). Putting together phylogenetic and ontogenetic perspectives on empathy. *Developmental Cognitive Neuroscience 2* (1) 1-24.

[8] Bos, P.A., Panksepp, J., Bluthé, R.M. & Van Honk, J. (2012). Acute effects of steroid hormones and neuropeptides on human social-emotional behavior: A review of single administration studies. *Frontiers in Neuroendocrinology 33* (1) 17-35.

[9] Moscarello, J.M. & LeDoux, J.E. (2013). The contribution of the amygdala to aversive and appetitive Pavlovian processes. *Emotion Review 5* (3) 248-253.

[10] Rilling, J.K. (2013). The neural and hormonal bases of human parental care. *Neuropsychologia 51* (4) 731-747.

[11] Eisenberger, N.I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience 13* (6) 421-434.

[12] McCarthy, J. (2007). From here to human-level AI. *Artificial Intelligence 171* (18) 1174-1182.

[13] Nisbett, R.E. & DeCamp Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review 84* (3) 231-259.

What role do non-reasoning, motivational systems play in human intelligence?

Much of the literature in Artificial Intelligence and the Philosophy of Artificial Intelligence is nonetheless focussed on reasoning through logic, and the corresponding brain structure: the cortex. This can perhaps be explained by a belief in human exceptionalism that stems from old philosophical considerations:[14] humans are taken as the archetype of intelligent creatures, so the emphasis goes to those structures and functions that separate humans from animals, with the enlarged cerebral cortex and its logical-reasoning capabilities being the obvious candidate. Such a focus on what makes humans special can provide an obvious blind spot for the factors that humans have in common with other animals that yet may still be fundamental to intelligence, because only the differences receive much attention. Another possible explanation is inherent to the tools available, namely computing itself: computers were designed to process arithmetic calculations and arithmetic and logic mesh very well. It is simply attractive to focus on the kind of tasks that the computer seems most capable of handling: applying logic rules in a rigorous fashion. This is more likely to produce quick and clean results, even if it may turn out to be a glorious dead end when all is said and done.

As was already mentioned, in many cases the focus in the philosophy of AI has been on computers replicating human behaviour and reasoning, while less attention has gone into those AI that learn. This is unfortunate, as learning is a key element of intelligence. A non-learning machine may be able to display intelligent behaviour in the situation it was programmed to perform in, but when that situation changes, that same intelligent behaviour can suddenly be quite dumb revealing the apparent intelligence to be a farce instead. This has often let to the conclusion that AI in general are incapable of humanlike intelligence. This is hardly fair, as the AI's on review have indeed been inadequate in that respect and therefore shouldn't function as the basis for pan-AI philosophical judgements. In my mind, this also means that any Artificial Intelligence worthy of the name needs to have the ability to learn. Still, even in learning AI, the pattern of a strong focus on logical rules emerges. A fair amount of computer programs have been given tools to learn, but most of these have functioned purely in relation to the outside world and the adjustment of learning mechanics generally comes from an external source. A fundamental aspect of human learning is therefore missing: motivation.

---

[14] Williams, J. (2007). Thinking as natural: Another look at human exemptionalism. *Society for human ecology 12* (2) 130-139.

Regrettably, the focus on reasoning and deduction through logic in the literature has left little room for questions concerning the non-cortical basis of learning. In my opinion, leaving out fundamental parts of how the brain learns and decides, results in a machine “intelligence” that is perhaps capable of impressive feats previously only performed by intelligent life forms, such as being able to play chess, but which is nonetheless hard to recognize as intelligent. Learning is a fundamental aspect of intelligence, yet learning is still often one of the most artificial aspects of Artificial Intelligence. AI's assign values for different situations and then adjust these values based on the end result. While this may superficially be somewhat like humans, AI do not learn because they are driven or motivated to do so, but rather because they have been hardcoded to adjust these values. They do not change these values due to some internal consequence, but rather the adjustment of these values is the only internal consequence. Although an AI with weighted functions already appears considerably more intelligent and natural in its behaviour, the disregard for the underlying structures that motivate the brain, leaving out the purpose and drive behind human learning, renders machine learning still a very artificial product with a less than natural feel. An AI produced without an integrated punishment and reward system is in my opinion not a “human” AI, but an “alien” AI. This leads to the follow-up question:

> Given modern insight into the role of punishment and reward-systems in biological intelligence, what contribution can motivational systems make to Artificial Intelligence from both a practical and a philosophical perspective?

The answer to these questions will hopefully open the door for a more “natural” Artificial Intelligence. During the past few decades, a larger emphasis has been placed on the development of neural nets. Neural nets are an abstraction of the neuronal networks that make up the brain. Their parallel processing has unlocked new ways of storing information: rather than storing a string of symbols to a fixed memory location, in Neural Nets the information is somehow stored in the connection strengths between the “Neurons”.[15] Although these Neural Nets are often simulated through an old fashioned serial processing computer, they could be produced in practice. In order to learn from a dataset, a second computer changes the connection-strengths between Neurons at random in an attempt to produce a predetermined desired response to a set of training examples.

---

[15] Henceforth I will use the noncapitalised “neuron” for the biological range of cell types, while I will use the capitalised “Neuron” for artificial, binary neurons.

Once the neural net has been instructed on this set of examples, it can be put to the test on untrained samples to see if it has learned the right thing. A lot of the aspects of human data-retention are mimicked in a superior way by Neural Nets when compared to old fashioned serial logic AI.[16] However, the learning mechanism itself is external to the trained Neural Net and also has no motivational component. Integrating a motivational system into this new Neural Net form of AI may serve to make it even more human-like. This leads to the final shape of my Research Question:

Research Question

In what manner can biological reward and punishment systems be integrated in the Neural Net-approach to creating Artificial Intelligence with humanlike learning and recognisable intelligence? What are the consequences of such a Natural AI for the field of Philosophy of AI?

Thesis

I will argue in this paper that intelligence is simply a compounded form of basic adaptability: being adaptable is what intelligence is all about. Adaptability itself can be broken down into its constituent parts: interaction, evaluation, storage and action adjustment. Adaptability is greatly benefitted by introducing punishment and reward-systems as these provide the necessary information to attach value to stored information. I will argue that human reward and punishment is based on homeostatic monitoring, which is itself grounded in the death and survival consequences inherent in natural selection. It is the interplay of punishment and reward-systems with homeostasis which allows for the successful creation of meaningful connections between outside stimuli and inside consequences. These connections allow for meaningful storage, provide meaning to interactions and in consequence allow successful and meaningful action adjustment beyond directly programmed responses. It is upon this foundation of internal consequence and therefore motivation that human intelligence, capable of higher level evaluation and decision-making is built. Integrating a similar system into the Neural Net approach, which by its nature strongly favours connections, will eliminate some of the more pressing philosophical issues with mechanical intelligences being deemed “unnatural” or “ungrounded” and therefore not truly intelligent. Motivation in the human brain works at a cellular, and therefore neuronal level, which allows for a pretty much seamless integration of motivational reinforcement learning as a training mechanism that regulates

---

[16] Copeland, B.J. (1993), *Artificial Intelligence: A philosophical introduction* (Oxford, 1993).

connection strengths in Neural Nets. This may provide researchers in the field of AI with new ways to implement self-learning programs and opens up new avenues of flexible AI instruction. The Philosophy of AI will also need to adjust to motivational learning in Artificial Intelligence. I will argue that rejection of Artificial Intelligence as potentially meaningful on the basis of its eventually symbolic nature does not hold, because adaptability, and with it intelligence, is not about the matter that composes it, physical or digital, but about the valued connections and consequences that can be supported. This provides an interesting answer to the Chinese Room Argument, the Simulation Objection and the Symbol Grounding Problem.

The goal

The goal of this paper is twofold. I aim to provide a list of desiderata for a recognisably adaptable and learning system based on intrinsic motivation, inspired by biology, and linking internal consequences to external factors. Coincidental with this goal is the effective removal of some of Philosophy of AI's greatest objections to Artificial Intelligence: the Simulation Argument and the Symbol Grounding Problem.

The argument

The argument will be built along the following structure:

Before moving to the inner workings of motivational systems, I will first give a brief introduction of what I mean by developing “Artificial Intelligence” in Chapter 1, namely the quest for Strong AI. While exploring what I mean by “Artificial Intelligence” I will also explore what I believe constitutes “intelligence”, for which I'll introduce a strongly reduced variant of Jack Copeland's “massive adaptability” which I call bare-bone adaptability. I will argue that biological intelligence is completely reliant on valued learning and that learning is the cornerstone of being adaptable. Through this argument I hope to establish that intelligence is a continuum of greater or smaller adaptability and that even the most primitive life form has something that can be equated to a very rudimentary intelligence, which allows it to adjust to changing circumstances and undertake actions beneficial to its survival. It is from this most rudimentary form of adaptability that I think human intelligence eventually stems.

After exploring these issues, I will delve into the biological (Chapter 2: Adaptability without a Brain) and neuroscientific (Chapter 3: Reward and the Brain) background of biological intelligence by illuminating the role evolution has played in the formation of basic learning mechanics in unicellular life, as well as the emergence of higher learning mechanisms in complex life forms such as humans. In concert with that, I will shortly delve into the role of hormones, the central nervous system and the brain. In this chapter on adaptability without the brain, I will first evaluate an important survival mechanism, homeostasis, and the impact it has on another important survival mechanism: the sensory apparatus of the bacteria that allows it to evaluate its environment. These two mechanisms combined have contributed greatly to adaptability. I will then show how the transition from single-celled organisms to multi-cellularity can maintain this interplay of homeostasis and sensory evaluation. It turns out that at the cellular level, organisms are capable of distinguishing good from bad, an important motivational tool for guiding their behaviour, and an ability that is later reused in communication between the body, the brain and individual brain-cells. In the chapter on the relationship between reward and the brain (Chapter 3), I will reveal the mechanisms at work in motivational learning at the macroscopic level. I will illuminate the function of reward-systems in the brain, as well as their hormonal basis, as an explanation for much of our natural learning processes, which often are neither strict reasoning nor even conscious. To illustrate this non-explicit-rule-following learning mechanism, I will illuminate some of the mechanisms through which humans learn in everyday life, in particular subconscious learning. Afterwards, the motivational connection between conscious learning and reasoning, and subconscious and emotional thought processes will be revealed.

After these two more biological chapters, I will draft up the rough schematics for a self-motivating “feeling” Artificial Intelligence. In Chapter 4 I will abstract the biological principles that underlie motivational learning to a level where they could be used in constructing an artificial motivational mechanism with a more natural feel. This first rough draft, called Motivated AI (MAI), will only have one homeostatic value to take into account, but the positive and negative associations that are derived from it will be fully instated allowing it to attach value to its interactions without outside help. Naturally, the proposed model will not be perfect yet, but it will illustrate some of the advantages that come from implementing motivational learning into AI. The objective of this chapter, which is strongly rooted in Chapters 2 and 3, is to create a list of desiderata for an AI system that incorporates a more natural learning mechanism. These requirements for the system can then function as a basis for further AI research, as well as opening up new avenues for philosophical research into the implications of having motivational-learning AI.

The consequences

In Chapter 5 I will then explore what philosophical objections may oppose calling a sufficiently advanced AI with inbuilt punishment and reward systems “intelligent”. I will argue that punishment and reward-systems help create an internal meaning which is grounded in internal and external reality. This position should assist in circumventing the symbol grounding problem, although an evolutionary mechanism that weeds out poor internal representation-to-reality correspondence is still required. I will also go over the physical differences between computer hardware and the lower levels of software and the “hardware and software” that make up biological intelligence. Rather than equating the computer hardware and software to biological wetware, I will argue that computer hardware and software is actually much more akin to physical forces, particles or DNA. Accepting this argument will allow for a serially processed digital computer to run a digital version of a parallel Neural Net without losing any philosophical credence. My argument will be that it is adaptability, the ability to learn from new situations and adjust behaviour accordingly and adequately is both the foundation and the distinguishing ability of any being or thing that has any claim to being called “intelligent”. Upon what this adaptability is constructed is of no real relevance, as long as the internal consequences are real.

Before I get to my concluding remarks, I will review three modern attempts at exploring human-like Artificial Intelligence. One, called Soar, is a very ambitious, symbolic, top-down and explicitly rule-driven cognitive model. The other two, called RMCLS and AuGMEnT are potential lay-outs for Artificial Neural Networks that use the broadcasting of a Global Reward Signal to modulate connection strengths between Neurons and thereby train them. Though much more limited in current design aspirations, the creators of these bottom-up Networks hope to get to the essence of biological reward-learning. I will review this small sample of recent endeavours in AI practice for possible overlaps and interconnections with the proposed MAI model. In essence, I will give a short summary of each AI project and the ways in which the new proposed model could positively impact their learning, adaptability and philosophical foundation.

Methodology.

This thesis will be based on a comprehensive, interdisciplinary study of the literature in the biological and neuroscientific fields as well as forays into the fields of Artificial Intelligence and the Philosophy of Artificial Intelligence. An important focus has been placed on review articles and books covering the subjects. Books and articles have been selected on the basis of relevance to the subject and academic quality.

Literature

Bos, P.A., Panksepp, J., Bluthé, R.M. & Van Honk, J. (2012). Acute effects of steroid hormones and neuropeptides on human social-emotional behavior: A review of single administration studies. *Frontiers in Neuroendocrinology 33* (1) 17-35.

Copeland, B.J. (1993). *Artificial Intelligence: A philosophical introduction* (Oxford, 1993).

Decety, J. & Svetlova, M. (2012). Putting together phylogenetic and ontogenetic perspectives on empathy. *Developmental Cognitive Neuroscience 2* (1) 1-24.

Eisenberger, N.I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience 13* (6) 421-434.

Fine, J. (1997). Deep Blue wins in final game of match; Chess computer beats Kasparov, world's best human player. http://faculty.georgetown.edu/bassr/511/projects/letham/final/chess.htm. MSNBC (retrieved 6 June 2014).

IBM 100 (2011). Icons of Progress; Deep Blue Overview. http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/. *IBM* (retrieved 16 May 2014).

Jackson, J. (16 February 2011). IBM Watson vanquishes human Jeopardy foes. http://www.pcworld.com/article/219893/ibm_watson_vanquishes_human_jeopardy_foes.html. *PCWorld* (retrieved 15 May 2014).

McCarthy, J. (2007). From here to human-level AI. *Artificial Intelligence 171* (18) 1174-1182.

Moscarello, J.M. & LeDoux, J.E. (2013). The contribution of the amygdala to aversive and appetitive Pavlovian processes. *Emotion Review 5* (3) 248-253.

Nisbett, R.E. & DeCamp Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review 84* (3) 231-259.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions.* (New York, 1998).

Rilling, J.K. (2013). The neural and hormonal bases of human parental care. *Neuropsychologia 51* (4) 731-747.

Williams, J. (2007). Thinking as natural: Another look at human exemptionalism. *Society for human ecology 12* (2) 130-139.

Chapter 1: Defining Artificial Intelligence

Programs such as Eliza, Parry, Shrdlu, the General Problem Solver (GPS), Sam[17] and the most famous computer conquerors of human champions Deep Blue[18,19,20] and Watson,[21] have achieved impressive levels of performance thanks to cleverly written algorithms and, quite often, brute computing power. They are illustrations of Artificial "Intelligence" that performs pretty well under the given circumstances. There is, however, something wrong with regarding these success stories as pinnacles of Artificial Intelligence in the way that I will be using the term: they are not really about intelligence and are instead about clever programming.

Watson is the most recent case of a computer coming out ahead in a match of man versus machine. In this case, the playing field in question was Jeopardy, a well-known game show contest of knowledge in which contestants answer a wide variety of trivia questions. Watson was built by IBM and equipped with a very large database and software capable of interpreting human sentences and their context-sensitive nature in order to excel at answering these trivia questions. Thanks to over 200 million pages of content and 6 million logic rules, Watson was often capable of producing the right answer and when pitted against human competition, it proved capable of doing so at a higher speed than his human opponents. Although it started out a bit slow, during the match its lead grew steadily and it absolutely trounced the human opposition. Apparently its logic rules and enormous database had made it more adept at producing the required information than the Jeopardy champions it was playing. And yet it still made a few obvious and stupid mistakes, such as offering "Toronto", a Canadian city, as the answer to a question asking for the largest US airport named after a World War II hero.[22] For a supposedly intelligent program, Watson had made a very silly mistake, as the correct answer was most definitely part of its database, as was the information that Toronto is

---

[17] Copeland, B.J. (1993). *Artificial Intelligence: A philosophical introduction* (Oxford, 1993).

[18] IBM 100 (2011). Icons of Progress; Deep Blue Overview. http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/. *IBM* (retrieved 16 May 2014).

[19] Fine, J. (1997). Deep Blue wins in final game of match; Chess computer beats Kasparov, world's best human player. http://faculty.georgetown.edu/bassr/511/projects/letham/final/chess.htm. MSNBC (retrieved 6 June 2014).

[20] Russell, S. & Norvig, P. (2010). *Artificial Intelligence A Modern Approach: Third Edition* (New Jersey, 2010) 29.

[21] Jackson, J. (16 February 2011). IBM Watson vanquishes human Jeopardy foes. http://www.pcworld.com/article/219893/ibm_watson_vanquishes_human_jeopardy_foes.html. *PCWorld* (retrieved 15 May 2014).

[22] Jackson, J. (16 February 2011). IBM Watson vanquishes human Jeopardy foes. http://www.pcworld.com/article/219893/ibm_watson_vanquishes_human_jeopardy_foes.html. *PCWorld* (retrieved 15 May 2014).

a Canadian city and not named after a World War II hero. Of course, the crux here is that Watson isn't really intelligent in the way that humans are, instead it has sophisticated algorithms that let it select a word from its database that is the most likely “answer” to the input it receives. Without knowing what either the question or the answer truly is or even knowing what a question truly is, Watson provides output that we then recognise as a (correct) “answer”. In many ways Watson is a quintessential Chinese Room (see Chapter 5), which explains its impressive answering capabilities as well as its otherwise puzzling stumble.

Watson is not alone in combining startling competence with shocking strike-outs. Eliza, Parry, Shrdlu, the General Problem Solver and Sam could all be stumped when their assignment was slightly outside of their capabilities. They lacked “common sense”, had no understanding of their programmed task and could be tricked into clearly displaying their lack of actual comprehension. Deep Blue was a very impressive chess computer, but it couldn't do anything else, nor did it really know what chess was: it just produced output based on input it received combined with extensive “training” that favoured certain outputs over others in certain situations. Watson was basically a very strong search engine, combined with natural language interpreting algorithms and a limitation to providing just one answer. It had no abilities outside its highly specialised purpose although its specialised purpose does allow for repurposing in other fields of data retrieval: Watson has been put to use in the field of medicine as an advisor to medical professionals.[23] From these examples, it appears obvious that programming millions of logic rules, coupled to an enormous database still leaves a lot to be desired for producing a well-rounded, truly intelligent AI. The examples of AI triumph mentioned above are simply not qualified for that title, but before we can get to the question of what would be required for a truly intelligent AI, it is perhaps fitting to first specify what I mean when I speak of Artificial Intelligence.

What is Artificial?

Artificial Intelligence is a broad term with a variety of meanings. The term naturally falls apart into two: “Artificial” and “Intelligence”. The word “Artificial” has a variety of meanings in and of itself, but for *Artificial Intelligence* two particular branches of meaning are of special importance:

---

[23] Upbin, B. (8 February 2013). IBM's Watson Gets Its First Piece Of Business In Healthcare. http://www.forbes.com/sites/bruceupbin/2013/02/08/ibms-watson-gets-its-first-piece-of-business-in-healthcare/. *Forbes* (retrieved 27 August 2014).

Artificial

1. made by human skill; produced by humans (opposed to natural): artificial flowers.
2. imitation; simulated; sham: artificial vanilla flavoring.[24]

The first branch of meaning, “produced by humans”, is pretty straightforward as long as we eliminate any natural reproduction mechanisms as a qualification for the “produced by humans”-clause. The field of Artificial Intelligence strives to create something intelligent through construction, rather than through the obvious “natural” means, i.e. giving birth. Anything “artificial” is “manufactured”, rather than “grown”. The second meaning is one that actually sparks a fair bit of debate among Philosophers of AI. Is any “AI” we manage to create an imitation, a simulation or a sham? After all, in creating an intelligence we seek to create one we would recognise as intelligent, otherwise there would be no way for us to know whether we succeeded. In order for any AI to pass that test, it must imitate at least some forms of human intelligence.

On the basis of this assumption and to do away with any tricky definitional questions rewarding the word “intelligence”, Alan Turing, an important founder of the AI field, proposed a test in 1950 where a computer actually plays an imitation-game in a natural language test.[25] The Turing Test directly measures AI performance in a human intelligence skill (namely human natural language and acting like a human, albeit in typewriting). If the AI is capable of consistently fooling a human investigator into thinking it is human, then it must be considered intelligent according to Turing. However, as natural language processing is but one of many shapes and forms of recognisable intelligence, even a machine that passes the Turing Test can be argued to not be truly intelligent, depending on its make-up. An example of this kind of reasoning can be found in John Searle's famous Chinese Room Argument, where he describes a computer that can give the perfect natural language answers without actually understanding what it is saying. According to Searle, the Chinese Room's apparent intelligence is actually a sham. Other debates focus more on the question whether a simulation of intelligence should be regarded as intelligent or not and how the symbols used by an AI can gain any intrinsic meaning or “grounding” for that AI. However, these philosophical issues are for a later moment (see Chapter 5), we are now only just determining what

---

[24] Dictionary.com. http://dictionary.reference.com/browse/artificial?s=t. *Dictionary.com* (retrieved 14 January 2014).

[25] Turing, A.M. (1950). Computing Machinery and Intelligence. *Mind 59* (236) 433-460.

is meant by “Artificial”, for which I will now give the following definition:

> *“Artificial Intelligence is an Intelligence created not through natural means, such as natural or semi-natural biological procreation, but rather through manufacture. Some uncertainty of whether this is actually possible, or that perhaps the best we can hope for is nothing more than a 'simulation', is already contained in the term, although not always expressed.”*

Although the above definition of artificial is still reasonably broad, for all practical intents and purposes AI's are nearly always conceived of as computers with or without a robot body. Although this thesis will review the importance of several key learning mechanisms for creating a proper Artificial Intelligence, it is useful to bear in mind that a digital computer is currently both the most popular and the most likely candidate for implementing these and any comments will be made keeping a computer framework in mind.

Before we explore the depths of the meaning of the word intelligence, a short look at the field of AI is in order. According to Russell and Norvig, the field of Artificial Intelligence is roughly divisible along the lines of at least four different sets of definitions for the developmental goals of AI research.[26]

- AI Systems with rational thought, where designs focus on determining the best possible outcome through applying strict and rigorous logic, regardless of the actual human method used. This approach is best conceived of as the “logic AI” approach.
- AI Systems with humanlike actions, where the designs emulate human behaviour, although the underlying processes that cause it do not need to be the same. This field can be called “social” or “mimicking AI”.
- AI Systems with rational actions, which strive for achieving the best possible outcome, even when there is no rational thought that determines the course of action. Systems like these acknowledge that even in situations where the outcome of actions cannot be

---

[26] Russell, S. & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach; Third Edition* (New Jersey, 2010) 2.

foreseen, undertaking any action can still be more rational than inaction, even if unfounded. Humans are a source of inspiration, but not a criterion by which the validity of a particular trick is measured. This field can be considered "practical AI",

and finally:

- AI Systems with humanlike thinking, where the designs focus on emulating human thinking processes and are rather averse to information-technology shortcuts that produce the same outcomes through entirely different means. This approach is often denoted as "natural AI".

It is this last sub-discipline of AI that carries the interest of this thesis.

Although there is some implicit appeal to learning as a part of intelligence in these set definitions, each of these descriptions focusses more on thought-processes and actions, rather than learning, perhaps presupposing it in "thinking". Regardless, all of these fields make use of the field of Machine Learning, which is concerned with constructing computer programs that automatically improve through experience.[27] Programs, in short, that learn and can apply this knowledge.

Artificial Intelligence today knows many practical applications, ranging from Google's useful search and auto-completion algorithms[28] to more niche systems, such as diagnostic tools in medicine[29]. For these applications of Artificial Intelligence, the restraints are not very severe: as long as any particular program achieves the intended result without wasting too many resources, the underlying process that guides this "intelligence" does not need to adhere to strict rules as to what actually qualifies as intelligent. These are specialised intelligences and the possibility of making such dedicated, intelligent programs is called "Weak AI", although a more gracious name would be "Expert Programs". These programs are masters in their own field, but outside of their very narrow band of expertise, they completely break down. None of these programs would be able to survive in a natural world, or even be able to conceive of one, as the tools in their arsenal are simply unsuitable and they were never designed with that purpose in mind.

However, from the dawn of Artificial Intelligence a different dream has pervaded the field,

---

[27] Mitchell, T.M. (1997). *Machine Learning* (Boston, 1997) xv.

[28] Mikolov, T., Sutskever I. & Quoc, L. (15 August 2013). *Learning the meaning behind words*. http://google-opensource.blogspot.nl/2013/08/learning-meaning-behind-words.html. *Google Knowledge* (retrieved 24 March 2014).

[29] Agah, A. (ed.) (2014). *Medical Applications of Artificial Intelligence* (Boca Raton, 2014).

pursued vigorously by some while disregarded as preposterous by others. This dream is that perhaps, one day, humans will create a computer with suitable programming that will equal us, or even surpass us, in all-round, actual intelligence. Through sheer computer power and human ingenuity, perhaps a computer can be made that truly thinks! This belief is what is called "Strong AI" and this paper seeks to contribute to this Strong AI-thesis. Before that is possible, however, we must first discuss the meaning of the word "intelligence".

What is Intelligence?

Intelligence is a very strange, and hard to define category. In philosophy it is traditionally placed under the "mental" aspects and the Philosophy of Mind, and away from the physical, mechanical workings of our bodies, a distinction that can be traced back to at least René Descartes in the seventeenth century. According to Descartes, the human essence can be split into the divisible, material and mechanical body and the indivisible, immaterial mind.[30] During the twentieth century, quite a few philosophers have argued against this separation of the "mental" and the "physical", better known as *dualism*, and have instead insisted that mental states can be reduced to physical phenomena.[31] The philosophical debate surrounding what should replace dualism is very interesting and quite complicated, but it is outside the scope of this thesis. I will instead lay out my basic assumption on this matter right now:

> *"Intelligence should not be regarded as a strictly "mental" quality. The physical state of the brain, and in fact the physical state of the body, has an inseparable impact on intelligence. There are many bodily processes that influence intellectual activity and it is wrong to try to understand intelligence in a purely mental frame of reference. Trying to do so closes doors that should remain open. Intelligence is firmly rooted, embodied if you will, in the hardware it resides in. Its sole biological function is to keep that body alive and, to this end, that body is an integral part of that intelligence."*

---

[30] Robinson, H. (2012). Dualism. http://plato.stanford.edu/archives/win2012/entries/dualism. In: Zalta, E.N. (ed.). *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition) (retrieved 24 March 2014).

[31] Searle, J. (2004). *Mind: A brief introduction* (Oxford, 2004) 47-81.

The above does not signal, however, any kind of human or biological exceptionalism: just because the biological body is a part of intelligence in us, it does not mean that there is only one kind of body that can support intelligence. I hope to display in the rest of this thesis why intelligence being embodied is a crucial factor in making it recognisable. In the meanwhile, we should seek to avoid any philosophical or popular preconceived notions and assumptions regarding intelligence. Perhaps it is more useful to replace the term “intelligence” with something that is less culturally laden. A good start, in my opinion, would be to swap the quest for intelligence out for a quest for Jack Copeland’s “being Massively Adaptable”. Copeland describes it as such:

> *“An organism's inner processes are massively adaptable if it can do such things as form plans, analyse situations, deliberate, reason, exploit analogies, revise beliefs in the light of experience, weigh up conflicting interests, formulate hypotheses and match them against evidence, make reasonable decisions on the basis of imperfect information, and so forth. Moreover it must be able to do these things in the boisterous complexity of the real world – not merely in a 'toy' domain such as Shrdlu's simple world of coloured blocks.”*[32]

Implicit in this definition, but not explicitly mentioned because we take them for granted, are three very vital parts of intelligence and adaptability as we know it: the ability to **interact** with the outside world (whatever that may be, although Copeland requires it to be the “real world”), the ability to **remember** what is important and why, and the ability to **adjust** actions on the basis of current interactions and earlier recollections. In fact, the above definition flows from compounding these three basic requirements. So what we really need for a bare-bone definition of adaptability is:

- A “being” must be capable of **interacting** with its environment (requiring some form of perception and some means of altering itself or the environment),
- A “being” must be capable of **storing** these interactions/perceptions, (more commonly known as having a “memory”),
- A “being” must be capable of **adjusting** its interactions based on previous

---

[32] Copeland, B.J. (1993). *Artificial Intelligence: A philosophical introduction* (Malden 1993), 55.

interactions/perceptions (more colloquially known as "learning").

These three requirements interact and overlap, allowing for complex and changeable behaviour. Still, they are not exhaustive of even bare-bone adaptability. After all, even if a being is capable of interacting with the environment and storing what happens, how does it determine what to adjust? How does it determine what is important and on what basis does it do so? In the above definition there is no call for establishing relative value, but relative value is necessary in determining the right choice or even in establishing what is important. A being must not only be capable of interacting with the environment and be able to store that information, but it must also be capable of evaluation: it needs some way to establish which interactions have been "beneficial" and which interactions have been "detrimental". A being, even one that is only bare-bones adaptable, needs to be able to store valued information. In other words, if a being must be capable of adjusting to its environment, the interactions and perceptions it attains need to be stored in a meaningful[33] way. This brings us to my final definition of bare-bone adaptability:

- A "being" must be capable of **interaction** with its environment (requiring some form of perception and some means of altering itself or the environment),
- A "being" must be capable of **evaluating** its interactions with the environment,
- A "being" must be capable of **storing** these valued interactions, (more commonly known as having a "memory"),
- A "being" must be capable of **adjusting** its interactions **based on these values** attained through previous interactions/perceptions (more colloquially known as "learning").

As massive adaptability has just been equated to full-blown intelligence, I will call this bare-bone adaptability a form of bare-bone intelligence. I will argue that even our highest-level reasoning skills eventually flow from these four basic prerequisites and that they do not exist in a vacuum. Instead, high-level reasoning can emerge in any organism sufficiently competent at the four given tasks. While "storing" and "adjusting" may be considered more important to our general ill-defined notion of "intelligence", adequately valued perception of and interaction with the environment is in

---

[33] "Meaningful" in this case refers to making connections between actions and their consequences.

fact a prerequisite of the other two. No useful learning is possible without detection of, and interaction with the environment (no matter how broadly or narrowly it is defined) and a means of evaluating that contact as no possible consequence or need for adjustment can otherwise be discerned. In the rest of this thesis, I will pay particular attention to the evaluation that is integral to bare-bone intelligence and learning, but so often overlooked. In order to find methods for implementing this in AI and the Philosophy of AI, I will explore how biological organisms evaluate their interactions, from the bare-bone intelligence found in simple organisms, to the complex intelligence found in humans.

But why even bother with determining, or creating, bare-bone intelligence, which enables behaviour which most people would not describe as intelligent, such as the actions of ants, or even of single-celled organisms, when it is higher intelligence that we're after? Because the distinction what is generally considered intelligent and what is not, is a lot like the distinction between what is a hill and what is a mountain. While a hill is definitely not a mountain, you cannot create a mountain without creating a hill first and this is precisely what is generally overlooked when discussing Artificial Intelligence: it is attempted to create and review an “intelligence” without the bare-bone that, in my opinion at least, is the backbone. Being “massively adaptable”, or truly intelligent, is only an increase in adaptability of the organism along a gliding scale, not a completely separate state of being. Humans and other relatively intelligent creatures known to us did not arise from a different spawn than “unintelligent” life and many of the mechanisms underlying “unintelligent” life may be a vital part of what we call intelligence now. In order to understand “massive adaptability”, we therefore need to understand “adaptability” first.

Objections against cheapening “Intelligence” by applying bare-bone adaptability

A possible objection to the above definition of bare-bone adaptability is that it allows an impressively wide variety of organisms some measure of “intelligence”, quite likely even all. It can even be argued that the most humble bacterium satisfies the demands in a very basic manner. After all, the cell-membrane that encloses bacteria is capable of detecting chemicals in its environment and in many cases even allows them to communicate among species members, and even across bacteria species.[34] The detection of particular harmful chemicals, or chemicals associated with

---

[34] Federle, M.J. & Bassler, B.L. (2012). Interspecies communication in bacteria. *The Journal of Clinical Investigation 112* (9) 1291-1299.

harmful substances or organisms, will trigger the bacterium to undertake evasive action, while the detection of beneficial substances will encourage the bacterium to stay put. Not only behaviour is affected either, detection of high bacterial numbers in the surroundings can alter target-gene expression in individual bacteria as well.[35]

This detection and response mechanism not only satisfies the first of the four demands (**interaction**), but the demands of **storage** and **adjustment** are also satisfied *without* the use of a brain or even a nervous system of any sort. The storage mechanism in a bacterium is not a bunch of neurons like our brain as a bacterium is much too small to contain other cells and in fact ceases to be a bacterium by definition the moment it would associate with neurons. The storage mechanism is still, however, a very familiar information storage and retrieval method: DNA. DNA allows bacteria to adjust to their environment across generations, as mutations in the DNA will allow for differing adaptations. Natural selection then eliminates, on average, those adaptations which did not improve or maintain the current survivability. Over time, the organism “learns” because its predecessors “learned” by process of elimination and the out-competing of the less fit. This transfer of stored knowledge and adaptations is called “vertical transfer”.[36] Of course this manner is very primitive and can only be called “learning” if we look from the perspective of the self-replicating DNA-strains, rather than the organism they support. Random mutation, combined with the death of organisms whose interactions and behaviours do not cope properly with the survival-threats of their environment, allows bacterial DNA to store “value” to behaviour across generations: those interactions that prolong life are good and maintained, while those behaviours that invite death are bad and removed from storage through organism death. In this most primitive of adaptabilities, the second demand of **evaluation** is externalised: the environment selects beneficial interactions, forcing adaptation through elimination. This type of evaluation is present at any level of biological adaptability and relies completely on the severe “teachings” of the environment in combination with random chance, but it provides a form of learning none the less.

However, it is not the only learning mechanism present in bacteria. There are in fact mechanisms that allow for much quicker adjustment to their surroundings and which evade quite a bit of untimely bacterial death. Recent study has demonstrated that bacteria are capable of

---

[35] Miller, M.B., Bassler, B.L. (2001). Quorum sensing in bacteria. *Annual Review of Microbiology 55*. 165-199.

[36] Lawrence, J.G. (2005). Horizontal and vertical gene transfer: The life history of pathogens. In: Russell, W. & Herald, H. (ed.). *Concepts in Bacterial Virulence; Contributions to Microbiology 12* (Basel, 2005) 255-271.

exchanging genetic information amongst each other, even between species. This process, called “lateral transfer”, allows the bacteria to learn snippets of DNA-code from one another, which can lead to quick changes and adaptations to new environments within the short lifetime of a single organism. It is, among other adaptations, the primary cause for bacterial antibiotic resistance.[37] This is also not a rare process, as bacterial genomes often contain a significant amount of foreign DNA.[38] Another method of quicker learning is the activation and deactivation of genes already present in the DNA, an adaptational method explored by the field of Epigenetics. These adjustments, which do not change the DNA code but do change which genes are actually expressed, are a quick response to environmental factors allowing the bacteria, and higher life forms, to “learn” what its environment is like and to adapt by switching on the genes that produce the more desired responses (through the production of the right amino acids).[39] Furthermore, bacteria even have a rudimentary temporal memory, part of their environmental detection system, that allows them to improve their navigation in response to positive or negative environmental stimuli.[40] I will expand on this mechanism in the chapter on adaptability without a brain.

It is therefore not all that farfetched to state that bacteria are capable of learning from their environment: they can interact with their surroundings and are capable of influencing them, as well as being able to adjust their actions to their surroundings and storing environmental information and working solutions. The basic point here is not that bacteria are intelligent in the same way that we are. Their level of adaptability is not comparable to ours: their genetic adaptation is much quicker than ours, although we do possess the same genetic abilities, while humans and other complex life forms have more ways of learning during individual lifetimes than the relatively inelegant forms of natural selection and genetic modification. The point is that learning is essential to adaptability and therefore intelligence and that it is ***not necessarily*** an exclusively “mental” category. It is not even strictly limited to the brain. Higher intelligence is a form of massive adaptability that seeks to implement changes in the way the organism behaves during its lifetime, rather than across

---

[37] Gyles, C. & Boerlin P., (2014). Horizontally transferred genetic elements and their role in pathogenesis of bacterial disease. *Veterinary Pathology 51* (2) 328-340.

[38] Dobrindt, U., Chowdary, M.G., Krumbholz G. & Hacker, J. (2010). Genome dynamics and its impact on evolution of *Escheria coli*. *Medical Microbiology and Immunology 199* (3) 145-154.

[39] Danchin, E., Charmantier, A., Champagne, F.A., Mesoudi, A., Pujol, B. & Blanchet, S. (2011). Beyond DNA: integrating inclusive inheritance into an extended theory of evolution. *Nature Reviews: Genetics 12* (7) 475-486.

[40] Magnab, R.M. & Koshland, D.E. Jr. (1972). The gradient-sensing mechanism in bacterial chemotaxis (temporal gradient apparatus/stopped-flow/S. Typhimurium/motility tracks/memory). *Proceedings of the National Academy of Sciences of the United States of America 69* (9) 2509-2512.

generations, but it must still evaluate effects of its behaviour on the body it controls. Intelligence, by its very nature, is therefore embodied in the body whose behaviour it is controlling. It is important to remember this when we venture into the world of microbiology. In the next chapter, we will look at a gliding scale of biological adaptability, or learning, but before we do so, I will conclude with a final definition for what I am referring to when I refer to “creating Artificial Intelligence”:

> *“Creating Artificial Intelligence is the quest to manufacture a machine which is characterized by its massive adaptability. It is a machine that is capable of interacting with its environment, capable of storing these interactions in a meaningful way and able to adjust its future interactions based on these learned experiences.”*

Literature

Agah, A. (ed.) (2014). *Medical Applications of Artificial Intelligence* (Boca Raton, 2014).

Copeland, B.J. (1993). *Artificial Intelligence: A philosophical introduction* (Oxford, 1993).

Danchin, E., Charmantier, A., Champagne, F.A., Mesoudi, A., Pujol, B. & Blanchet, S. (2011). Beyond DNA: integrating inclusive inheritance into an extended theory of evolution. *Nature Reviews: Genetics 12* (7) 475-486.

Dictionary.com (2014). http://dictionary.reference.com/browse/artificial?s=t. *Dictionary.com* (retrieved 14 January 2014).

Dobrindt, U., Chowdary, M.G., Krumbholz G. & Hacker, J. (2010). Genome dynamics and its impact on evolution of *Escheria coli*. *Medical Microbiology and Immunology 199* (3) 145-154.

Federle, M.J. & Bassler, B.L. (2012). Interspecies communication in bacteria. *The Journal of Clinical Investigation 112* (9) 1291-1299.

Fine, J. (1997). Deep Blue wins in final game of match; Chess computer beats Kasparov, world's best human player. http://faculty.georgetown.edu/bassr/511/projects/letham/final/chess.htm. MSNBC (retrieved 6 June 2014).

Gyles, C. & Boerlin P. (2014). Horizontally transferred genetic elements and their role in pathogenesis of bacterial disease. *Veterinary Pathology 51* (2) 328-340.

IBM 100 (2011). Icons of Progress; Deep Blue Overview. http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/. *IBM* (retrieved 16 May 2014).

Jackson, J. (16 February 2011). IBM Watson vanquishes human Jeopardy foes. http://www.pcworld.com/article/219893/ibm_watson_vanquishes_human_jeopardy_foes.html.

PCWorld (retrieved 15 May 2014).

Lawrence, J.G. (2005). Horizontal and vertical gene transfer: The life history of pathogens. In: Russell, W. & Herald, H. (ed.). *Concepts in Bacterial Virulence; Contributions to Microbiology 12* (Basel, 2005) 255-271.

Magnab, R.M. & Koshland, D.E. Jr. (1972). The gradient-sensing mechanism in bacterial chemotaxis (temporal gradient apparatus/stopped-flow/S. Typhimurium/motility tracks/memory). *Proceedings of the National Academy of Sciences of the United States of America 69* (9) 2509-2512.

Mikolov, T., Sutskever I. & Quoc, L. (15 August 2013). *Learning the meaning behind words*. http://google-opensource.blogspot.nl/2013/08/learning-meaning-behind-words.html. *Google Knowledge* (retrieved 24 March 2014).

Miller, M.B., Bassler, B.L. (2001). Quorum sensing in bacteria. *Annual Review of Microbiology 55* (2001) 165-199.

Mitchell, T.M. (1997). *Machine Learning* (Boston, 1997) xv.

Robinson, H. (2012). Dualism. http://plato.stanford.edu/archives/win2012/entries/dualism. In: Zalta, E.N. (ed.). *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition) (retrieved 24 March 2014).

Russell, S. & Norvig, P. (2010). *Artificial Intelligence; A Modern Approach; Third Edition* (New Jersey, 2010).

Searle, J. (2004). *Mind: A brief introduction* (Oxford, 2004).

Turing, A.M. (1950). Computing Machinery and Intelligence. *Mind 59* (236) 433-460.

Upbin, B. (8 February 2013). IBM's Watson Gets Its First Piece Of Business In Healthcare. http://www.forbes.com/sites/bruceupbin/2013/02/08/ibms-watson-gets-its-first-piece-of-business-in-healthcare/. *Forbes* (retrieved 27 August 2014).

Chapter 2: Adaptability without a Brain

So why look at biological intelligence? The answer is simple: the only beings we know of that are generally considered "intelligent" are of biological origin. Now that we have established that intelligence is rooted in biological adaptability it makes sense to look at the biological origin of intelligence, or rather its constituent part "adaptability". By establishing what methods biological organisms use to establish their adaptability, we can hopefully extract useful information as to how to make Artificial Intelligence. As adaptability primarily focusses on learning, I will focus on the role of biological evolution in the development of learning with a special interest in the mechanisms that allow organisms to connect environmental events and their own actions to the proper set of consequences. The biological methods of evaluation may prove instrumental in creating an adaptable AI and will be the focus of my investigation.

Evolution and learning

It could be argued that evolution is a process of learning. In fact, this is precisely what I argued in Chapter 1 where I showed that adaptability is not just a mental attribute in the traditional sense, but a physical attribute as well. All biological organisms are geared to survive long enough and in large enough quantities in order to reproduce. As already proposed by Charles Darwin and Alfred Russel Wallace in the mid-nineteenth century, this is not because organisms are designed that way, but because those organisms that did not sufficiently meet the criteria have gone extinct through a process best known as "natural selection".[41] It is important to realise that there is no organizing force in evolution steering organisms down one evolutionary path or another, but natural selection does function as a learning mechanism as organisms adapt to their new environment through internal storage of successful survival techniques and the deletion or suspension of detrimental aspects. Through the ultimate consequences of death and procreation, natural selection also "grounds" internal information storage to the outside world. After all, organisms whose internal mechanisms do not reflect at least successful avoidance of death and successful pursuit of survival, including procreation, will go extinct. Internal representation of external factors is therefore grounded in the external environment through this evolutionary process. It is this "natural selection"

---

[41] Gregory, F. (2008). Chapter 17: New ideas about life and its past. *Natural science in western history* (Boston, 2008).

that provides a fundamental source of value for organism interaction-evaluation: in the end it is their death or survival that distinguishes good interactions from bad. In the following section, I will delve into an important mechanism for meaningful internal representation of death and survival: homeostasis. It is my hypothesis that it is the monitoring of homeostasis that allows biological organisms to attach meaning to their internal representations by making the proper connections between interactions and their internal consequences: a crucial connection for effective self-contained learning.

Two short warnings

The driving force behind “natural selection” is only a random and changing set of restrictive circumstances that cut off all mutations that hinder the survival and reproduction of a particular organism too much. Although this is a more accurate way of representing what actually happened and still happens throughout the course of evolutionary history, many scientists and non-scientists use more goal-oriented language to describe what takes place for simplicity's sake. So although giraffes with longer necks tended to be better nourished and therefore on average survived longer and reproduced more effectively, passing their genes on in greater number than their short-necked counterparts, most would simply say that giraffes “developed longer necks in order to reach the top of the trees”. This greatly simplifies the description and streamlines communication, but, like most simplifications, it also clouds the true mechanism behind the giraffe's evolution: giraffes with longer necks were simply less handicapped in dealing with their environment and passing on their genes than giraffes with shorter necks. “Natural selection” did not so much “select” the long-necked giraffes as it “deselected” the short-necked specimens.

Although the shorthand is inexact and at times misleading, it is in many cases a much less convoluted way of speaking about particular subjects and mechanisms. Therefore, usage of goal-oriented language may surface in the following section, but the reader should keep in mind the random training environment behind natural, low-level adaptation. The natural environment in essence grants a near infinite amount of trials that the biological organisms must use to train their survival mechanisms.

Another warning is required for the assumption that there is a definite hierarchy to the levels of advancement found in life forms. Although through scientific history, the assumption has often been made that there is a *scala naturae* where some “lower” organisms today are considered primitive exemplars of aeons past, this assumption is impossible to maintain in modern day evolutionary biology. Instead, creatures belong in clades that separated from each other in the evolutionary tree at given points in time. No matter how early the separation, every organism alive today is part of a crown-species, a species that has gone through the same billion years of random mutation and natural selection that humans and their direct ancestors have.[42] This means that detecting traits in unicelled organisms and then concluding that those traits must have been present in our unicelled common ancestors is a dangerous path to follow. Regretfully, there are painfully few ways of exploring the capabilities and behaviours of long extinct microorganisms. In the following section I will assume that traits shared by *all* living organisms are not the result of convergent evolution, but most likely due to having originated in a single common progenitor species.

Three short definitions

The following two terms will surface on a regular basis in the following chapter. Their meaning is not necessarily as straightforward as it appears, so a short definition of their use is in order:

- Reward: a reward is the internal value signal representing a survival value for an organism. This survival value can be bad or good, often depending on the circumstances. Whenever the word “reward” is used in a general context, it should be taken to include both reward and punishment.
- Survival: survival in an evolutionary context is less about the survival of an individual and much more about the survival of its genetic make-up, better known as DNA. Using this definition handily includes procreation in survival-necessities, which allows for a better description of natural selection. A survival value can be positive or negative, often depending on the circumstances.
- Signifier: a signifier is an external signal (such as a detected chemical) that can be used as a reward-event predictor. It “signifies” the availability of reward or impending punishment.

---

[42] Murray, E., Wise, S. & Rhodes, S. (2011). Chapter 4: What can different brains do with reward? In: Gottfried, J.A. (ed.). Neurobiology of Sensation and Reward (Boca Raton, 2011).

Staying alive: the constant struggle for homeostasis

All living organisms, from the smallest single-celled organisms to the most massive of plants and animals, have a delicate internal balance that must be maintained in order for them to stay alive. This balance, called homeostasis, is not a matter of keeping a particular bodily value as low as possible, or as high as possible, but one of keeping a value within a specific range.[43] Anything too high or too low is a negative condition as it severely increases the chances of death. Anything within the optimum range is a positive condition as it signifies sustained survival.[44] Furthermore, organisms don't have a single parameter for which to maintain homeostasis, but a multitude including, but not limited to: temperature, pH-value, hydration, and nutritional values such as usable energy and the availability of required chemical compounds. Homeostasis is a concept that permeates the field of biology and serves as the explanation as well as the driving force for a wide variety of self-regulatory mechanisms that maintain internal balance.

Maintaining this internal balance is crucial and tricky. If one of the parameters, such as temperature, falls too low or rises too high, cellular damage starts to occur and internal processes such as metabolism may no longer work as required. The correct temperature can be found on a gradient, with values too low and too high both being detrimental to creature survival. Somewhere in between lie values that are sustainable, although some of these values may still produce better results than others. Some of these homeostasis-parameters are very strict. For instance, the human blood pH-value is kept within a narrow range: 7.37 to 7.43, with an ideal value of 7.40.[45] Any deviation outside this band triggers a range of internal problems that will lead to an untimely death if left unattended. To offset these problems, the organism experiences internal drives, it needs to detect that something bad is going on, and it must then act upon this if it is to survive. However, actions themselves also require the expenditure of resources. This means that any organism that undertakes action automatically risks disruption of homeostasis if it undertakes that particular action for a time without being compensated for this resource loss with a gain, or at least a draw, in survivability. Actions are detrimental to homeostasis unless they somehow benefit the homeostatic balance.

---

[43] Craig, A.D. (2003). A new view of pain as a homeostatic emotion. *Trends in neuroscience 26* (6) 303-307.

[44] Damasio A. & Carvalho G.B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience 14* (2) 143-152.

[45] Lewis, James L. III (October 2013). Acid-base regulation. http://www.merckmanuals.com/professional/endocrine_and_metabolic_disorders/acid-base_regulation_and_disorders/acid-base_regulation.html?qt=&sc=&alt. *The Merck manual* (retrieved 2 June 2014).

To make matters worse, inaction will also automatically lead to homeostasis disruption in all organisms.[46] This is due to the fact that all organisms have some form of metabolism where they convert chemical compounds or photons, henceforth referred to as nutrients, into energy. These nutrients are required for DNA-replication and for the repair of DNA-damage caused by environmental factors. This not only serves to sustain the current individual, but is also required for cellular growth as well as ensuring the long-term survival of the DNA through procreation as any creature not procreating will eventually die out. 100% Efficiency in this conversion is impossible due to the random damaging environmental factors that force the need for reparation, as well as the physical impossibility to transmute one nutrient into another without at least some loss of energy. As perhaps is most fittingly expressed through the law of conservation of energy, the internal processes of the organism itself, combined with the impossibility of attaining 100% efficiency, will cause internal homeostasis to be disrupted as nutrient availability without outside influx will dry up. Enduring inaction will therefore inevitably prompt organism-action if it seeks to survive, otherwise internal and external factors can and will disrupt internal homeostasis and through that process send the organism on a path towards death.

Avoidance of death is something strongly promoted by natural selection. So strongly even, that death without procreation could be considered the antithesis to evolution: any DNA that has evolved and still exists today has had to implement a mechanism that counteracts the death-process or it would simply no longer be around. To maintain their internal homeostasis, organisms have internal negative feedback loops that detect deviations and activate countermeasures.[47] As mentioned, organisms are constantly depleting their internal environment, this means that the eventual tools for correcting homeostatic disruption need to come from the external environment. On the other hand, external environments can also disrupt internal homeostasis if they are sufficiently hostile. It is therefore vital that the organism has a way of mitigating detrimental environmental factors and seeking out helpful environments. If an environment is too hot or too cold, maintaining internal homeostasis becomes too difficult and the organism needs to undertake action to counteract this imminent threat. Even an environment that contains a high concentration of nutrients can be problematic, as many nutrients can be toxic in high quantities.[48] However, when

---

[46] Some mosses, seeds and micro-organisms can lie dormant for millennia, waiting for circumstances to improve, but eventually even they will succumb to randomly occurring damage.

[47] Craig, A.D. (2003). A new view of pain as a homeostatic emotion. *Trends in neuroscience 26* (6) 303-307.

[48] Hathcock, J. N. (1989). High nutrient intakes – the toxicologist's view. *Symposium Upper limits of nutrients in infant*

any internal balance is disrupted, an organism can also use outside help to restore internal homeostasis. For instance, when the organism is becoming dehydrated to the point where damage will start to occur, it becomes important for it to acquire water. So much so that the risk of death due to the need for water may overwrite the need to avoid other potentially lethal hazards.[49] It becomes essential for an organism to take risks. After all, any undertaken action includes the risk that it is more detrimental to homeostasis than the benefits it provides, not only through aversive external effects, but also through a higher degree of resources spent. The mechanism through which action is promoted and this homeostasis-oriented action-taking is encouraged, is called reward and punishment. It serves as a translation between internal needs and external factors, as well as driving internal motivation to change external factors to more beneficial or less detrimental ones. It is likely that feelings have arisen in this context of maintaining homeostasis.[50,51]

Homeostasis seems therefore to be the perfect foundation for connecting actions to their internal consequences and providing them with value. Serious disruption of internal homeostasis brings the ultimate consequence for any organism, death, while the restoration of homeostatic parameters to their proper range increases survival. Evaluating the impact of actions on homeostasis is therefore vital to organisms and absolutely central to adaptation during an organism’s lifetime. Monitoring internal homeostasis and making connections between homeostatic disruptions and external factors as well as organism-actions therefore allows organisms to attach consequence to their actions and with it evaluate them. With these evaluated actions they can start taking informed decisions, in other words they can now learn from their actions through a method both quicker and more efficient than natural selection and organism death.

Reward learning in single-celled organisms

To illustrate the importance of combining external input with valued judgement on the basis of homeostasis, let us return to some of the smallest and least intelligent organisms on the planet. Bacteria are generally not considered intelligent and most people are only familiar with their generational methods of adaptation, i.e. strict mutation and natural selection. However, even these

---

formulas (November 7-8, 1988) 1779-1784.

[49] The case of the drinking wildebeest and the crocodile lying in wait comes to mind.

[50] Damasio A. & Carvalho G.B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience 14* (2) 143-152.

[51] Craig, A.D. (2003). A new view of pain as a homeostatic emotion. *Trends in neuroscience 26* (6) 303-307.

single-celled organisms have an internal homeostasis to maintain and it seems therefore likely that they have a way of adapting to circumstances during their own lifetime.

As discussed in the section on Artificial Intelligence, adaptability boils down to four basic points:

- A "being" must be capable of **interaction** with its environment (requiring some form of perception and some means of altering itself or the environment),
- A "being" must be capable of **evaluating** its interactions with the environment,
- A "being" must be capable of **storing** these valued interactions, (more commonly known as having a "memory"),
- A "being" must be capable of **adjusting** its interactions **based on these values** attained through previous interactions/perceptions (more colloquially known as "learning").

All these four requirements are requirements posed by natural selection upon any being that seeks to survive long enough to procreate. Being adaptable is what evolution is all about and interaction with the environment is very important. Any organism that seeks to adjust its interactions with the environment to increase its odds of survival, needs to perceive non-lethal input from the environment and then execute a, hopefully appropriate, response. However, it is important that any organism that wants to adjust to its environment, is also capable of assigning meaning to the signals it picks up from both the environment as well as internal signals. This is a mechanism so crucial, even simple bacteria use it, not just through the passing on and exchange of genes that has already been discussed in Chapter 1, but even through direct environmental monitoring. Although I will discuss bacteria, members of the Prokaryote domain, the following also applies to unicelled organisms that fall under the Eukaryote domain, the same domain that humans, plants, animals and algae fall under. I have decided to stick with bacteria in the coming section in part because it makes for an easier read. Much more importantly though, I want to show that adaptability, reward and punishment is something that is shared by all life, not just the Eukaryotes. With this, I hope to evade the exceptionalism that has plagued the Philosophy of Mind, Intelligence and AI, where humans, great apes, primates and mammals have all been ascribed special and unique properties at one point or another that are hard to maintain when compared to supposedly out-group species.

Interpreting external signals is fairly straightforward for single-celled organisms. An environmental signal signifies something that either:

- increases your chances of dying, a negative signal, or
- increases survivability, a positive signal, or
- does neither, a neutral signal.

It seems therefore obvious that bacteria judge incoming signals on two vectors: a positive effect value and a negative effect value. For single celled organisms and organisms without a distinct neural net, sensations and reward signifiers are basically the same thing. So much so that constructing sensors for chemicals or other potential signals that do not carry a positive or a negative connotation could be considered redundant and a waste of resources from an evolutionary standpoint.[52] Further down the line I will discuss why neutral signals will still be picked up by organisms, but for now it is useful to note that good and bad are rooted in the very foundation of life: good at its best promotes life, bad at its worst terminates it. Neutral signals require no particular change in behaviour, while signals that signify danger require a quick opposing response (such as rapid motion in the opposite direction of the signal). Signals that signify beneficial environments (such as food) require approaching responses.

As already stated, bacteria need to maintain internal homeostasis like all living organisms. As only a thin membrane separates their vulnerable internal mechanisms from the outside world, monitoring change in their environment is very important. They need to be able to detect danger, or positive circumstances, so they can adjust their behaviour accordingly. There is plenty of scientific evidence that they indeed do this. Bacteria capable of movement, respond by altering their movement patterns when they detect changes in temperature, light, salinity, oxygen and specific metabolites and other signalling molecules. Movement in response to the last is called *chemotaxis*.[53]

In order to respond to chemical stimuli, bacteria developed one of the first senses, perhaps the first sense altogether: a sense of smell. Receptors on the outside of the membrane are capable of

---

[52] Gottfried, J.A. (2011). Preface. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Chicago, 2011).

[53] Baker M.D., Wolanin, P.M. & Stock, J.B. (2005). Signal transduction in bacterial chemotaxis. *BioEssays 28* (1) 9-22.

binding chemical molecules in the environment and can then release intracellular communication in a way very similar to our own.[54] However, due to their respective size in comparison with the size of the molecules they are sensing, bacteria are often too small to adequately measure particle density with their limited surface area. There is simply not enough room to fit the required amount of sensors to accurately detect concentration densities at this scale, let alone being able to detect gradient differences between one side of the bacterium and the other. To make matters worse, the random fluctuations in concentrations of chemicals at the bacterial size-scale is impractically high, making it near impossible for bacterial "senses" to detect which way avoids death by a single measuring moment in time.[55] Due to size restrictions, bacteria seem unable to reliably tell which way is safer or more beneficial and yet it is obvious from their behaviour that they reliably move away from danger and towards attractive stimuli. However, this paradox only occurs when one assumes that bacteria are incapable of comparing changes in their environment across time-intervals.

Research has uncovered that bacteria possess mechanisms for the detection of temporal gradients, that is to say, they are able to compare concentrations of signal transmitters over a time interval and then evaluate whether the new situation is an improvement or not.[56] Many motile bacteria[57] possess two basic modes of movement: a coordinated, mono-directional burst of movement and a "tumble" mechanism that rotates them into a random new direction.[58] When they are present in an environment with a uniform distribution of a positive signifier, no matter the concentration, they will alternate between movement along an almost straight line and random tumbling at a default rate.[59] This is their normal state, it is required for the detection of changes as well as more straightforward survival: as has been mentioned, permanent internal homeostasis is

---

[54] Gottfried, J.A. & Wilson, D.A. (2011). Chapter 5: Smell. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

[55] Magnab, R.M. & Koshland, D.E. Jr. (1972). The gradient-sensing mechanism in bacterial chemotaxis (temporal gradient apparatus/stopped-flow/S. Typhimurium/motility tracks/memory). *Proceedings of the National Academy of Sciences of the United States of America 69* (9) 2509-2512.

[56] Magnab, R.M. & Koshland, D.E. Jr. (1972). The gradient-sensing mechanism in bacterial chemotaxis (temporal gradient apparatus/stopped-flow/S. Typhimurium/motility tracks/memory). *Proceedings of the National Academy of Sciences of the United States of America 69* (9) 2509-2512.

[57] Bacteria with the ability of self-propulsion.

[58] Chatterjee, S., da Silveiram R.A. & Kafri, Y. (2011). Chemotaxis when bacteria remember: Drift versus diffusion. *PloS Computational Biology 7* (12) Special section 5. 1-8.

[59] Magnab, R.M. & Koshland, D.E. Jr. (1972). The gradient-sensing mechanism in bacterial chemotaxis (temporal gradient apparatus/stopped-flow/S. Typhimurium/motility tracks/memory). *Proceedings of the National Academy of Sciences of the United States of America 69* (9) 2509-2512.

impossible for life forms, they always consume energy and so must find new energy to replenish the old. This provides an internal drive to undertake action.

However, when introduced to a higher concentration of the same positive signifier, bacteria will change their behaviour: they increase the lengths of their mono-directional burst, which results in less frequent random tumbling. The same occurs when they detect a lowering concentration of a negative signifier. This increases their movement towards the positive signifiers and away from the negative ones. When bacteria are introduced to a lower concentration of positive signifiers, or a higher concentration of negative signifiers, they will instead shorten the duration of their straight runs, which results in more frequent tumbling to change direction until a more beneficial concentration is detected[60]. The combination of these two factors leads to bacteria moving effectively towards a positive signifier and away from negative signifiers, as they rapidly change direction when their current direction leads to a negative temporal gradient, while staying more true to their direction when a positive temporal gradient is detected. Once their changed movement method no longer appears to provide any advantage, that is to say, when they no longer detect any temporal changes in positive or negative signifiers, bacteria revert to their normal state of straight runs and random tumbles.[61] Bacteria, in other words, compare the old situation to the new and decide whether it has improved, has grown worse or has stayed the same and then adjust their behaviour accordingly. This is another vector of adaptability: bacteria keep track of value-changes over time which requires them to temporarily store a survival value, interact with the environment, compare the two values and then adjust their behaviour accordingly. So far, this is done on predetermined avoidance and approach signifiers, but no permanent learning of the living organism is yet involved. Adjustment of approach or avoidance on the basis of negative or positive signifiers taken across a time-differential, is a very basal form of adaptation present in bacteria.

All well and good, bacteria are able to adjust, but this does not cover the whole range of their reward-sensing apparatus. It can actually learn as well and by a much quicker method than genetic exchange and mutation. Research has uncovered that bacteria prefer environments that “smell” like the environment in which they grew up.[62] The chemical signals of their original

---

[60] Chatterjee, S., da Silveiram R.A. & Kafri, Y. (2011). Chemotaxis when bacteria remember: Drift versus diffusion. *PloS Computational Biology 7* (12) Special section 5. 1-8.

[61] Magnab, R.M. & Koshland, D.E. Jr. (1972). The gradient-sensing mechanism in bacterial chemotaxis (temporal gradient apparatus/stopped-flow/S. Typhimurium/motility tracks/memory). *Proceedings of the National Academy of Sciences of the United States of America 69* (9) 2509-2512.

[62] Gottfried, J.A. & Wilson, D.A. (2011). Chapter 5: Smell. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and*

environment, even when they don't signify any particular positive effect as expressed by the DNA-code that stores the information of generations past, are preferred over other neutral signals. The bacterium has learned to associate these chemical signals with a positive thing, presumably a safe and stable environment (something that occurs in macroscopic organisms as well). As a consequence, the response each individual bacterium displays to any given signal depends on the history of that particular cell.[63] Bacteria are also capable of learning new smells. Neutral signals that coincide with the occurrence of reward signifiers, can take on the value of those reward signifiers themselves, becoming positive or negative signifiers in the process.[64] This process of learning is fundamental in the survival of motile bacteria and the main reason for detecting neutral signals. Bacteria are capable of long-term evaluation of new chemicals, based on their interactions with the environment and pre-established valued chemicals. An impressive feat suggestive of a rudimentary learning that can certainly be qualified of bare-bone adaptability.

However, it is also suggestive of something deeper. Rather than a one on one, straightforward coupling of reward signifiers with the corresponding good or bad, it seems that bacteria are capable of decoupling chemical sensing and the values attached to the sensed chemicals. It appears that they instead have mechanisms to signify good or bad, that are separate of the chemicals they give value. Although I have been unable to find this mechanism, it seems that external signifiers are mapped to this “reward matrix”. The presence of a separate reward-system seems established even in these relatively simple life forms. Bacteria are not just capable of detecting whether change occurs, or even whether it is beneficial or detrimental, they are even capable of adjusting their evaluation based on training. This means that the four component parts of adaptability (interaction, evaluation, meaningful memory and action-adjustment) are indeed all present even during bacterial lifetime adaptability. The evaluation, or “meaning-giving” component of the stored information appears to reside in a reward and punishment matrix that has, as of yet, remained unidentified.

Reward (Boca Raton, 2011).

[63] Baker M.D., Wolanin, P.M. & Stock, J.B. (2005). Signal transduction in bacterial chemotaxis. *BioEssays 28* (1) 9-22.

[64] Gottfried, J.A. & Wilson, D.A. (2011). Chapter 5: Smell. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

Reward and change

Perhaps now the question arises: what about rewards signalling that the current state is stable and good? There is no reward-signal mentioned for this in the bacterial example above. Whether in a homogenous environment of positive signifiers, or one of negative signifiers, bacteria do not adjust their behaviour from the norm. Instead they tumble and swim about with their regular pattern. It is only when they are introduced to a change over time, that they adjust their behaviour to enhance or counteract this change, depending on its nature.[65] The explanation for this is simple.

As mentioned previously, organisms need to maintain internal homeostasis. This homeostasis is the norm, it is normal and should not change. This means that initial internal change is bad. An organism whose food stocks are dwindling, or whose pH-value is increasing, or who suffers from any other deviation from the norm, is experiencing a negative change. This negative change has an absolute negative value: if it persists, it increases chances of death. To prevent these deviations, organisms have internal negative feedback mechanisms that are triggered when their homeostasis is disrupted. This mechanism sets in motion changes that counteract the imbalance. If the mechanisms somehow exacerbate the matter, they are bad as they increase the chance of death. However, if they manage to restore the original homeostasis, their change is good as they promoted survivability. Although death is absolute, the perception of organisms of things bad or good is always referring to change, and the threat thereof, in internal homeostasis over time.

This means that feedback mechanisms only operate when there is a change on which to act. They are alerted by a negative reward-signal when the situation is getting worse, the part of the organism that gives off this signal is usually counted as part of the feedback mechanism, while they are encouraged by a positive reward-signal when the situation is improving. When the current situation is stable and no change is required, there is no signal to give because there is no action to undertake and the release of reward signals stops through a negative feedback mechanism. Reward and punishment, triggered by signifiers and internal disruption, serve a motivating role exclusively and do not cause actions in a state of equilibrium such as established homeostasis.

But why not always strive for perfection in the outside environment? Although internal cellular homeostasis is set within near-fixed parameters, external homeostasis is not quite as fixed.

---

[65] Magnab, R.M. & Koshland, D.E. Jr. (1972). The gradient-sensing mechanism in bacterial chemotaxis (temporal gradient apparatus/stopped-flow/S. Typhimurium/motility tracks/memory). *Proceedings of the National Academy of Sciences of the United States of America 69* (9) 2509-2512.

As the cellular membrane separates the internal from the external, the external homeostasis can differ from the internal necessities as long as the organism is able to cope with the differences through internal mechanisms. This means that the external requirements for homeostasis are not quite as strict as the internal requirements. Although any organism has a preferred optimum of external values, they also have a wide range in which their survival is perhaps not optimal, but still good enough. This is a necessity, as, unlike their internal environment, the external environment is very difficult to manipulate and the perfect environment may simply not exist. It is therefore important that organisms only spend energy on intensive interaction, deviation from the most basic actions required to stay alive in a homeostatic environment, when they have an indication that such action will have beneficial effect, i.e. when there is something bad to avoid, or something good to seek out. In other words: when they are motivated by external signifiers and their internal reward system to undertake action.

An evolutionary change has allowed organisms a bigger influence on their external environment though. Cooperation between cells has greatly increased malleability of the environment while also causing a host of new problems. It is through this change that the reward systems have taken on a new role.

Communication between cells

Since their origin, single-celled organisms have had an interesting problem. Due to the nature of their breeding, or really that of any breeding organism, a successful organism in a suitable environment is unlikely to find itself alone for very long. Due to cytokinesis, the division of a single full-grown bacterial cell into two new bacterial cells, bacteria often find themselves living close to other cells of the same species and often even the same genetic make-up. This leads to various problems, such as the rapid depletion of required resources and the production of potentially harmful amounts of unusable waste products. Furthermore, reward-signifiers become harder to detect when isolated by layers of other organisms and their absorption of the signals. Many motile bacteria and protozoa[66] therefore take action to create some distance between themselves and other, competing bacteria.

However, this is not always the case. Several single-celled organisms that we know of are

---

[66] Single-celled organisms belonging to the Eukaryotes, rather than the Prokaryotes-group.

known to form clusters on occasion, or even habitually. According to the fossil record, this type of multicellularity has occurred in prokaryotes and eukaryotes for several billion years.[67] This can be explained through the fact that there are circumstances in which living in a group of likeminded, or even identical individuals can be beneficial. One of these advantages is that organisms living in a group can exert greater pressure on their environment. The combined output of millions of bacteria is guaranteed to make more of an impact than that of individual bacteria.[68] Clustering together can make cells less vulnerable to predators, it can help them conserve nutrients, it also allows them to divide labour, or even expands their range of metabolic opportunities.[69] Another advantage is that living in a group with likeminded individuals can also provide an early warning system as long as the individual cells know the signs to look for. This leads to improved environmental detection. Living in a group, chemical signals could be used to warn of deadly threats that provide very little warning of their own. Even threats so aggressive and stealthy that they are already demolishing the outlying cells can now be signalled, as long as the signals are able to outpace the damaging factor. The key to survival in a group is therefore communication and in unicellular organisms capable of forming clusters a large part of their sensing apparatus seems indeed geared towards intercellular communication.[70] This section deals with setting up meaningful interactions between individual cells and it can be kept relatively short, because it is really not that complicated.

Meaningful communication between micro-organisms

Every single-celled organism absorbs nutrients and signifiers from the environment, while releasing left-over chemicals. As signifiers are also simply chemicals, and bacteria are quite capable of picking these up, the most primitive mechanisms for inter-organism communication was already in place before communication was most likely attempted. Communication likely started by an accidental repurposing of the chemo-sensory array and the associated reward-signalling mechanism to analyse waste-products of organisms of the same species or different species. Their increased presence indicates increased organism-density, which can be a positive or negative factor depending

---

[67] Grosberg, R.K. & Strathmann, R.R. (2007). The evolution of multicellularity: a minor major transition? *Annual Review of Ecology, Evolution, and Systematics 38*. 621-654.

[68] Park, S., Wolanin, P.M., Yuzbashyan, E.A., Silberzan P., Stock, J.B. & Austin R.H. (2003). Motion to form a quorum. *Science 301* (5630) 188.

[69] Grosberg, R.K. & Strathmann, R.R. (2007). The evolution of multicellularity: a minor major transition? *Annual Review of Ecology, Evolution, and Systematics 38*. 621-654.

[70] Baker M.D., Wolanin, P.M. & Stock, J.B. (2005). Signal transduction in bacterial chemotaxis. *BioEssays 28* (1) 9-22.

on the needs of the individual so it makes sense for this survival mechanic to start monitoring this. It seems logical that, for instance, the waste-product of a successful metabolism of nutrients became a positive indicator of nutrients in the environment, while the waste-products of cellular damage and repair became an indication of danger. Through their ability to distinguish chemicals, cells became capable of detecting warning signs generated by other organisms. Thanks to their internal reward/punishment matrix, combined with environmental information (other signifiers), cells were capable of grounding them with meaning. Organisms particularly adept at generating and perceiving these warnings as well as understanding them, had a greater adaptive value, letting them procreate more effectively and outcompete the organisms that did not. Eventually, cells became quite capable of communicating with other cells through chemical signalling. The internal consequences of external signals allowed the cells to give value to the signals of other cells and therefore ground them with survival-meaning: positive or negative signifiers on one or more homeostatic axons.

Meaningful communication within macro-organisms: the development of neurons and the brain

Even in groups of single-celled organisms, signalling is fairly straightforward. Releasing chemicals into the surrounding environment functions as a communication mechanism between individual cells within groups. However, the real communicational challenges and opportunities arose when some cells started a closer cooperation, with the evolution of organisms that were composed of a multitude of differentiated cells working tightly together for their common survival. Although it is still unsure through which mechanisms single-celled organisms gave rise to complex[71] multicellularity, the fact of the matter is that it did, and on several occasions to be precise.

Multicellular organisms arose from single-celled organisms on at least 25 separate occasions.[72] However the rise of complex multicellular organisms with a differentiated cellular structure is much rarer, having only arisen a handful of times on separate occasions. Examples of separate groups of multicellular life forms are plants, insects and animals, who have each evolved

---

[71] “Complex” multicellularity refers to organisms composed out of differentiated cells, in contrast to for instance biofilm producing grouping single-celled organisms that live in groups.

[72] Grosberg, R.K. & Strathmann, R.R. (2007). The evolution of multicellularity: a minor major transition? *Annual Review of Ecology, Evolution, and Systematics 38*. 621-654.

from a different lineage.[73] As complex multicellularity provides several interesting evolutionary advantages, such as size, functional specialization and division of labour, it is somewhat surprising that only a handful of lineages have persisted and branched out.[74] An important factor in achieving complex multicellularity is establishing meaningful and expedient communication between cells of different types and functions.

For cells cooperating in a single complex organism, the feedback loop between the cellular death of one cell and the death of another is not so strong as among uniform multicellular organisms. Although harm may befall some cells in an organism, this does not spell immediate doom to others, it may even be required for the survival of the organism. Multicellular organisms even feature mechanisms that allow for programmed cell death, a suicide trigger that forces cells to sacrifice themselves for the benefit of the group.[75] Malfunction of this mechanism is one of the problems in cancerous growth, showing the importance of individual cell death in complex multicellular organisms. It is therefore adamant that a mechanism is created that directly mediates between the survival of single cells and the ability to accept some cellular damage for greater gains. The complex organism that is composed of these single cells requires valued detection of the outside world and internal problems on a scale that surpasses single-cell survival. This new evaluation-mechanism is rooted in two forms of cellular communication. The first is a form that was already prevalent amongst the single celled organisms: communication through chemicals released into a plasma that connects individual cells. The second technique is new: some cells specialised into oblongated, sensory cells whose only purpose became to detect danger and distress on a level higher than that of single cells and then rapidly communicate this to other cells for adequate action. These cells, the first primitive neurons, are the ancestors of all neural systems.

As discussed above, cells in a multicellular organism have the individual ability to register and evaluate aspects of their surroundings that they have inherited from their unicellular ancestors. However, cells within multicellular organisms have the challenge of a varied exposure of cells to external and internal environments. On the other hand, complex multicellular organisms provide the opportunity for cells to specialise and one of these possible specialisations is dedication to sensory-

---

[73] Grosberg, R.K. & Strathmann, R.R. (2007). The evolution of multicellularity: a minor major transition? *Annual Review of Ecology, Evolution, and Systematics 38*. 621-654.

[74] Grosberg, R.K. & Strathmann, R.R. (2007). The evolution of multicellularity: a minor major transition? *Annual Review of Ecology, Evolution, and Systematics 38*. 621-654.

[75] Grosberg, R.K. & Strathmann, R.R. (2007). The evolution of multicellularity: a minor major transition? *Annual Review of Ecology, Evolution, and Systematics 38*. 621-654.

functions: detection of potential death and survival signals at the organism scale.[76] Neurons are one such a specialisation. These specialised cells have taken over the function of detection and long-distance communication that requires a certain degree of speed, as well as greater accuracy. Through the net formed by neurons, signals from one part of the body can quickly travel to another part of the body, enhancing organism survival through greater speed.

When thinking of neurons, it is easy to think of only the brain, but no brain is required for neurons to function on a basic level. Neurons came first, the brain came second and in fact a large number of species function with neuron cells, but without a central governing neuron organ. Many complex organisms in fact do not even have a centralised nerve system, such as members of the *Cnidaria*, which includes jellyfish. These creatures instead possess a diffuse network of neural connections, allowing cells in one part of the body to communicate and cooperate with cells located elsewhere in a rudimentary fashion. These communications allow for simple approach and avoidance behaviour, but it's hard to establish if the organism has any kind of integrated “feeling” experiences beyond the cellular level.

From the above it follows that not all neurons are brain-neurons. An example of non-brain neurons are sensory neurons, such as those specialised in photoreception or chemoreception. Sensory neurons bear cilia or microvillar structures on their surfaces, which are connected to complex membrane structures. They detect the outside environment through means that are strongly reminiscent of those applied by bacteria, but they then communicate by electrical potential through synapses located on their axons (long tail-like structures, part of the neurons body, that allow the signal to be transported far away before being transmitted to other neural cells) or via synapses to adjacent neural cells.[77]

Interestingly enough, these communications still require chemical communications to work. Although neurons build up an electric potential across their body, most of the actual communication from neuron cell to neuron cell across synapses is mediated by the release of chemicals called neurotransmitters. These neurotransmitters are what prompt neurons to undertake or abstain from

---

[76] Jacobs, D.K, Nakanishi, N., Yuan, D., Camara, A., Nichols, S.A. & Hartenstein V. (2007). Evolution of sensory structures in basal metazoa. *Integrative and Comparative Biology 47* (5) 712-723.

[77] Jacobs, D.K, Nakanishi, N., Yuan, D., Camara, A., Nichols, S.A. & Hartenstein V. (2007). Evolution of sensory structures in basal metazoa. *Integrative and Comparative Biology 47* (5) 712-723.

action.[78,79]

As complexity increases, it becomes harder and harder for cells that experience the benefits of a particular action to communicate this to the cells responsible for the action. For instance, cells that are being overheated will have trouble rewarding the brain-neurons responsible for alleviating the pain by applying cooling water. The old method of simply dumping their rewarding chemicals into the surrounding liquids in the hopes of them reaching the to-be-rewarded cells, has two major drawbacks:

- It is **slow**. This can lead to a dissociation between the actual beneficial act and the evaluating reward. Without adequate coupling between the two, negative or positive actions will not be properly signalled or reinforced, which disables the adaptability requirement of proper evaluation. It is also quite possible for cells to receive the wrong message, encouraging or discouraging them to undertake actions that were not meant to be valued, resulting in a kind of anti-adaptability.
- It is **inaccurate**. Not just the cells and pathways responsible for the beneficial action may be rewarded, but cells and actions that have nothing to do with the positive effect are reinforced as well. This is undesirable, as it makes it impossible to specialise the beneficial behaviour. Do note, that to some extent this still happens, even in a centralised nervous system with a dedicated reward system.

So reward, the cellular representation of survival value, needs to be dealt out more accurately and in a much speedier manner when distributed within a complex multicellular organism. Otherwise beneficial action by a group of neurons in the brain will be unrewarded or wrongly rewarded and therefore unvalued or valued improperly. To mediate between signals from remote parts of the body and the brain neurons responsible for actions, the new neuronal network developed a specialisation. Some neurons became tasked with dealing out positive and negative signalling chemicals to other neurons in appropriate situations, representing the body experience in the brain. These chemicals and the connections they enforce are, in my view, the basis for our experience of reward and

---

[78] Fields, R.G. & Stevens-Graham, B. (2002). New insights into neuron-glia communication. *Science 298* (5593) 556-562.

[79] Kalat, J.W. (2004). *Biological Psychology; 8th Edition* (Belmont, 2004) 53-58, 60-61.

punishment. From both an evolutionary and an ontogenetic perspective, this experiential aspect of body-representation in the mind can be considered the lowest level of both mind and consciousness, as it provides an integrated bodily experience in the brain.[80]

Conclusion

Thanks to cellular evaluation-systems, single-celled organisms are motivated to take actions or cease taking actions. This evaluation takes place on the basis of homeostasis. Maintaining homeostasis is required for survival, while letting a disrupted homeostasis go unmanaged will result in death. By combining homeostatic measurements with interactions with the environment and storing the now valued results, it becomes possible to create meaningful connections that are ultimately rooted in the survival/death paradigm that dominates evolution. Successful mapping of external events to internal consequences grants even microscopic organisms the ability to attach value to new signifiers. In other words, they become capable of learning on a very rudimentary level. To support this learning, single-celled organisms are equipped with a mechanism that evaluates the impact of environmental signifiers by making valued connections with homeostatic disruptions and with other already valued environmental signifiers. Due to this mechanism, single-celled organisms can assign survival values to previously meaningless external signals. It is through this method of connecting external events to internal consequences that single-celled organisms are able to adapt to changing circumstances during their lifetime rather than by the less direct natural selection mechanism. In essence, it provides a bare-bone adaptability that goes beyond random chance. Being able to make valued connections (evaluation) between external signals and organism actions (interaction) appears to reside in the necessity for organisms to maintain homeostasis. By monitoring homeostasis, an organism can evaluate the outcome of actions that would otherwise require the final evaluation: natural selection through death of the ill-adjusted.

Due to the nature of homeostatic monitoring and the necessity of recognising and valuing signifiers that are not pre-valued in the DNA-code, the reward/punishment signal is presumably separated from the received signifiers. This microscopic ability for attaching value to signifiers also allows for macroscopic valued feedback between single-celled organisms living in groups, as well as the differentiated cells present in complex multicellular life. Cells can now communicate among

---

[80] Damasio A. & Carvalho G.B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience 14* (2) 143-152.

one another by interpreting the chemical signals that they release and grounding them on the internal meaning provided by their own homeostatic evaluation. In the development of complex multicellularity, a new type of communication cell was differentiated: the neuron. These neurons were capable of making connections amongst each other and centrally directing organisms by steering coordinated actions between groups of cells. In order to do so, they developed the ability to motivate other cells into action, a motivation quite possibly reliant on the already present reward/punishment matrix and definitely dependent on chemical signalling. Although neurons famously transmit their information along their cellular body through electricity, any communication between neurons instead comes down to the exchange of chemical signals called neurotransmitters. These neurotransmitters allow cells to communicate amongst each other in a meaningful way, enabling them to encourage or discourage action (such as the firing of a neuron) under particular circumstances. This allowed for a new level of adaptability, but also required a new level of organisation. The next chapter deals with this new organisational level: reward and punishment in organisms with brains.

Literature

Baker M.D., Wolanin, P.M. & Stock, J.B. (2005). Signal transduction in bacterial chemotaxis. *BioEssays 28* (1) 9-22.

Chatterjee, S., da Silveiram R.A. & Kafri, Y. (2011). Chemotaxis when bacteria remember: Drift versus diffusion. *PloS Computational Biology 7* (12) Special section 5. 1-8.

Craig, A.D. (2003). A new view of pain as a homeostatic emotion. *Trends in neuroscience 26* (6) 303-307.

Damasio A. & Carvalho G.B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience 14* (2) 143-152.

Decety, J. & Svetlova, M. (2012). Putting together phylogenetic and ontogenetic perspectives on empathy. *Developmental Cognitive Neuroscience 2* (1) 1-24.

Fields, R.G. & Stevens-Graham, B. (2002). New insights into neuron-glia communication. *Science 298* (5593) 556-562.

Gottfried, J.A. (2011). Preface. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Chicago, 2011).

Gottfried, J.A. & Wilson, D.A. (2011). Chapter 5: Smell. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

Gregory, F. (2008). Chapter 17: New ideas about life and its past. *Natural science in western history* (Boston, 2008).

Grosberg, R.K. & Strathmann, R.R. (2007). The evolution of multicellularity: a minor major transition? *Annual Review of Ecology, Evolution, and Systematics 38*. 621-654.

Hathcock, J. N. (1989). High nutrient intakes – the toxicologist's view. *Symposium Upper limits of nutrients in infant formulas* (November 7-8, 1988) 1779-1784.

Jacobs, D.K, Nakanishi, N., Yuan, D., Camara, A., Nichols, S.A. & Hartenstein V. (2007). Evolution of sensory structures in basal metazoa. *Integrative and Comparative Biology 47* (5) 712-723.

Kalat, J.W. (2004). *Biological Psychology; 8th Edition* (Belmont, 2004) 53-59.

Lewis, James L. III (October 2013). Acid-base regulation. http://www.merckmanuals.com/professional/endocrine_and_metabolic_disorders/acid-base_regulation_and_disorders/acid-base_regulation.html?qt=&sc=&alt. *The Merck manual* (retrieved 2 June 2014).

Magnab, R.M. & Koshland, D.E. Jr. (1972). The gradient-sensing mechanism in bacterial chemotaxis (temporal gradient apparatus/stopped-flow/S. Typhimurium/motility tracks/memory). *Proceedings of the National Academy of Sciences of the United States of America 69* (9) 2509-2512.

Murray, E., Wise, S. & Rhodes, S. (2011). Chapter 4: What can different brains do with reward? In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

Park, S., Wolanin, P.M., Yuzbashyan, E.A., Silberzan P., Stock, J.B. & Austin R.H. (2003). Motion to form a quorum. *Science 301* (5630) 188.

Chapter 3: Reward and the Brain

When humans act, they often do so because they *feel* the drive to do so. Unlike computer programs, which automatically and unequivocally act according to their software protocols because they *must*, humans act because they, at one level or another, *want* to. They have been motivated to undertake actions and from the results of these actions they learn what to do better in future situations. Unlike computer actions, which are set in stone and often not open to any kind of value judgement, human actions have to be linked to values in order to be adaptive.

When humans learn, they can do so passively or actively. Passive learning occurs at a mostly subconscious level and is strongly dominated by rewards and punishments experienced as the result of, or in co-occurrence with, actions. Humans do not even need to be consciously aware of these signifying factors for them to still play a role in their learned behaviour. Motivation, driven by reward and punishment, therefore plays a key role in passive learning. Active learning, on the other hand, is a much more conscious experience. When it comes down to memorising a list of German verbs or learning how to fix your car by watching instructional videos posted on YouTube, conscious motivation plays a central role. Knowledge that failure to learn the verbs may for instance result in a bad grade in school or communication errors with a vital business partner can produce negative, or avoidance motivation. On the other hand, knowing that learning to fix your car can save a lot of money on repairs or allow for a purposeful pastime can provide positive, approach motivation. Fear of the consequences of failure and the anticipation of the benefits of success are the motivating factors for undertaking action and deciding which action to take.

As learning is an important part of our intelligence, the crucial part in my view as explained in Chapter 1, and evaluation is crucial to determining what to learn and why, reward and punishment play a very large role in human intelligence. After exploring the implementation of positive and negative evaluation at the cellular level, it now is time to explore reward and punishment at the organism level. It will come as no surprise that the organ most involved with reward is also the organ widely regarded as the seat of our learning intellect: the brain. So how does processing reward benefit from the presence of a brain? Centralised information processing, combined with a comparative function, can allow complex organisms to make complex choices based on the multitude of reward signifiers they detect. The brain provides the ability to integrate information from multiple sensory organs and can, through reward learning, establish reward-expectancy. Due to the fact that a brain allows for projection over time, more commonly known as the ability to plan (no matter how weakly developed in some organisms), it can also decide not to engage in reward-driven behaviour by postponing it to a later moment or abandoning it entirely if other options also offer motivation. In the most limited case, a centralised brain allows merely for the selection among several rewarding alternatives, in a wider case, the brain can make long term decisions based on current and prospective rewards and punishments. In the more extreme cases, a brain can even use the meaning provided by reward to give meaning to experiences far removed from the primitive death/survival paradigm it has been founded on, such as appreciating music or architecture.[81] In the following chapter, I will discuss how this is possible.

Language use

“Feelings” and “emotions” are terms that often come up when discussing reward outside of a Behaviourism-context. I will use these terms because I do not agree with Behaviourism, which has fallen into disfavour in the academic world. In fact, I think that humans are not the only type of animal to experience emotions and have feelings: other animals with central nervous systems almost certainly experience them as well. I believe this to be the case based on the large overlaps in both physiological make-up (which I will provide some evidence for in the section on Reward in the Brain) and behavioural components which are too similar to be ignored. That said, there is of course a chance that animals do not experience all the emotions and feelings we have, experience them differently or even experience feelings or emotions we don't have (we know for a fact that many animal senses cover different parts of spectra than ours, or are even different senses entirely, such as a shark's electro sense). Luckily, only the presence of feelings is required, not their exactly identical nature.

Regretfully, “feelings” and “emotions” don't have strongly separated definitions in the literature. Although the use of the word “emotions” always includes the meaning of “affect”, which implies the creature or its actions were affected, the word “feeling” is used more ambiguously: sometimes it is used as a synonym to “emotion”, but at other times it is purely used to describe “unaffective” sensory input. To put it in other words, emotions “move” the brain while feelings may move the brain, but may also simply “inform”. This difference can be illustrated by appealing to

---

[81] Murray, E., Wise, S. & Rhodes, S. (2011). Chapter 4: What can different brains do with reward? In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

everyone's everyday experience of sight: seeing is generally not experienced as affective, until you see something that holds affective value or your eyes are oversaturated with light such as when looking directly at a bright light source. Another example, this time from the neuroscientific literature, is the possible removal of the unpleasant component from the feeling of pain. It is possible to only experience the sensory information that signals pain, without experiencing the unpleasant feeling that accompanies it.[82] In the following chapter, I will treat feelings as having an affective value unless otherwise specified. Non-valued sensations I will call exactly that: sensations. The experience of pain and other affective sensations that are usually called feelings, I will refer to by the more customary term of “feeling” although they could also be called an “emotion” instead.[83] Due to the nature of the following chapter, which focusses solely on learning through reward and punishment, the non-affective “sensations” will be largely left out of the discussion.

Homeostasis

Just like single-celled organisms, multicellular, complex organisms with a central nervous system (CNS) need to maintain homeostasis (see Chapter 2) not just at a cellular level but also at the organism level. Because the brain is responsible for most behaviour in CNS-organisms, homeostasis needs to be represented in the brain in order to allow adaptability to account for survival values. Homeostasis is indeed represented in the brain by brainstem structures that monitor internal homeostasis through the bloodstream and lymphatic system, and guide automated internal actions as well as activating higher functions when automated action is insufficient for restoring homeostasis. A second channel of homeostasis monitoring is provided by the sensory neurons distributed throughout the body which, among other things, monitor light intensity and colour (external) or signal the occurrence of physical damage (internal). Thanks to the brain's representation of homeostasis, there are two prime motivations for human behaviour:

- **Internal motivation**, which consists of *drives* such as hunger. Drives are triggered by the need to maintain bodily homeostasis which ensures short-term survival, as well as triggers that provide more general pro-fitness such as maintaining muscle tissue through use. Drives

---

[82] MacDonald, G. & Leary, M.R. (2005). Why does social exclusion hurt? The relationship between social and physical pain. *Psychological Bulletin 131* (2) 202-223.

[83] Craig, A.D. (2003). A new view of pain as a homeostatic emotion. *Trends in neuroscience 26* (6) 303-307.

also take care of such long-term survival necessities such as procreation. Drives require the detection of internal homeostatic disruption. A possible example is the reduction of salt levels in the bloodstream, which triggers a drive to ingest salt.[84]

- **External motivation**, which consists of *incentives* such as external signifiers of positive or negative factors that require creature action to obtain or avoid.[85] External motivation relies on detecting advance warnings of imminent internal danger or potential reward that requires action. It is an extra layer of adaptability that allows for earlier responses. Although the chemical or other triggering signifier comes from an external source, the associated motivation comes from within the body. A possible example is the detection of the presence of a predator by sight, sound or smell, where the signifier is external, but the drive to avoid pain and death is internal.

As discussed in the previous chapter, these two prime motivations are built upon the two primordial consequences that underlie all motivated behaviour of living organisms: death and survival. Both serve as a grounding point for a variety of rewards and punishments, thanks to their representation in homeostasis. Some examples for how they affect our feelings are in order.

The Death-consequence:

- Pain is the predictor and reward-signal of death. Pain indicates damage, or impending damage, at the cellular level as well as the organism level which may pose a threat to the physical survival chances of the pain experiencer. The strength of the signal is often an indication of the amount of damage sustained and correlates to the increased chance of death. Of course the feelings associated with the Death-consequence are not limited just to pain. Other feelings include hunger, thirst, temperature and itch, which all act as immediate motivators.[86] It is worth noting that the experience of pain consists of two separate components: the sensation of pain and the affect of pain. While pain sensation is gathered by the pain receptors present throughout the body, to inform the brain about ongoing tissue damage, the affect of pain harbours the actual motivation part. This is the uncomfortable

---

[84] Gray, P. (2002). *Psychology; Fourth Edition* (New York, 2002) 188-189.

[85] Gray, P. (2002). *Psychology; Fourth Edition* (New York, 2002) 188-189.

[86] Craig, A.D. (2003). A new view of pain as a homeostatic emotion. *Trends in neuroscience 26* (6) 303-307.

feeling that accompanies pain.[87] I will come back to this later on. Pain in humans is an emotion that motivates behaviour to re-establish homeostasis or at least to prevent further damage.[88] It functions as a reactive motivator.

- Fear is an emotion that drives humans to avoid pain and other death-signifiers, and with them death. Fear indicates increased chances of pain and is instrumental in avoiding pain, it therefore functions as a predictive motivator. It is grounded by being valued through pain, which is then grounded in death.

The sight of leprosy-sufferers with their missing fingers, toes, feet, hands or even limbs is all too familiar from images encouraging donations to support the treatment of leprosy. What many people don't know is the way in which leprosy causes the destruction of body parts: rather than causing them to fall off directly, leprosy causes loss of sensation by damaging nerves. The loss of sensation that follows from this makes leprosy-sufferers insensitive to those vital warning signs (called pain) that prevent and indicate damage to tissue. This leads to secondary infections which do most of the visual and structural damage.[89] This illustrates the importance of pain-receptors and pain avoidance in order to avoid serious damage and death.

The Survival-consequence:

- Pleasure is the predictor and signifier of survival. Pleasure indicates an improvement to the physical survival chances of the organism or its DNA and is grounded in survival and procreation. The amount of experienced pleasure is indicative of the survival-value. Pleasure is designed to motivate survival promoting behaviour and acts as a reactive motivator. Example types of pleasure are a sweet taste when hungry, salt when salt-deprived, and the satisfaction achieved through acts of procreation.
- Attraction is the drive to approach pleasure and with it survival. Attraction indicates increased chances of pleasure and it mostly motivates people by the feelings of happiness it promises they experience when giving in to that attraction. Attraction functions as a

---

[87] MacDonald, G. & Leary, M.R. (2005). Why does social exclusion hurt? The relationship between social and physical pain. *Psychological Bulletin 131* (2) 202-223.

[88] Craig, A.D. (2003). A new view of pain as a homeostatic emotion. *Trends in neuroscience 26* (6) 303-307.

[89] American leprosy missions (2014). Leprosy frequently asked questions. http://www.leprosy.org/leprosy-faqs/. American leprosy missions (retrieved 6 March 2014).

predictive motivator. It is grounded by being valued through pleasure, which is grounded in survival.

Pain and pleasure

Pain and pleasure are quite likely two different signals given off by the reward centres. So unlike the attributes of cold and heat in physics, where cold is simply a lack of heat, pain is not the absence of pleasure or vice versa. This is visible, for instance, in the hormonal signalling of pregnant females. Some hormones ready the female for maternal behaviour by decreasing fear and avoidance of infant-related stimuli, while others increase attraction towards infant related stimuli.[90] As both hormone sets positively impact the female attitude towards infants, one by decreasing the bad, the other by increasing the good, two different approaches can modify behaviour allowing for some system redundancy. Similarly, a particular neurotransmitter opposition is posited to exist between dopamine and acetylcholine, where the first encourages approach, while the second fosters avoidance of substances.[91] Other evidence comes from the field of reinforcement learning, where it has become clear that learning from positive and learning from negative feedback is at least separable in some cases, suggesting two different signals.[92] Other differences lie in their differing effects: unlike the pleasure signal, pain signals interrupt ongoing behaviour.[93] They are also much more expedient at promoting quick learning and quick responses aimed at terminating, reducing or escaping the source of threat.[94] In order to facilitate this, learning through pain appears to take a different and quicker path through the amygdala as well, resulting in quicker but less damage-resistant learning than the more pleasurable path.[95]

That said, it is important to realise that the exact relationship between pain and pleasure is very complicated and still far from understood. From personal experience, most humans will most

---

[90] Decety, J. & Svetlova, M. (2012). Putting together phylogenetic and ontogenetic perspectives on empathy. *Developmental Cognitive Neuroscience 2* (1) 1-24.

[91] Hoebel, B.G., Avena, N.M. & Rada, P. (2007). Accumbens dopamine-acetylcholine balance in approach and avoidance. *Current Opinion in Pharmacology 7* (2007) 617-627.

[92] Fellows, L.K. (2011). Chapter 16: The neurology of value. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

[93] Eccleston, C. & Crombez, G. (1999). Pain demands attention: a cognitive-affective model of the interruptive function of pain. *Psychological Bulletin 125* (3) 356-366.

[94] Eisenberger, N.I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience 13* (6) 421-434.

[95] Moscarello, J.M. & LeDoux, J.E. (2013). The contribution of the amygdala to aversive and appetitive Pavlovian processes. *Emotion Review 5* (3) 248-253.

likely admit that it is quite possible to have mixed feelings towards a particular object or subject. From this experience we can posit that pain and pleasure are indeed not the same system, but they are certainly related. A positive factor and an equal negative factor do not cancel each other out, but instead result in a mix bag of emotion. Exploration of human subcultures such as sadomasochism, demonstrates that it is even possible to experience painful sensations as pleasurable and vice versa, provided enough retraining of the reward system has taken place.

Another distinction can be found in the relativeness of pleasure-values in particular. Where some kinds of pain are negative across the board (such as the pain that results from mutilation, barring extreme corner cases), pleasure value is much more dependent on the internal state of the organism.[96,97] Of course, quite a few pain types are also situation-dependent, such as the discomfort and pain that may be associated with temperature-sensing: an icy bottle is likely to be experienced as unpleasant or even painful on a cold day, while it may be a sweet release on a hot day.[98]

All of this only goes to show how important homeostasis is for these basic feelings and emotions. The very purpose of life is to survive and procreate, while avoiding death. Maintaining homeostasis is an extremely important part in this and allows for the development of affective values. To see how brains use homeostasis and feelings to improve biological adaptability, we should continue onto learning.

Emotion and learning: assigning value to experience

There are at least two important methods through which the reward system provides motivation: hedonic impact and incentive salience.[99]

Hedonic impact refers to the direct effects of contact with a particular substance, for instance the consumption of a sandwich. If this sandwich satisfies a particular short-term drive, such as hunger, or has other beneficial homeostatic effects such as restoring salt-levels, the reward system will release neurotransmitters that give off a pleasurable sensation. This may be just a good taste which is an increase of the pleasure factor, or it can be combined with a quenching of the negative

---

[96] Fellows, L.K. (2011). Chapter 16: The neurology of value. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

[97] Gray, P. (2002). *Psychology; Fourth Edition* (New York, 2002) 188-189.

[98] Craig, A.D. (2003). A new view of pain as a homeostatic emotion. *Trends in neuroscience 26* (6) 303-307.

[99] Berridge, K.C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology 191* (3) 391-431.

feelings associated with hunger, a decrease of the punishment factor.[100] This process rewards the organism for undertaking good actions by both positive affirmation and the reduction of negatively charged drives. Hedonic impact is more commonly referred to as how much organisms "like" something and it can be positively impacted by hunger or other drives that signal bodily deprivation.[101] Hedonic impact is the direct experience of a reward or punishment and can be greatly increased if it satisfies internal homeostatic shortages. It is directly coupled to, and valued by, homeostasis, which is grounded on the survival/death-consequences that underlie all living creatures.

Incentive salience is the motivational power of pre-existing knowledge about rewards associated with undertaking a particular action, such as the consumption of a particularly tasty hamburger. Because the previous reward was so good, the brain has associated a good reward with a particular input, motivating the creature to seek out that particular pleasure when confronted with it. This mechanism is predictive and comes into play after learning an association. Because a previous experience has been good, the brain assumes that repeating it will also be beneficial, motivating new goal-directed behaviour. This mechanism is responsible for acting on signifiers that remind the organism of a particularly tasty (rewarding) sandwich, even when it's not that hungry, or seeking out a particular drug, even when the chemical dependency has been broken.[102] Incentive salience is more commonly known as how much creatures "want" something,[103] it relies on consciously or subconsciously remembered experience and produces an anticipated reward. It is triggered by external factors and provides external motivation to undertake action.

These two methods combined have a great impact on organism adaptability. As has become apparent, incentive salience is created from hedonic impact thanks to a learning process. During interactions of the organism with its environment, the hedonistic, positive factor becomes associated with stimuli indicative of the new positive situation. These stimuli then become a trigger for reward-expectancy themselves through a process called reward learning. The same, but inverted, goes for negative experiences and negative stimuli. A negative stimulus triggers a negative reaction from the intrinsic drives for self-preservation. This negative reaction is linked to any present

---

[100] Craig, A.D. (2003). A new view of pain as a homeostatic emotion. *Trends in neuroscience 26* (6) 303-307.

[101] Gray, P. (2002). *Psychology; Fourth Edition* (New York, 2002) 188-189.

[102] Saunders, B.T., Yager, L.M. & Robinson T.E. (2013). Cue-evoked cocaine "craving": role of dopamine in the accumbens core. *The Journal of Neuroscience 33* (35) 13989-14000.

[103] Gray, P. (2002). *Psychology; Fourth Edition* (New York, 2002) 188-189.

signifiers through reward learning, providing them with incentive salience to take preventive measures before actual harm has been done.

The learning mechanism through which this occurs is called reward learning and is presumably governed by its own hormonal associates, although which ones impact which part of the process still remains unclear.[104] In fact, the body has a multitude of neurotransmitters associated with reward and punishment motivation, of which most functions, interactions and other inner workings are far from unravelled.[105,106]

Through hedonistic impact, brains learn the survival or death value of eating particular foods or taking other particular actions. Through reward learning they are able to store these positive or negative values in memory, complete with situational information such as salt-deprivation (eating large quantities of salt when you are not salt-deprived is not a good idea, and the brain should make sure the new memory does not encourage this behaviour). Observations that relate to this experience are also stored with emotive value and become reward signifiers themselves. Whenever the creature then observes the signifiers, they can trigger incentive salience which motivates the organism to adapt its behaviour. In the end, incentive salience is derived from earlier experienced hedonic impact, which is derived from homeostatic monitoring, which is ultimately derived from the survival/death mechanic.

Emotion and learning: what to remember?

Reward and punishment not only works on direct action, or even just by creating incentive salience. It has other memory properties too. When a desired behaviour is taught to animals in the lab, rewards and punishment are often used as studied variables, as well as training aides. In doing so, scientists discovered that reward and punishment have not one, but two major reinforcement-effects on memory.[107]

Reward/aversion-learning is strongly coupled to the approach/avoidance effect detectable in all organisms. In my opinion, reward/aversion is necessary for triggering approach/avoidance

---

[104] Berridge, K.C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology 191* (3) 391-431.

[105] Barbano, M.F. & Cador, M. (2007). Opioids for hedonic experience and dopamine to get ready for it. *Psychopharmacology 191* (3) 497-506.

[106] Berridge, K.C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology 191* (3) 391-431.

[107] White, N.M. (2011). Chapter 3: Reward: What is it? How can it be inferred from behaviour? In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

effects in multicellular organisms, although some scientists prefer to strongly separate the two by declaring reward/aversion to be a conscious process, while approach/avoidance does not have to be.[108] I disagree with reward or aversion being something experienced on a conscious level, due in part to the myriad of rewards we take on a daily basis that reinforce our behaviour without our consciousness noticing them. While they are generally experienced unconsciously, these small rewards can become conscious if we pay particular attention to them, but regardless of conscious attention, their reinforcing aspects work. Regardless, positive or negative signifiers will produce approach or avoidance reactions.[109] This is the method of learning that has been described as **reward learning** in the preceding section.

In humans as in animals, another interesting learning process takes place on the basis of reward/punishment. Many people are able to recall where they were and what they were doing when they heard of the sudden death of a loved one even though their location and activity in all but the most extreme cases was unrelated.[110] This form of traumatic memory can even occur with more impersonal but still emotionally impactful events, such as the assault on the Twin Towers in 2001. The reason for this is that events of reward significance strongly improve memory in biological organisms. This is not limited to humans or even to traumatic events. The learning curve of rats can be accelerated by applying punishment such as shocks, or rewards such as food for a hungry rat, just after or prior to the training task. A very interesting mechanic lies hidden in the fact that it doesn't have to be a reward-value that correlates to the training task. Reward and punishment are interchangeable when training a rat to, for instance, find food in a maze. A rat that has walked into a corridor without food (a negative result) before being removed from the maze and shocked or given food, will in *both* cases better remember that that particular corridor is empty. This may seem counterintuitive, but it is a strong indicator that reward-signifiers reinforce memory directly.[111] This method of learning is called **memory modulation**. The unrelated reward-significant context reinforces the whole process of memory creation, including the parts that had no actual relevance to the reward.

There is a strong hint here that allows for some philosophising how reward works within

---

[108] White, N.M. (2011). Chapter 3: Reward: What is it? How can it be inferred from behaviour? In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

[109] Avoidance includes freezing.

[110] Actions responsible for the death of the lovedone obviously do not qualify.

[111] White, N.M. (2011). Chapter 3: Reward: What is it? How can it be inferred from behaviour? In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

complex organisms. Rather than being specifically coupled to only the circumstances and instances that actually contributed to its occurrence, reward is coupled to the state of the brain and a period of time, even when the gained reward obviously has nothing to do with the learned information. Although this may at first seem odd, it actually makes perfect sense. The brain operates in a world where there are many uncertainties. Its function is to improve survivability by strongly increasing adaptability to organism surroundings: it must connect actions to consequences in order to provide the values on which it can make decisions in the future. However, due to the uncertainty inherent in the world, on which inner workings the brain is largely clueless, the brain cannot know what particular factor led to the experience of a reward-signifier. Although it is possible that the eating of that mealy piece of fruit led to stomach cramps, it could be due to other environmental causes, such as a vile stench that was present, the colour of the walls, a stomach virus or, to name something both invisible and extreme, radioactive radiation. To complicate matters further, a time factor may be involved. A delay between action and the valued effect it facilitates is present in a great many of stimuli-reward relations. In order to learn the relevant combination, the brain must therefore learn all the potential signifiers, across time and space, on the assumption that on repeated tasks eventually the real signifier will be the most enforced. I will discuss a possible mechanism through which it does this later on.

So, because the brain cannot know in advance which information is actually relevant to the emotional experience, it tries to store all information that could be relevant. Furthermore, because a reward signal was received, the brain knows that there was something worth remembering (be it something bad or something good). The stronger the reward signal, the more important the creation of a strong memory. This explains why a hungry mouse that receives a reward signifier while looking for food, will better remember that the corridor was empty if it receives a positive or negative signifier in close time-proximity to the event, regardless of what kind of signifier it received. In order to learn from encountered rewards and punishments, the brain must cast a wide memory net. As it can't be sure what actions and consequences are connected precisely, it must enforce a wide variety of connections and value them in the hope that the right connection is among them and will be reinforced more often than the others. This process is responsible for many memory effects, illuminating the massive role that reward plays in brain adaptability. The brain is there to link action to consequence and reward learning (evaluation) is how it does it. It's now time to look at reward learning in psychological practice.

Conditioning: subconscious learning through reward

Conditioning, the most well-known example of reward-learning, was made famous by the experiments done by the Russian Ivan Petrovich Pavlov (1848-1936). Like other important scientific breakthroughs, his discovery was something he initially considered a problem with his experiment: the fact that the dogs he was experimenting on started to salivate in response to signals that preceded the actual administration of food, hindered his research on digestive reflexes in these animals. However, his experiments on dogs soon revealed their natural precondition to attach a pre-existing reflex to new conditioning stimuli.[112] If, for example, the sound of a bell preceded the delivery of food to the dog, the dog would very rapidly learn to associate the two, leading to pre-emptive salivation as soon as the bell was rung.

This form of conditioning, where a stimulus that previously did not elicit a reflexive response starts eliciting a reflexive response after being paired with a stimulus that already elicits that response, is called *classic conditioning*. Other tests have revealed this mechanic to be very present not just among animals, but among humans as well. If, for instance, humans are exposed to a bright flash of light, a negative stimulus that triggers a defensive mechanism in the muscles surrounding the eyes, paired with a clicking sound, the clicking sound will start triggering a blinking response without the flash of light being present.[113] Classical conditioning ties predictive outside stimuli to behaviour. It is, in essence, reactive towards outside stimuli and geared towards providing a quick and adequate response.

A second form of conditioning through the use of rewards or punishments is *operant*, or *instrumental conditioning*. In the case of operant conditioning, the consequences of a response decrease or increase the likelihood that the response occurs again. For instance, when behaviour, such as touching a pointy cactus, is immediately followed by physical pain, it is less likely that that behaviour will occur again.[114] Operant conditioning ties consequences to behaviour, when that behaviour has been, or appears to have been, instrumental in causing the consequences. It is designed to increase the effectiveness of actions initiated by the organism itself, increasing its survival value and decreasing its risks of death.

---

[112] Gray, P. (2002). *Psychology; Fourth Edition* (New York, 2002) 99.
[113] Gray, P. (2002). *Psychology; Fourth Edition* (New York, 2002) 98.
[114] Gray, P. (2002). *Psychology; Fourth Edition* (New York, 2002) 98-99.

Reflexes are mediated by the nervous system as quick responses to disruptions of homeostasis. They are important for keeping organisms alive, as conscious, thoughtful activity is often too slow to correct for a sudden imbalance such as tripping. By responding quickly and automatically, reflexes prevent a lot of potential damage which is in accordance with maintaining homeostasis and adaptability. However, reflexes that do not actually prevent damage are bad. The same reflex can be dangerous in one situation, such as eliciting a fearful scream when a predator is sighted as that may draw its attention, while potentially beneficial in another, eliciting the same scream to alert others for a cooperative defence or retreat. As speed is still of the essence, the body has several ways to modify or suppress reflexive responses. It learns to associate outside stimuli with outside rewards or threats, as well as associating detrimental behaviour with negative consequences, and beneficial behaviour with positive consequences. Reflexes can even be suppressed or completely vanish when their relevance declines, an example of this is the suckling-reflex,[115] which rapidly loses survival value as the child ages beyond early infancy. In order to change reflexive behaviour, or rather in order to learn and adapt to the environment, the body harnesses the reward system. Avoidance of painful stimuli and consequences and the pursuit of pleasurable stimuli and consequences promotes certain behaviours and not others through a learning process completely dependent on the pain-reward system. Being able to condition reflexes is useful because it allows for the activation of counter-measures based on learned indicators that may appear before the actual stimulus has occurred.[116] That said, conditioning is not limited to reflexive behaviour. To understand the tight relation between physical consequences and emotions, it is useful to know that conditioning not only works with physical consequences, but also with emotions. Pairing an unconditioned stimulus with an existing fear or positive emotion causes conditioning along the same lines, showing that it is the value of the stimulus, not the stimulus itself that allows for affective pairing.[117] More on this will follow in the section on higher reward learning.

Unconditioning

Conditioned associations can also be removed again through a similar learning process. If a conditioned stimulus no longer predicts the unconditioned stimulus, it starts to lose its association

---

[115] Kaneshiro, N.K. (12 April 2013). Infant reflexes. http://www.nlm.nih.gov/medlineplus/ency/article/003292.htm. *Medline Plus* (retrieved 6 March 2014).

[116] Gray, P. (2002). *Psychology; Fourth Edition* (New York, 2002) 105.

[117] Gray, P. (2002). *Psychology; Fourth Edition* (New York, 2002) 104.

and the corresponding reflexive response through a process called extinction. Extinction is, perhaps surprisingly, not a process of forgetting a conditioned response, but instead one of learning to no longer display the conditioned response. Extinction is really a learned inhibition of the reflexive response, so another acquired information-instance, rather than the deletion of earlier association. This is evidenced by two remarkable occurrences: if the unconditioned stimulus has been extinguished, but then hasn't occurred for a long enough period, it may trigger the conditioned response again as it spontaneously recovers from the “unlearning”. Just as the cessation of unconditioned stimuli can weaken a conditioned response, the lack of responses to inhibit makes the inhibit-response weaker. Another proof is the immediate re-emerging of the conditioned response if the conditioned stimulus is paired once again with the unconditioned stimulus. Just a single pairing is enough to re-establish the conditioned response.[118]

The evolutionary and reward-system associations are clear. If a conditioned stimuli loses its learned value by no longer being associated to a reward/pain experience, the positive/negative experience associated with it needs to be suppressed to prevent unnecessary actions. In effect, the parts of the brain involved in error detection detects a negative discrepancy. It then uses reinforcement mechanics to reinforce the suppressing neurons, perhaps by releasing the associated negative value neurotransmitters. However, because the conditioned stimulus has been an effective predictor in the past, the brain retains its information just in case it becomes useful again in the future. Perhaps additional information is required to narrow down the predictive value of the conditioned response, or some other additional predictor can be found. It is therefore useful for the brain to inhibit the learned behaviour, rather than destroy it.

Conditioning also has a generalisation effect. Stimuli that resemble the conditioned stimulus, will also trigger the conditioned behaviour. The more they resemble the original stimulus, the more likely the conditioned behaviour is to occur and the stronger the reaction will be. This is likely part of the brains insecurity about the world and its actual states, which require it to cast a wide net to catch signifying stimuli. If, however, the resembling stimulus is an accurate predictor of a lack of unconditioned stimulus, it will be strongly discriminated from the effective unconditioned stimulus and will trigger no, or the inverse behaviour.[119]

Again, expanding the range of conditioned stimuli that are viewed as predictive is useful,

---

[118] Gray, P. (2002). *Psychology; Fourth Edition* (New York, 2002) 102-103.
[119] Gray, P. (2002). *Psychology; Fourth Edition* (New York, 2002) 103.

but only in so far as they will actually still predict the unconditioned stimulus. Once again, widening the selection criteria is useful for learning through a less than complete dataset, but when the widened criteria fail to predict, they are swiftly eliminated through a lack of reward-pain valence and error-detection. However, the wide net may also be due to the way in which reward learning works at the cellular level. I will address this in the section on reward at the neuronal level.

A case of subconscious learning

The best illustration, perhaps, on just how subtle our reward-pain systems influences our learning of the right behaviours when it comes to operant conditioning, can be found in an experiment by R. F. Hefferline et al in 1959. In this experiment, adult participants were listening to music that was occasionally disrupted by static noise, a very unpleasant experience. Some of the participants were told nothing about the interjected static, but were instead informed that it was an experiment on the effects of music on body tension, while others were informed on the static and the fact that it could be disengaged by a specific response and were tasked with finding out what that specific response was that would turn it off. Interestingly, both groups increasingly displayed the behaviour (a twitching of the thumb) that would cut the static short, yet neither group could inform the experimenters on what method performed the feat. Instead, the misinformed group merely reported a decrease in static, unwitting of their contribution to its reduction, while the informed group didn't know what they did to reduce it.[120]

No conscious puzzle-solving had solved the puzzle and many participants were not even aware that there was a puzzle to be solved. And yet, the puzzle was solved: the participants successfully reduced the unpleasant static without even being aware of doing it. They subconsciously learned the value of twitching their thumb in response to the bothersome static that was disrupting the music. This is a clear illustration of how painful experiences and the body's innate desire to avoid or reduce those, can and will trigger learning processes on even the subconscious level through the coupling of circumstances to results. The subconscious brain managed, with the help of the reward system, to interact with the environment, store successful interactions in a meaningful way (i.e. only using the memory when the static played) and alter its

---

[120] Gray, P. (2002). *Psychology; Fourth Edition* (New York, 2002) 110. Although later experimenters did criticise the experiment, the conclusions were upheld by later, more stringent experiments, such as:
Laurenti-Lions, L., Gallego, J., Chambille, B., Vardon, G. & Jacquemin, C. (1985). Control of myoelectrical responses through reinforcement. *Journal of the Experimental Analysis of Behavior 44* (2) 185-193.

behaviour, all so that an annoying disruption of auditory function would cease. The reward system had combined with unconscious action to improve organism adaptability without any higher brain function. It is now time to turn to some of the higher learning functions.

Higher learning

Reward and punishment is very important for higher level learning as well. A simple philosophical reflection on our internal learning will suggest as much. Whenever humans are in a learning environment, one or both of the reward system's main pathways are always present. When we learn foreign languages in school, we do so often because we face repercussions such as bad grades and reprimands if we do not invest the time to do so. Alternatively, knowing that learning a foreign language is an investment in our future survival[121] due to the expanded options and skillsets it enables, can provide positive motivation to learn them. The knowledge of future punishment or reward functions as a motivational crutch that motivates our actions. There is also a more primitive motivation at work. We may actively like or dislike the learning activity. Some people may enjoy learning about the rules of a complicated game as they find that to be fun, while others may hate having to go through a rulebook as they find that to be boring. Some have learned that the act of studying is inherently rewarding through association, while others have learned the opposite value: that studying is boring and to be avoided.

Likewise, in the classroom, a good teacher engages and educates their students not just by explaining carefully, or providing the correct exercises, but also by creating an interesting environment where participation is encouraged. Intrinsic learning, learning that engages the learner's natural interests, is very important and can be reduced in effectiveness by providing external motivations that are unsuited to the learner.[122] These more basic motivations compete with the higher order long-term reward mechanics in determining why we do what we do. It is no accident that the concept of reward and punishment comes back time and again in learning methods designed for humans.[123] Intrinsic reward and external reward are very important to human learning, because they signal what is important for the human to do and to learn.

---

[121] In western societies with a properly functioning social security, this "survival" is less about physical survival and more about quality of life and social status.

[122] Armstrong, J.S. (2012). Natural learning in higher education. In: Seel, N. M. (ed.). *Encyclopedia of the sciences of learning* (2012) 10p (page numbers unknown).

[123] Armstrong, J.S. (2012). Natural learning in higher education. In: Seel, N. M. (ed.). *Encyclopedia of the sciences of learning* (2012) 10p (page numbers unknown).

Reward and planning

So far so good. We have established that biological adaptability is irrevocably intertwined with reward and punishment in micro-organisms, as well as demonstrated that reward and punishment also have a significant impact on the behaviour and adaptations of multicellular organisms including humans. Reward and punishment play important roles in learning in biological organisms. However, involvement does not mean they are a prerequisite. As the purpose of this thesis is to illustrate the role of reward and punishment in not just learning, but also higher level intelligence such as we recognise in humans, it is important to return to the adaptability and intelligence debate offered in Chapter 1. In this chapter I proposed to replace the vague concept of *intelligence* with the somewhat more defined concept *massive adaptability* as presented by Jack Copeland. I proposed that massive adaptability is not an inherently different form of adaptability than regular adaptability, but rather a more complicated and complex form. Massive adaptability is, in my view, composed of layers of adaptability intertwined and intersected to form a more complex whole. From that statement I posed that intelligence as we recognise that in humans, is not just built on the building blocks provided by earlier life forms, but fundamentally constructed out of them. If this is true, and reward and punishment are truly fundamental, we should see a serious breakdown of higher order intelligent processes such as planning when the lower level processes such as reward and punishment break down.

Indeed, planning and decision-making suffer greatly when the reward system is damaged. Given that planning involves increasing risk the further ahead the brain projects and the role of emotions in decision-making becomes more and more prevalent the greater the uncertainty of the outcomes, emotion and projected emotion can be assumed to play an important role in the planning and execution of planned tasks.[124] Emotion, reward and punishment have been demonstrated to play an enormous role in complex decision-making,[125] including making decisions in social contexts.[126] We use past experience, memories laden with affective value by the reward system, to generate expectations about the future. This is visible in the activation of the same brain networks in both the

---

[124] Bechara, A. (2004). The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage. *Brain and Cognition 55* (2004) 30-40.

[125] Quartz, S.R. (2009), Reason, emotion and decision-making: risk and reward computation with feeling, *Trends in Cognitive Sciences 13* (5) 209-215.

[126] Rilling, J.K. & Sanfey, A.G. (2011). The neuroscience of social decision-making. *Annual Review of Psychology 62*. 23-48.

act of remembering and the act of predicting.[127] Insight, the ability to predict outcomes based on similar situations experienced in the past and learn from outcomes that did not fit the prediction, is dominated by the Orbitofrontal Cortex which also plays an important role in the experience of reward. Furthermore, substances that interfere with the default functioning of the reward system, such as the drug cocaine, also interfere with this insight learning function.[128] Malfunctioning of the reward-systems may result in the inability to make decisions, or the tendency to make deeply flawed ones. Patients who suffer from brain damage in these areas often are still capable of performing actions, but are no longer able to determine the why, which can lead them to contextually inappropriate behaviour.[129]

Without the reward-systems intact, human adaptability and with it intelligence suffers a tremendous hit. Many higher level functions we associate with intelligence become much harder or quite impossible to execute. Under the definition of bare-bone adaptability I have given in Chapter 1 and the assumption that massive adaptability is built from bare-bone adaptability, it will come as no surprise that the disruption of one of the four pillars (interaction, evaluation, storage and adjustment) will severely impact intelligence in general. That these functions do not cease entirely may be due to the widespread representation of reward in the brain: it is hard to knock-out all reward systems and representations, and the structures built on them before this happened are still influenced by evaluation's previous presence. Now that we have discussed the workings of reward in organisms, from conditioning to planning, it is time to explore its location in the brain.

Reward in the brain: feelings and emotions

In spite of many years of research, the exact location of reward in the brain is still unknown. Several regions have been strongly implicated in experiencing and processing reward, which at the very least demonstrates its importance to human cognitive function. In this section I will explore some of these regions and their functions.

The importance of reward and its associated feelings and emotions is perhaps best illustrated

---

[127] Fellows, L.K. (2011). Chapter 16: The neurology of value. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

[128] Lucantonio, F., Takahashi, Y.K., Hoffman, A.F., Chang, C.Y., Bali-Chaudhari, S., Shaham, Y., Lupica, C.R. & Schoenbaum, G. (2014). Orbitofrontal activation restores insight lost after cocaine use. *Nature Neuroscience 17* (8) 1092-1099.

[129] Fellows, L.K. (2011). Chapter 16: The neurology of value. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

by its presence in the most fundamental part of the brain, the brainstem. This part of the brain is located at the very base of the brain where the spinal cord meets the cranial cavity. It is the oldest component part of the brain and it is absolutely vital to survival. Damage to the brainstem has far-reaching consequences for human functionality and the region appears to be very much involved with feelings and emotion. Lesions in the dorsal (posterior) half of the upper brainstem are associated with severe conditions such as coma and vegetative states, where feelings and even sentience are abolished, while lesions in the ventral (anterior) half of the upper brainstem cause locked-in syndrome, where feelings and consciousness are preserved, but physical action is impossible.[130] Needless to say, the disruption of either set of functionalities is deadly when untreated. To demonstrate that the lack of feelings without the brainstem is not simply due to total brain shutdown, it is useful to note that inducing feelings in humans during experiments shows activation of brainstem structures. On a darker note, mammals whose cortex has been removed still exhibited coherent, goal-oriented behaviour consistent with feelings.[131] Electrical stimulation of certain brainstem regions can elicit behaviours consistent with emotional responses imbued with positive and negative valence in mammals. This also occurs in humans, with the added benefit that they can and will report experiencing the corresponding feelings. A key role for the brainstem appears to reside in triggering and supporting emotion and feeling.[132]

On the basis of this evidence, my earlier rejection of the idea that feelings are somehow limited to humans, or even to mammals seems all the more reasonable. Non-human mammals, birds, reptiles and even phylogenetically older species definitely and clearly display behaviour completely consistent with emotions and feelings. From a brain-oriented approach, these species show dramatic differences with humans at the level of the cerebral cortex. Although the danger of anthropomorphising animals is always present in biology and psychology, the brainstem, the presumed vital area for feelings, is essentially conserved in layout, design and function suggesting that feelings are not exclusive to humans and it is very likely that they have long been present in evolution. It seems fair to conclude that animals most likely share the basic feelings and emotions

---

[130] Damasio A. & Carvalho G.B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience 14* (2) 143-152.

[131] It is always difficult to establish feelings in animals as they are such a personal experience, but this philosophical quagmire can be easily expanded to throw doubt on the presence of feelings in fellow humans. A notion that seems to me to be quite absurd.

[132] Damasio A. & Carvalho G.B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience 14* (2) 143-152.

that we experience.[133] This is consistent with my position that feelings, as an extension of reward and punishment, are deeply rooted in adaptability. It seems that the oldest and most vital part of the brain has as one of its core tasks the integrated control of reward-feelings and their relation to signifiers.

However, feelings are so important for information processing that they are not limited to the brainstem alone. The limbic system has also been strongly implicated in correlating rewards with events and handling the experience. Many regions and structures within the limbic system play an important role and fire up when positive or negative experiences are encountered. The limbic system is also strongly connected to cortical regions which can reinforce or inhibit feelings and emotions to a certain degree.[134,135] Examples of limbic regions involved are the hypothalamus, the amygdala and the striatum, while at the cerebral cortex level the insula, the anterior cingulate cortex (ACC), the dorsal anterior cingulate cortex (dACC), ventromedial prefrontal cortex (PFC) and orbitofrontal cortex (OFC) been shown to play an important role in valuing outcomes with feelings and emotions.[136,137,138,139,140,141,142,143]

Although it is very difficult to pinpoint the exact location of the neuronal structures necessary for the development of feelings, only the brainstem structures seem absolutely vital, an

---

[133] Damasio A. & Carvalho G.B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience 14* (2) 143-152.

[134] Bush, G., Vogt, B.A., Holmes, J., Dale, A.M., Greve, D., Jenike, M.A. & Rosen, B.R. (2002). Dorsal anterior cingulate cortex: a role in reward-based decision making. *Proceedings of the National Academy of Sciences 99* (1) 523-528.

[135] Damasio A. & Carvalho G.B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience 14* (2) 143-152.

[136] Damasio A. & Carvalho G.B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience 14* (2) 143-152.

[137] Decety, J. & Svetlova, M. (2012). Putting together phylogenetic and ontogenetic perspectives on empathy. *Developmental Cognitive Neuroscience 2* (1) 1-24.

[138] Craig, A.D. (2003). A new view of pain as a homeostatic emotion. *Trends in neuroscience 26* (6) 303-307.

[139] Bush, G., Vogt, B.A., Holmes, J., Dale, A.M., Greve, D., Jenike, M.A. & Rosen, B.R. (2002). Dorsal anterior cingulate cortex: a role in reward-based decision making. *Proceedings of the National Academy of Sciences 99* (1) 523-528.

[140] Fellows, L.K. (2011). Chapter 16: The neurology of value. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

[141] Lucantonio, F., Takahashi, Y.K., Hoffman, A.F., Chang, C.Y., Bali-Chaudhari, S., Shaham, Y., Lupica, C.R. & Schoenbaum, G. (2014). Orbitofrontal activation restores insight lost after cocaine use. *Nature Neuroscience 17* (8) 1092-1099.

[142] Chikazoe, J., Lee, D.H., Kriegeskorte, N. & Anderson A.K. (2014). Population coding of affect across stimuli, modalities and individuals. *Nature Neuroscience 17* (8) 1114-1122.

[143] Schoenbaum, G., Roesch, M.R., Stalnaker, T.A. & Takahashi, Y.K. (2011). Chapter 15: Orbitofrontal Cortex and Outcome Expectancies: Optimizing Behavior and Sensory Perception. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

important clue that the neuronal reward system is as old as the brain itself.[144] A wide selection of brain structures are involved with the processing of feelings, which suggest they play an important role in the brain indeed. According to my thesis that reward and punishment are vital for proper adaptive behaviour, this result can hardly be called a surprise. Although the question of where reward and punishment is located still has an unsatisfying answer, the matter of how feeling evaluation may be established is even more perplexing. In the following section I will discuss how reward-learning may function on the neuronal level.

Reward learning at the neuronal level: a philosophical explanation

The reward-system has interesting effects at the microscopic scale. Release of reward-signalling hormones has been shown to trigger neuronal growth and specialisation in the brain areas associated with the external signal picked up by the senses. When a reward is for instance coupled with a particular sound, not only the part of the brain responsible for reviewing the reward is fine-tuned and plastic, the part of the brain that processes the raw signal gets modified by reward-growth as well. Through reward learning, the neurons associated with processing the relevant information are fine-tuned and the brain area adapts and can even expand. Even the primary sensory areas are therefore influenced by reward, responding to reward-information with appropriate growth.[145,146] This shows that reward has a direct promotional effect at the cellular level. Apparently neurons that receive rewarding chemicals are strengthened in their behaviour and growth in their area is encouraged.

This is not odd when you consider that neurons are still first and foremost cells. That means that they stem from cells that had their own internal positive and negative signifier matrix such as found in microorganisms (see Chapter 2). They are also very susceptible to changes in their environment.[147] It seems to me to be reasonable to assume that the positive/negative matrix presumed to exist in Chapter 1 has been preserved to facilitate intercellular communication,

---

[144] Damasio A. & Carvalho G.B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience 14* (2) 143-152.

[145] Weinberger, N.M. & Bieszczad K.M. (2011). Chapter 1: Introduction: From traditional fixed cortical sensationism to contemporary plasticity of primary sensory cortical representations. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

[146] Camalier, C.R. & Kaas, J.H. (2011). Chapter 9: Sound. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

[147] Cook, N.D. (2008). The neuron-level phenomena underlying cognition and consciousness: synaptic activity and the action potential. *Neuroscience 153* (3) 556-570.

especially since neurons have to travel and grow connections guided by chemical signals after they have been created. The ultimate consequence has also been preserved at the cellular level: neurons that fail to establish sufficient connections will suffer cellular death.[148] It is also known that the chemical neurotropic that prevents neuronal death also increases branching of incoming axons, which enables the creation of more connections, an important part of memory creation.[149] That means that neurons can be individually encouraged or discouraged to pursue certain kinds of action. I posit that it is this encouragement that tweaks neuronal connectivity, which results in enforcing beneficial connections, while increasing the inhibition on detrimental ones. Rewarding chemicals have a particular effect on neurons that are active or have recently been active: these neurons are encouraged in their behaviour.

Neurons that fire in beneficial circumstances are therefore being promoted again and again, which enforces behaviour at the macroscopic level. The arrangement of neurons into smaller networks that have their own specialisations also allows for learning associations in uncertain environments. Imagine a very restricted environment. In this environment only four different stimuli can be detected: a flash of light, a burst of sound, a smell and a touch. Say the “organism” living in this environment detects a flash of light, a burst of sound, and a touch, while it also experiences a reward (homeostasis is improved). The brain releases reward-transmitters and the neurons that fired for the light, sound and touch stimuli are reinforced with a positive value that puts their “significance” with respect to the experienced reward at (1). On a second trial, the organism experiences a rewarding sensation, but this time, the organism has detected a smell, a burst of sound and a touch. The sound and touch neurons will receive another encouragement which puts them at (2), while the smell neurons level with the light neurons which are not enforced by this encounter and stay valued at (1). A third encounter with the rewarding sensation seals the deal: the organism now detects a burst of sound and a flash of light, while experiencing the reward. The sound neurons are reinforced most strongly (3), while the other sensations trail at (2) or less, resulting in the outcome that the sound burst will now be the most potent signifier of the incoming reward (see Figure 3.1).

---

[148] Kalat, J.W. (2004). *Biological Psychology; 8th Edition* (Belmont, 2004) 109.
[149] Kalat, J.W. (2004). *Biological Psychology; 8th Edition* (Belmont, 2004) 111.

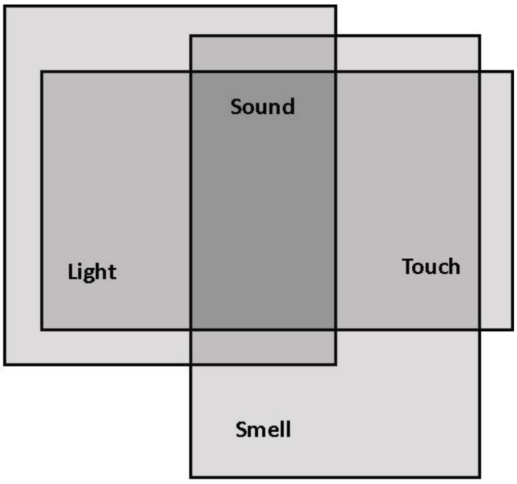

Figure 3.1: Overlapping reward signals. The resulting stimulation of neurons associated with light, sound, touch and smell after three trials paired with reward, where the sound was the actual signifier, while light, touch and smell were only randomly paired. Trial 1 featured a light, a sound and a touch. Trial 2 featured a sound, a touch and a smell. Trial 3 featured a light and a sound. All trials featured a reward paired with the sensory data. By simple up regulation of active neurons during reward-signified trials, the organism can determine that sound is the most important factor.

This mechanism can be helped along by the guided attention of the animal involved. Depending on the way in which an animal associates a reward with certain circumstances, the brain-area responds with growth or not. The way an animal processes information has a direct impact on the brain areas that are stimulated by reward to grow.[150] This is indicative of a process in the brain that can guide the transmission of the reward and preselect the brain-areas that are susceptible to reward stimulation. Let us now turn to the potential mechanisms that produce these reward-signals.

Value-assigner and Arbiter

On the basis of the information contained above, as well as the information gathered from Chapter 2, I will now posit two hypothetical reward-learning mechanisms that could explain the very important functions of reward-learning in the brain. The first is the reward/punishment matrix itself, which I will call the "Value-assigner". Some areas of the brain seem specialised in releasing positive or negative neurotransmitters that affect the rest of the brain. It is likely that this mechanism produces the reward/punishment values that give meaning to organism actions and sensory input. The Value-assigner signals other parts of the brain when they have presumably

---

[150] Weinberger, N.M. & Bieszczad K.M. (2011). Chapter 1: Introduction: From traditional fixed cortical sensationism to contemporary plasticity of primary sensory cortical representations. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

performed positively, or when they have performed negatively instead. This burst of value-giving chemicals also indicates what is important and what is not: events that trigger negative or positive reward will always be important to remember, so they can respectively be avoided or approached later on. I speculate that value signalling neurotransmitters always have a signal enhancing effect: either they enhance firing that encourages the firing of other neurons, or they enhance the firing that inhibits the firing of other neurons. Because reward signals are released around the time of the creation of the memory and the brain cells storing that memory through their connections are therefore active, the memory will be automatically reinforced because the cells themselves are strengthened in their connection.

The Value-assigner releases its signals when prompted by another mechanism that makes the actual comparison on the basis of which positive or negative are defined and signalled. The oldest mechanism that performs this function is most likely the homeostatic monitor, which may very well reside in the brainstem. This brain-mechanism is simply concerned with comparing past, current and ideal homeostatic states of the body. Based on the relations between the three, this brain component, which I will call the “Arbiter”, can decide whether the most recent actions or received external signifiers are to be associated with a positive, a negative or a neutral change. Most likely, the Arbiter has an internal representation of the ideal homeostatic values, as well as a representation of past measurements. It can then compare these to newly measured homeostatic values. This allows the Arbiter to judge whether:

- the new homeostatic values are better or worse than the old,
- the new homeostatic values require further action (because homeostasis is still not achieved).

The first part of this Arbiter judgement can form the basis for positive or negative signals to other brain cells: if homeostatic values have been improved, their activity has led to a positive outcome and they need to be encouraged to do the same behaviour again if the situation calls for it. The reverse goes for homeostatic deterioration, which needs to be inhibited. This can be done by sending negative signals to active cells or instead sending positive signals to cells that need to inhibit those active cells. This second option would be in concordance with the way unconditioning seems to work: a strengthening of inhibition rather than a weakening of excitation, perhaps inhibitory brain cells are rewarded by a different neurotransmitter. In this functionality the Arbiter takes on the role of action-evaluator.

The second part of the Arbiter judgement forms the basis for creature action. As discussed in Chapter 2, action can be detrimental to creatures unless there is something to gain. The Arbiter can, on the basis of homeostatic imbalance, prompt for action in the unbalanced category: there is a disruption in homeostasis and action is required to compensate for it. In this manner, the Arbiter functions as an action-driver and motivates the taking of actions through homeostatic monitoring.

Branching off from the first Arbiter-functionality, other options of using reward-learning can bloom as well. It is known that the brain features networks that track errors by comparing actual outcomes to expected outcomes.[151] Two brain regions hypothesised to have this functionality are the OFC[152] and the ACC, though, sadly, a lot of uncertainty about the roles of each brain region involved in reward still exists.[153] Perhaps it is these specialised Arbiters that are capable of directing reward to a more restricted area such as to explain the directed attention effect mentioned earlier (see: reward learning at the neuronal level: a philosophical explanation). The parts of the brain involved in an accurate outcome will receive a dosage of rewarding neurotransmitters, which at the cellular level informs the neurons that their action or inaction contributed to a beneficial outcome. This causes them to strengthen the connections that were involved. Because all neuronal connections that were involved in the beneficiary action where reinforced, this microscopic reward/punishment mechanism in effect reinforces the macroscopic behaviour, enforcing the strength of the prediction. If the prediction turns out to be in error, different neurotransmitters will be released, to discourage whatever action or inaction the particular brain cells have undertaken, in effect inhibiting the macroscopic behaviour and weakening the original prediction in a manner reminiscent of unconditioning.

Together, Arbiter-modules and Value-assigner modules could encourage macroscopic behaviour by monitoring homeostasis and then encouraging cellular activity. Behaviour at the creature level may thus be explained through simply neuronal adaptation thanks to reward-systems

---

[151] Murray, E., Wise, S. & Rhodes, S. (2011). Chapter 4: What can different brains do with reward? In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

[152] Schoenbaum, G., Roesch, M.R., Stalnaker, T.A. & Takahashi, Y.K. (2011). Chapter 15: Orbitofrontal Cortex and Outcome Expectancies: Optimizing Behavior and Sensory Perception. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

[153] Fellows, L.K. (2011). Chapter 16: The neurology of value. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

that are grounded in monitoring and maintaining homeostasis. Especially the Value-assigner module comes off as being extremely versatile and multi-purposable: releasing excitatory signals in brain areas can potentially be exploited by multiple Arbiters making use of the same Value-assigner system. In the following paragraph, I will discuss one such possible borrowing of such Value-assigner functionality in the human brain.

“Hacking” the Value-assigner

Perhaps the best way to show the multiple applicability of the reward system is to demonstrate how it may have been relatively recently repurposed in nature. One prime example of the malleability of the reward system can be found in an evolutionary adaptation residing in mammals, which is very strongly present in humans. Neuroscientist research has uncovered a strange relationship between affective physical pain and the affective pain caused by social emotions.[154]

Broad research during the 20th century has revealed the importance of social ties for the welfare and survival of pretty much all mammalian species. Unlike the young of most reptiles, mammal infants are generally completely dependent on other members of their species for their nutrition, protection and other care. Mammals living in groups also have the shared responsibility for gathering food, the care of infants and protection from predators, which is crucial to each individual's survival. This means that a threat, or actual damage, to social bonds can be just as dangerous as actual physical harm to the individual which explains why social bonds need to be protected.[155,156] An excellent way to motivate an individual to protect its social bonds is to wire this social survival mechanism into that age-old survival mechanism: the reward-system.

As social connections are broad and their survival impact can reach quite far, the definition of “social pain” must be taken broad as well. Social pain in the oncoming section is defined as: experiences that signal loss, or potential loss, of social connection or value. From an evolutionary standpoint of group dynamics, these losses of social connection or value indicate an increased survival risk. This means that both in situations where the subject receives (perceived) damage to social standing due to his own actions or lack of action, as well as in situations where social bonds

---

[154] MacDonald, G. & Leary, M.R. (2005). Why does social exclusion hurt? The relationship between social and physical pain. *Psychological Bulletin 131* (2) 202-223.

[155] Eisenberger, N.I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience 13* (6) 421-434.

[156] MacDonald, G. & Leary, M.R. (2005). Why does social exclusion hurt? The relationship between social and physical pain. *Psychological Bulletin 131* (2) 202-223.

are severed although the individual is clearly not to blame such as the unavoidable death of a “loved one”, the individual is likely to experience social pain. As these are negative factors that increase the chances of untimely death, they should be coded as negative experiences and may therefore be coupled to negative feelings such as pain. It makes sense for the body to wire the emotions of social loss through the physical pain system. After all, the pain system has been established to prevent damage by motivating the cessation of action, or promoting it. Trying to keep social bonds alive and satisfied with your behaviour limits social damage, and pain, the mechanism that motivates the prevention of physical damage, seems a useful fit to make this happen.[157] Social pain does not generally hijack the entire pain experience. It especially triggers the uncomfortable part of the pain sensation, while leaving out most of the sensory somatic components, which is further proof that the reward-matrix associated with physical pain has been repurposed for social pain. However, in cases of extreme social pain, many people even report somatic symptoms such as an actual heartache, making the relation between social and physical pain even clearer. In the same manner, the pleasure system rewards us for establishing new positive social bonds or successfully maintaining current ones. Several hormones are released when we experience positive social interactions, most famous amongst them being oxytocin, which not only reduces social stress, but also decreases physical pain.[158]

There is plenty of evidence that a repurposing of the pain/pleasure matrix, the Value-assigner for short, is indeed what has taken place in biological organisms. Research has provided both direct and indirect evidence that experiences of social pain indeed rely on some of the same neurobiological substrates that are also vital for experiencing physical pain.[159]

I will start out with the indirect evidence. In natural languages around the world, the words to describe physical pain and social pain are quite often the same. In English for instance, physical pain analogies are often used for social pains, such as “hurt feelings”, “broken hearts” etc. This suggests a potentially universal overlap in the experience of social and physical pain.[160] Universal overlaps in language may well be due to universal overlaps of experiences, suggesting that the

---

[157] Eisenberger, N.I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience 13* (6) 421-434.

[158] Eisenberger, N.I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience 13* (6) 421-434.

[159] Eisenberger, N.I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience 13* (6) 421-434.

[160] Eisenberger, N.I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience 13* (6) 421-434.

underlying cause is not so much cultural as it is biological. Aside from being strongly associated with physical discomfort, experiences of social pain can be just as detrimental to actual individual health as physical pain: both chronic physical pain sufferers, and those who are socially isolated or have suffered societal loss are more likely to commit suicide than control groups.[161,162] Both physical as well as social pain are known to cause anxiety disorders, which are characterized by a heightened focus on possible harm and harm avoidance. Two concerns have been shown to lie at the root of anxiety disorders: fear for possible physical harm and the corresponding pain, and fear for possible social harm which includes rejection or evaluation.[163] Depression, another mental illness with strong social connotations, can be caused by both physical and social pain.[164] Another argument is that people asked to recall prior episodes of social pain report as much pain experienced as when they are recalling physical pain. Moreover, following the death of a loved one, a term representing strong social bonds and therefore high social “capital”, bereaved people not only report feeling intense psychological pain but often complain of somatic pain as well.[165] There's more. Patients who suffer from chronic pain also experience more social pain than control subjects. Patients who suffer from higher levels of daily pain, also have higher levels of anxious attachments and are more concerned about being rejected by others. The reverse is also true. People who are more sensitive to social pain also report more somatic symptoms and physical pain.[166]

This does not only apply to the sick, but also occurs in the healthy. People who report higher levels of physical pain following the same negative stimulus also suffer more when they are socially excluded.[167] Those who have a particular mutation of the mu-opioid receptor called OPRM1 polymorphism have both a heightened physical pain perception, as well as demonstrating higher social pain when faced with rejection, which is supported by having more detectable activity in the

---

[161] Tang, N.K.Y. & Crane, C. (2006). Suicidality in chronic pain: a review of the prevalence, risk factors and psychological links. *Psychological Medicine 36* (5) 575-586.

[162] Mee, S., Bunney, B.G., Reist, C., Potkin, S.G. & Bunney, W.E. (2006). Psychological pain: a review of evidence. *Journal of Psychiatric Research 40* (8) 680-690.

[163] Eisenberger, N.I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience 13* (6) 421-434.

[164] Mee, S., Bunney, B.G., Reist, C., Potkin, S.G. & Bunney, W.E. (2006). Psychological pain: a review of evidence. *Journal of Psychiatric Research 40* (8) 680-690.

[165] Eisenberger, N.I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience 13* (6) 421-434.

[166] Eisenberger, N.I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience 13* (6) 421-434.

[167] Eisenberger, N.I., Jarcho, J.M., Lieberman, M.D. & Naliboff, B.D. (2006). An experimental study of shared sensitivity to physical pain and social rejection. *Pain 126* (1-3) 132-138.

related brain areas.[168] Opiates, best known for their pain-relieving effects, not only reduce physical pain, but also reduce separation distress behaviours in non-human mammals and humans.[169] Finally, oxytocin, heralded as the love-hormone, is a social bonding hormone and is released when someone is being comforted by a loved-one. It reduces sensitivity to both social and physical pain.[170]

It seems clear that mechanics that monitor and motivate social interaction has indeed repurposed the Value-assigner module for its own reward-related purposes, allowing for successful motivated behaviour on the basis of older neurological scaffolding. The homeostatic values represented by somatic pain and pleasure, death and survival also underlie human social interaction. Likewise, the experience of social damage and physical damage overlap. This demonstration of malleability of the Value-assigner in the connections it values, pleads for the versatility of its implementation. It is becoming more and more feasible that projection of reward onto more abstract concepts, such as art, also becomes possible through associations with more physical valued processes. This clears the way for implementing the Value-assigner as a module that can be inserted into self-teaching Neural Net AI.

Conclusion

Reward systems are present throughout living organisms. As established in Chapter 2, even the simplest bacterial cells are able to detect changes in their environment and are capable of connecting these changes with an internal evaluation system that determines whether the change is bad or good and that steers behaviour accordingly. Although many of these affective values are instinctual associations, it is possible to unlearn this information as well as learning new affective values for new, or previously meaningless compounds. Bacteria can then undertake action, as well as learn additional circumstantial information that can improve its reward/punishment prediction. This capacity to detect environmental signals and connect them with internal values has allowed cells to start communicating and even cooperating with each other. With the emergence of complex multicellular organisms and central nervous systems, this traditional method of chemical signalling

---

[168] Way, B.M., Taylor, S.E. & Eisenberger N.I. (2009). Variation in the mu-opioid receptor gene (OPRM1) is associated with dispositional and neural sensitivity to social rejection. *Proceedings of the National Academy of Sciences of the United States of America 106* (35) 15079-15084.

[169] Eisenberger, N.I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience 13* (6) 421-434.

[170] Decety, J. & Svetlova, M. (2012). Putting together phylogenetic and ontogenetic perspectives on empathy. *Developmental Cognitive Neuroscience 2* (1) 1-24.

has become too slow and inaccurate. Specific cells specialised into what we now call neurons, which function as quick relays for important information. Through electrical signalling, these neurons can cause cells in different parts of the body to quickly act upon communication sent by sensory neurons.

As complexity grew further, a centralised governing system emerged. This “brain” was tasked with making macroscopic decisions to promote the health of the entire community of cells: the organism. In order to be able to do this, the innate cellular ability to distinguish reward and punishment was presumably utilised. Specific functions of the brain, generalised as the “Arbiter” were dedicated to evaluating the impact of the environment and organism actions on internal homeostasis, as well as comparing the outcomes of predictions made by the brain with the actual outcomes as perceived in the environment. These evaluations are based in monitoring homeostasis, which is itself based in the ultimate grounding: survival or death through natural selection. Which parts of the brain can act as Arbiters is still very much a topic of research. After judging whether the outcome represents an improvement or degradation of performance, the Arbiter then informs another department of cells with an evaluative function, which I generalised as the “Value-assigner” of positive or negative valence. These specialised neurons then release their rewarding signals, predominantly neurotransmitters, into the corresponding brain areas with a process that likely produces feelings in our day to day experience. Feelings produced in this manner function as internal representation and communication of value. They are rooted on internal consequences, provided by the homeostatic monitoring performed by the Arbiter, and grant meaning to internal information storage of organism interactions with its environment.

The mechanism through which the Arbiter and Value-assigner provide value to internal information storage works at the same level where information storage itself takes place: the connections at the neuronal level. Neurons that have been active when the Value-assigner produces its feelings-inducing transmitters, receive these signals which serve as encouragement for their behaviour, whether it has been the inhibition of other cells, or excitation of other cells. By delivering neuron-level reward every time a positive marker is encountered, all correlating actions and sensations are encouraged. By the randomisation of the environment, the neurons involved in performing and registering the relevant interactions are encouraged most, as they will be rewarded in all positive situations, which separates them from neurons that have only been activated due to chance simultaneous activation. This stronger selection of actually relevant neuron firings vs nonrelevant neuron firings can be more pronounced by allowing negative instances to send a negative, or inhibitory signal that will adjust the value wrongful connections. The possible modularity of the Value-assigner in reward-learning can be demonstrated by the use of physical pain reward circuits for social pain motivation, encouraging the use of the Value-assigner as a reward-matrix that can be accessed by several different networks in order to use its rewarding properties.

Upon this relatively simple mechanism of large scale, cellular-level rewards, it is possible to build great adaptability. Reward and punishment play a decisive role in the way humans and other central-nervous-system organisms learn. It promotes unconscious learning, better known as conditioning, as well as conscious “explicitly motivated” learning. It is this reward-system that allows organic creatures to learn so effectively, as well as giving an inherent valence and grounding to all learned information through the connection between homeostasis (the basis for reward and punishment) and survival and death. Reward, or value, stored as connection-strengths in neural memory, also plays an important role in predicting future outcomes and deciding preferences on short as well as longer timescales. When aspects of reward-learning are disabled, adaptable behaviour up to human levels of intelligence becomes disrupted, showing its importance in both bare-bone and massive adaptability. The reward-system, built on homeostatic monitoring and cellular communication, is both responsible for action prompting as well as providing action and perceptive valence. Reward-systems may therefore be a useful addition in self-training Neural Nets and may also strongly impact the philosophy of any AI built on them. Let us now turn to a first model of such an AI.

Literature

American leprosy missions (2014). Leprosy frequently asked questions. http://www.leprosy.org/leprosy-faqs/. *American leprosy missions* (retrieved 6 March 2014).

Armstrong, J.S. (2012). Natural learning in higher education. In: Seel, N. M. (ed.). *Encyclopedia of the sciences of learning* (2012) 10p (page numbers unknown).

Barbano, M.F. & Cador, M. (2007). Opioids for hedonic experience and dopamine to get ready for it. *Psychopharmacology 191* (3) 497-506.

Bechara, A. (2004). The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage. *Brain and Cognition 55* (2004) 30-40.

Berridge, K.C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology 191* (3) 391-431.

Bush, G., Vogt, B.A., Holmes, J., Dale, A.M., Greve, D., Jenike, M.A. & Rosen, B.R. (2002). Dorsal anterior cingulate cortex: a role in reward-based decision making. *Proceedings of the National Academy of Sciences 99* (1) 523-528.

Camalier, C.R. & Kaas, J.H. (2011). Chapter 9: Sound. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

Chikazoe, J., Lee, D.H., Kriegeskorte, N. & Anderson A.K. (2014). Population coding of affect across stimuli, modalities and individuals. *Nature Neuroscience 17* (8) 1114-1122.

Cook, N.D. (2008). The neuron-level phenomena underlying cognition and consciousness: synaptic activity and the action potential. *Neuroscience 153* (3) 556-570.

Craig, A.D. (2003). A new view of pain as a homeostatic emotion. *Trends in neuroscience 26* (6) 303-307.

Damasio A. & Carvalho G.B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience 14* (2) 143-152.

Decety, J. & Svetlova, M. (2012). Putting together phylogenetic and ontogenetic perspectives on empathy. *Developmental Cognitive Neuroscience 2* (1) 1-24.

Eccleston, C. & Crombez, G. (1999). Pain demands attention: a cognitive-affective model of the interruptive function of pain. *Psychological Bulletin 125* (3) 356-366.

Eisenberger, N.I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience 13* (6) 421-434.

Eisenberger, N.I., Jarcho, J.M., Lieberman, M.D. & Naliboff, B.D. (2006). An experimental study of shared sensitivity to physical pain and social rejection. *Pain 126* (1-3) 132-138.

Fellows, L.K. (2011). Chapter 16: The neurology of value. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

Gray, P. (2002). *Psychology; Fourth Edition* (New York, 2002).

Hoebel, B.G., Avena, N.M. & Rada, P. (2007). Accumbens dopamine-acetylcholine balance in approach and avoidance. *Current Opinion in Pharmacology 7* (2007) 617-627.

Kalat, J.W. (2004). *Biological Psychology; 8th Edition* (Belmont, 2004).

Kaneshiro, N.K. (12 April 2013). Infant reflexes. http://www.nlm.nih.gov/medlineplus/ency/article/003292.htm. *Medline Plus* (retrieved 6 March 2014).

Laurenti-Lions, L., Gallego, J., Chambille, B., Vardon, G. & Jacquemin, C. (1985). Control of myoelectrical responses through reinforcement. *Journal of the Experimental Analysis of Behavior 44* (2) 185-193.

Lucantonio, F., Takahashi, Y.K., Hoffman, A.F., Chang, C.Y., Bali-Chaudhari, S., Shaham, Y., Lupica, C.R. & Schoenbaum, G. (2014). Orbitofrontal activation restores insight lost after cocaine use. *Nature Neuroscience 17* (8) 1092-1099.

MacDonald, G. & Leary, M.R. (2005). Why does social exclusion hurt? The relationship between social and physical pain. *Psychological Bulletin 131* (2) 202-223.

Mee, S., Bunney, B.G., Reist, C., Potkin, S.G. & Bunney, W.E. (2006). Psychological pain: a review of evidence. *Journal of Psychiatric Research 40* (8) 680-690.

Moscarello, J.M. & LeDoux, J.E. (2013). The contribution of the amygdala to aversive and appetitive Pavlovian processes. *Emotion Review 5* (3) 248-253.

Murray, E., Wise, S., & Rhodes, S. (2011). Chapter 4: What can different brains do with reward? In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

Quartz, S.R. (2009). Reason, emotion and decision-making: risk and reward computation with feeling. *Trends in Cognitive Sciences 13* (5) 209-215.

Rilling, J.K. & Sanfey, A.G. (2011). The neuroscience of social decision-making. *Annual Review of Psychology 62*. 23-48.

Saunders, B.T., Yager, L.M. & Robinson T.E. (2013). Cue-evoked cocaine “craving”: role of dopamine in the accumbens core. *The Journal of Neuroscience 33* (35) 13989-14000.

Schoenbaum, G., Roesch, M.R., Stalnaker, T.A. & Takahashi, Y.K. (2011). Chapter 15: Orbitofrontal Cortex and Outcome Expectancies: Optimizing Behavior and Sensory Perception. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

Tang, N.K.Y. & Crane, C. (2006). Suicidality in chronic pain: a review of the prevalence, risk factors and psychological links. *Psychological Medicine 36* (5) 575-586.

Way, B.M., Taylor, S.E. & Eisenberger N.I. (2009). Variation in the mu-opioid receptor gene (OPRM1) is associated with dispositional and neural sensitivity to social rejection. *Proceedings of the National Academy of Sciences of the United States of America 106* (35) 15079-15084.

Weinberger, N.M. & Bieszczad K.M. (2011). Chapter 1: Introduction: From traditional fixed cortical sensationism to contemporary plasticity of primary sensory cortical representations. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

White, N.M. (2011). Chapter 3: Reward: What is it? How can it be inferred from behaviour? In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

Chapter 4: Modelling Motivated Artificial Intelligence

Reminder: I will use the capitalised "Neuron" for the constituents of Neural Nets, while the noncapitalised "neuron" refers to the biological cell type that inspired them.

Now that it is clear that value in biological organisms is derived from homeostasis and its link to the ultimate consequences of natural selection (death and survival) it is time to make a first draft of an AI that is capable of assigning values on its own. We can now take the first steps towards creating a model of an AI that is designed to associate positive or negative values with its actions on the basis of internal consequences. In the current chapter I will attempt to create the basics of Arbiter and Value-assigner functionality that can be integrated in a modular fashion with memory and action-decision mechanics. As creating an actual AI is beyond the scope of this thesis, I will only produce a theoretical, and, regretfully, incomplete model of an AI that learns and acts on the basis of homeostatic feelings.[171] I will dub this AI model "MAI", short for Motivated AI, because this AI will be driven by homeostatic disturbance backed by actual cessation of function, rather than performing actions because it's directly programmed to do so. MAI will not require an outside source to establish value for it, but will instead provide its own value on the basis of which it will be capable of adjusting. What I hope to show is that such an AI would be capable of displaying adaptable behaviour that is also firmly connected to reality through consequences.

As discussed in the section on Artificial Intelligence, adaptability boils down to four basic points:

- A "being" must be capable of **interaction** with its environment (requiring some form of perception and some means of altering itself or the environment),
- A "being" must be capable of **evaluating** its interactions with the environment,
- A "being" must be capable of **storing** these valued interactions, (more commonly known as having a "memory"),
- A "being" must be capable of **adjusting** its interactions **based on these values** attained through previous interactions/perceptions (more colloquially known as "learning").

---

[171] The word "feelings" in this context is not meant to imply emotions, only the negative or positive evaluation of information based on homeostatic feedback.

Although interaction and storing will both feature in this section, the emphasis will be placed on the application of evaluation as a means to adjust the AI's behaviour. MAI will have an internal reference frame by which it can judge its interactions, attach value to what it stores and from there make useful adjustments. This is in opposition to the standard AI that lacks an internal reference frame and relies on other, often external, processes to make adjustments. Rather than use standard AI Neurons, which both inhibit and excite, I will use more nature-inspired Neurons that either inhibit or excite, a functionality that can be duplicated with regular AI Neurons by locking either the inhibitory or the excitatory connections to zero.

MAI requirements

In order to build an AI that can support Arbiter and Value-assigner functionality, some basics need to be present in its design.

MAI will need some measures that will provide representation of the outside world. I will list within brackets what part of adaptability they support:

3. “Neurons” that the program uses as **sensors** to keep track of outside parameters. These outside parameters could be extremely limited, such as the ability to detect whether it is light or dark, or a full spectrum of detection including sight, touch, hearing and other sensor-input available or unavailable to humans (interaction), and
4. “Neurons” that represent **actions** targeted at the outside world (interaction). These actions could vary from the limited flicking of a light switch, to a much broader scope of actions such as complicated movement patterns.

Some measures representing the inside world are also required. These measures include:

2. Some range of “**homeostatic**” **parameters** wherein the program must try to stay (evaluation),
3. “Neurons” that the program uses as **sensors** to keep track of these parameters (evaluation),
4. Some automated, time-dependent process that **disturbs homeostasis** to drive action-

selection and simulate natural homeostatic disruption, this can instead be achieved by using an actual draining battery that can be recharged (evaluation), and

5. A rising **degradation-counter**, which will deactivate MAI if it stays outside the homeostatic zone for too long, connecting it to an ultimate consequence: death (grounding evaluation). This counter will slowly go down if MAI spends enough time at the required homeostasis to simulate repair.

Other required program capabilities:

- Some form of **memory-storage**, capable of storing internal and external perceptions, as well as actions (storage),
- An a**ction-selection** mechanism that selects between available actions based on action-memory (interaction), combined with relevant action-prompts provided by:
- An **Arbiter**-mechanism. This mechanism must be capable of ascertaining homeostatic imbalance in order to trigger the action-decision mechanism. It must also be capable of judging action-impact on homeostasis from the provided internal perception. It must be able to assess not just in which direction they adjust homeostatic values, but also whether this is good or bad in the given situation (evaluation). The Arbiter then informs the action-selection mechanism and the:
- A **Value-assigner**-mechanism, which commits this verdict to memory through signal modulation, making the memory meaningful by attaching values that are directly connected to consequences (evaluation, storage and adjustment), and finally
- A **priming**-mechanism, internal to the Neuron itself, which allows it to be modulated by signified reward. Inactivity of this priming-mechanism prevents Neuron-connections that were not part of the reward-producing action from being reinforced (evaluation, storage and adjustment).

One last prerequisite, that is easy to forget but very much a necessity:

- The possibility of the **external environment** affecting homeostasis (interaction and consequences), such as light or darkness affecting energy levels.

The hardware

Ideally, a real-world, physical Neural Net would be created to support the learning mechanisms above. However, physical Neural Nets are prohibitively expensive to create and maintain. Additionally, they have physical limitations that are hard to get around: Neural Nets can become bulky very quickly, the “growth” of new “Neurons” and connections between them is very hard to implement and there is a high risk for unplanned physical damage to name but a few problems inherent in physical Neural Nets. With the MAI approach to learning, there is also a very real chance of damage inflicted upon the system by MAI itself. This is due to the fact that MAI is supposed to detect real, physical consequences and act on them: program failure to avoid damage could result in the destruction of a physical net, making a physical net unsuitable for MAI-experimentation. Even successful evaluation may result in limited damage as MAI finds out the consequences. Therefore, at least in the foreseeable future, it seems more practical to settle for a computed Neural Net. I will argue in the chapter on Philosophy that this has no negative philosophical implications for the status of the program's reality, adaptability or even its grounding.

Establishing homeostasis

Seeing that this is only a first attempt, it seems wise to keep things relatively simple and give MAI only one homeostatic parameter to keep track off: the amount of electric energy. This particular example parameter is chosen for two reasons:

- It has **native survival consequences** for a computer, due to their electricity dependence, which enhances realism,
- It forms a very easy **analogue for biological organisms** who require their own type of energy to stay functional, including later options for including storage of reserves,[172]

Like in biology, too little energy will lead to a less than optimal performance on a computer. In the more extreme cases the value could fall low enough to cause a cessation of function as simply not

---

[172] As an aside, some microscopic organisms are actually thought to survive on nothing but electricity itself. Brahic, C. (16 July 2014). Meet the electric life forms that live on pure energy. http://www.newscientist.com/article/dn25894-meet-the-electric-life-forms-that-live-on-pure-energy.html?page=1#.U8fh_LGvTX4. *New Scientist 2978* (retrieved 21 July 2014).

enough energy is available to continue functionality. Also as in biology, where too much unbound glucose can destroy cells or an overstuffed intestinal tract may rupture, a too high dose of electric energy (in the form of electric current) can cause damage to the physical computer hardware decreasing performance and ultimately causing irreparable breakdown. Therefore, by picking energy as the functional homeostatic value, a degree of realism is preserved if dire consequences such as the cessation of a faulty program are enforced. Because MAI will initially be run on virtual hardware (a simulated Neural Network), no physical energy restraints will penalise poor decisions and failure to rectify them. Instead, simulated restraints, where the insufficient or excessive power reduces MAI's efficiency and starts the process of “degradation” which eventually shuts MAI down, will serve as the limiting factor. I will discuss the philosophical consequences of this in Chapter 5.

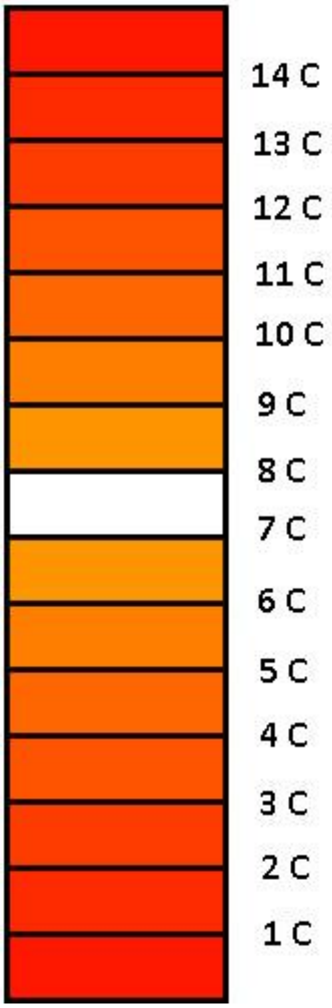

Figure 4.1: Illustration of the used example homeostatic range. The white range indicates no signal: between 7C and 8C no corrective action is required so no signal is given. Increasing redness away from the 7C to 8C range indicates increasing survival threat.

Here an obvious objection appears already against the claim that this AI would be new or special. Modern computers already have programs running that monitor power surges and overheating (which is an important side-effect through which high energy destroys computer hardware). These programs will often also take actions to prevent the computer hardware from its untimely destruction: CPU rates will be throttled to reduce energy consumed and through that reduce the heat produced. These programmes may even force the computer to go into a hard shut down to prevent damage from occurring. Of course, this resembles the behaviour of MAI in some ways, as MAI will also try to self-regulate. Surely MAI is not a unique program in every way and of course computers run programs that monitor their physical health as predetermined by outside programming. However the manner in which this function is integrated is fundamentally different: unlike normal computers, where the heat-monitor is a separate program that runs in the background and only controls the CPU, the fan and the ON/OFF switch, the integration of this monitoring function in MAI will be complete: every act MAI performs, every link it makes, everything will be associated with its effect on the energy running through the computer. The importance of every action is the impact it has on the available energy and MAI will adapt not only reduce the damage caused by power surges and overheating, but also to prevent them.

Let us return to creating MAI.

The internal detection array:

In order for MAI to detect danger to its homeostasis, MAI needs to monitor what state the energy value (C) is currently in (henceforth actual state), as well as knowing what state it should be in (homeostasis). Let's arbitrarily place the ideal energy for the machine's hardware between values 7C and 8C. Anything below or above is detrimental to prolonged survival and needs to be adjusted to prevent damage or even cessation of function. The larger the difference between the ideal values and the actual value, the more serious the threat (see Figure 4.1). Obviously, MAI now needs sensors to tell its actual state. These sensors for the brain in the human body are neurons, it seems only apt that we replace them with "Neurons" similar to those found in Neural Nets (with the exception that these will be exclusively excitatory). At first glance, MAI seems to need only three, or perhaps even only two Homeostatic Neurons:

- Neuron A fires when energy < 7C,
- Neuron B fires when energy > 8C, and
- Neuron C, which fires when 7C ≤ energy ≤ 8C. Neuron C is not actually necessary and biologically unlikely. Research and every day experience suggests that biological organisms are not informed if homeostasis is maintained, but only receive signals of change and threat. Provided Neuron A and B function properly, the program doesn't need to be informed that it's currently in the safe zone, it just needs to know if it's in the red zones and if its situation is improving or growing worse. (See Chapter 2).

An alternative way to model this is with another set of Neurons:

- Neuron X fires when energy > 7C,
- Neuron Y fires when energy < 8C, and
- Neuron Z, which fires when 7C ≤ energy ≤ 8C isn't true. Again this third Neuron is not quite

> as necessary, as it only informs the body that it is currently not in the safe zone. This required information is already available (if both X and Y fire, homeostasis is between 7C and 8C). That said, there is some biological suggestion for a function like Neuron Z. Touching extreme heat or cold, for instance, often feels the same (i.e. painful temperature) for a short time before more specific information arrives (i.e. hot or cold). This allows for a quick reflexive cessation of contact where a quickness to act is more important than a detailed reason. It is also possible that this function resides within the Arbiter (see below) instead.

It is possible to double-layer the two groups of neurons, which provides redundancy.
If Neuron A and Y fire (see Figure 4.2), the Arbiter knows that energy is below 7C. If Neurons B and X fire, the Arbiter knows that energy is above 8C. If Neurons X and Y fire and A and B do not, the Arbiter knows that energy is between 7C and 8C, or rather, that homeostasis is currently being maintained. From biological sources we know that there is no signal for achieved homeostasis, which suggests that the biological equivalent of Neuron X, Y and C-types don't exist, as those fire during achieved homeostasis. Their biological absence is possibly due to the fact that firing neurons cost more energy than non-firing neurons while they would, in this case, provide no useful information. The redundancy matter in nature is probably fixed instead by having more neurons (for implementation, see below). Regardless of biological realism, because the Neurons A and B suffice, I will disregard the C, X, Y and Z possibilities for now, to simplify what is to come.

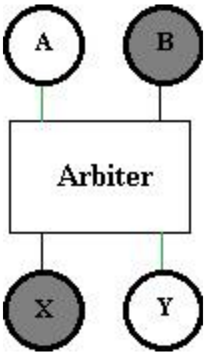

Figure 4.2: Example of Neuron activation when energy has dropped below 7C. Neuron's A (below 7C) and Y (below 8C) are active, as is the Arbiter, while Neurons B and X lie dormant.

This orchestration will work for the first basic internal monitoring task: knowing when the balance is off and in which direction it is off. It allows the program to take countermeasures that it knows are effective in changing the balance, or to experiment with new countermeasures if it doesn’t know of a method yet. However, in this second case it can quickly run into problems: most methods will not result in an immediate drop or rise in energy, so how does MAI know it is on the right track or not? With only these sensors, the computer is unable to decide what is a beneficial action and what is a detrimental action, unless that action completely changes the state MAI is in (see Table 4.1).

MAI doesn't necessarily have time to test each method long enough to see what its effects are, if it has any effect at all. Worse, in many cases it won't be able to detect a change when such a change definitely occurs. In order to dramatically shorten the timespan for arriving at a conclusion as well as enabling a wider range of possible conclusions, we need to expand our Neuronal assembly to allow for a much more accurate reading. Let us assume that a total shutdown is imminent at 3C and takes considerably longer in the range between 3C and 7C (although progressively quicker as it reaches 3C). A linear line of Neurons derived from Neuron A to monitor <3C, <4C, <5C and <6C could provide the program with much quicker and versatile feedback.

Table 4.1: Display of Arbiter judgements of possible states at T2 if T1 was a "too low". Note the lack of valued difference between states 3C to 6C and 9C and the resulting lack of valued association.

Condition at T1	HS Neurons firing T1	Arbiter	Possible T2	HS Neurons firing T2	Arbiter	Improvement/ Worsening	Association
Energy 5.1C	A	"Too low"	Energy 3C	A	"Too low"	A → A, Unknown/No	Neutral
			Energy 4C	A	"Too low"	A → A, Unknown/No	Neutral
			Energy 5C	A	"Too low"	A → A, Unknown/No	Neutral
			Energy 6C	A	"Too low"	A → A, Unknown/No	Neutral
			Energy 7C			Improvement	Positive
			Energy 8C			Improvement	Positive
			Energy 9C	B	"Too high"	A → B, Changed to a different negative.	Positive/ Negative

Let us split Neuron A into Neuron group A with the following properties:

- Neuron A1 fires when energy < 7C,
- Neuron A2 fires when energy < 6C,
- Neuron A3 fires when energy < 5C,
- Neuron A4 fires when energy < 4C, and
- Neuron A5 fires when energy < 3C,

This group allows the Arbiter a significant second layer of information. It can now not only establish if the present Energy value C is below the desired threshold, but the Neurons also allow it to judge how far below the threshold the threat is by both signal source (A5 is more serious than A1) as well as signal intensity (the higher the number of firing Neurons, the more severe the threat).

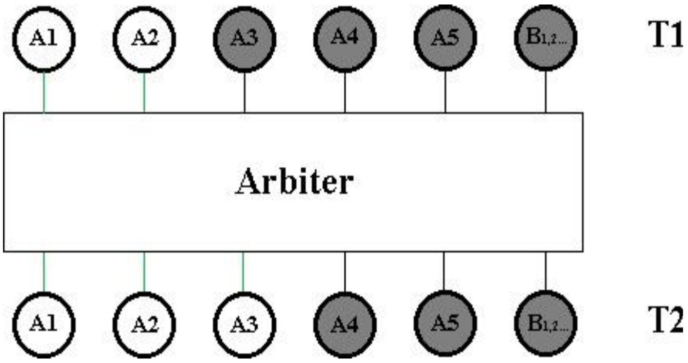

Figure 4.3: Different Arbiter inputs at T1 and T2. Note that both from Neuron-source and from number of active Neurons (white circles depict firing Neurons, grey circles depict inactive Neurons), the Arbiter can tell both when change occurs and what the scale of the change is.

Most importantly though, the system now has more possible states to compare. Rather than being limited to “Homeostasis”, “Too low” and “Too high”, with this extra layer of information, the system now also has access to more nuanced comparative statements such as “still too low, but an improvement”, “still too low, no change” and “still too low, now even worse” (see Table 4.2). With these more refined statements, the system is much more capable of producing a positive or negative statement about the effects of the most recent events. Only in a very limited case where the C doesn’t change, or doesn't change enough, to change the firing pattern will the Arbiter be unable to evaluate if there was any change at all. Naturally, the smaller the activation gaps between Neurons, the more justified MAI will be to conclude that a particular action has had no positive or negative effect when no change in the Neuronal fire pattern is detected. The more accurate the reading, the faster the program can make informed decisions built on internal valence, at the cost of increased complexity and maintenance.

Table 4.2: The same situation T1 and possible T2's, but with a more refined sensory apparatus. Note how the Arbiter is able to make a better judgement as well as provide more valued associations.

Condition at T1	HS Neurons firing T1	Arbiter	Possible T2	HS Neurons firing T2	Arbiter	Improvement/ Worsening	Association
Energy 5.1C	A1,2	"Too low"	Energy <4C	A1,2,3,4	"Too low"	A1,2 → A1,2,3,4 Much worse	Very Negative
			Energy >4C, <5C	A1,2,3	"Too low"	A1,2 → A1,2,3 Worse	Negative
			Energy >5C, <6C	A1,2	"Too low"	A1,2 → A1,2 Unknown/No	Neutral
			Energy >6C, <7C	A1	"Too low"	A1,2 → A1 Improvement	Positive
			Energy >7C, <8C			A1,2 → Great Improvement	Strong Positive
			Energy >8C, <9C	B1	"Too high"	A1,2 → B1 Great Improvement/ New Problem	Strong positive/ Weak Negative
			Energy >9C, <10C (etc.)	B1,2	"Too high"	A1,2 → B1,2 Great Improvement/ New Big Problem	Strong Positive/ Strong Negative

Modelling

Based on these sensory requirements and the list of other requirements listed in the earlier chapter, we can set up a schematic representation of the relations between the different factors that are included in MAI or directly affecting it:

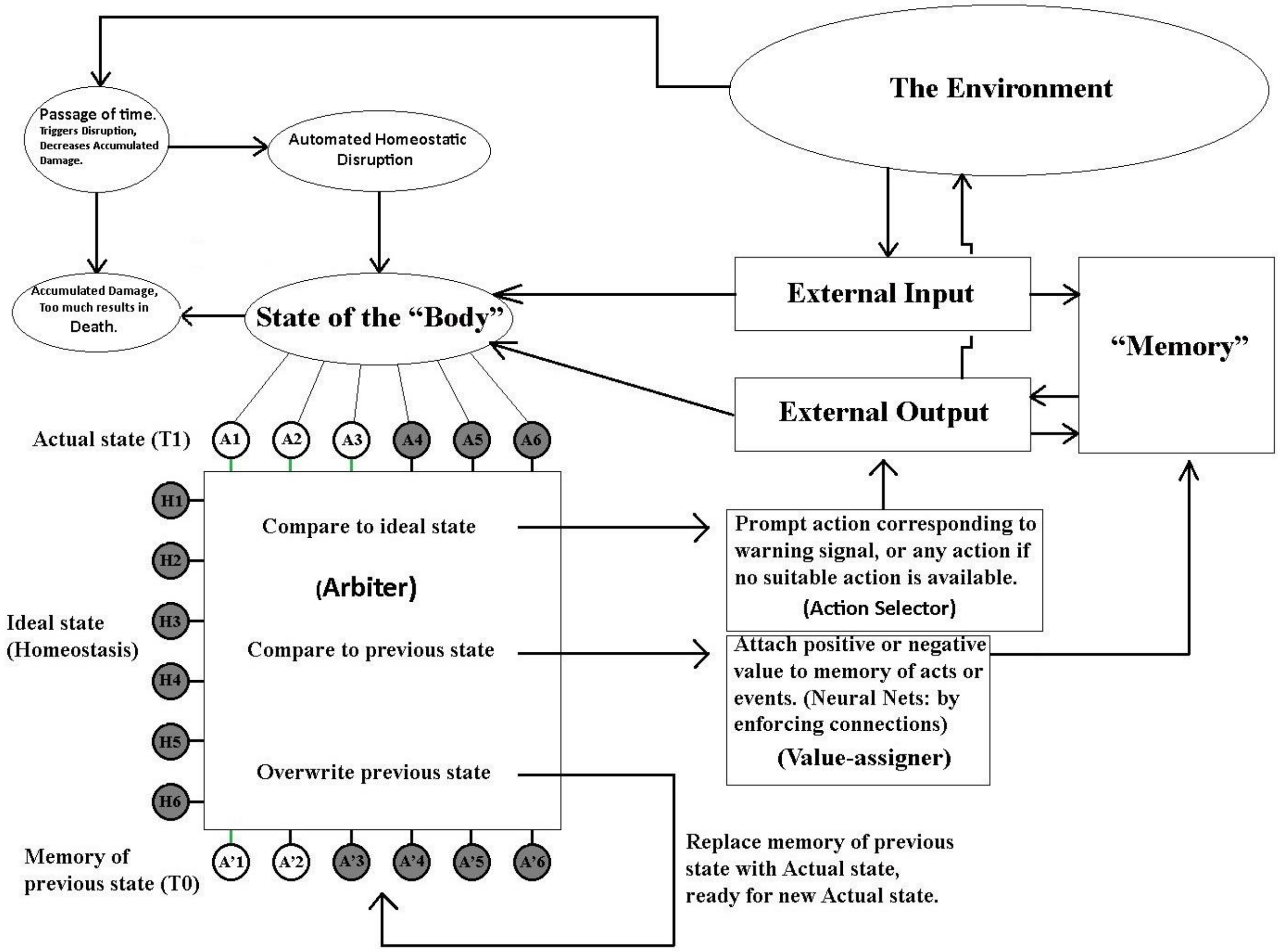

Figure 4.4: Schematic depiction of interaction between the different modules that comprise MAI. White circles depict firing neurons, grey circles depict inactive neurons. Arrows depict direction of influence.

Now that we know what kind of information-input the Arbiter requires, we can devise a possible solution for the comparative functions it executes. I will now demonstrate a set of possible Neuron-based mechanisms that would respond to homeostatic change in an adequate fashion. These are in no way exhaustive, but should give a suitable starting point for the experimental exploration of MAI.

I will start with the comparative functions of the Arbiter. In the following illustrations, I sketch a possible Neuron-based network that can compare old homeostatic information to the current information and send the correct signal to the Value-assigner to be transmitted. In order not to clutter the illustration, I will only depict the possible paths and states that are associated with one single Homeostatic Neuron, Neuron A1. Every homeostatic sensory Neuron has a comparable connection to the Value-assigner, so the amount of active connections that enter the Value-assigner can dictate the value-strength transmitted by its reward signal. For comparison across two moments, it is clear that Neuron A1 provides for a total of four possible different outcomes:

- It fired at the first moment (A'1), but not at the second (Figure 4.5, on the left),
- It fired at the second moment (A1), but not at the first (Figure 4.5, on the right),
- It fired at both the first and the second moment (Figure 4.6, on the left), or
- It fired at neither the first, nor the second moment (Figure 4.6, on the right).

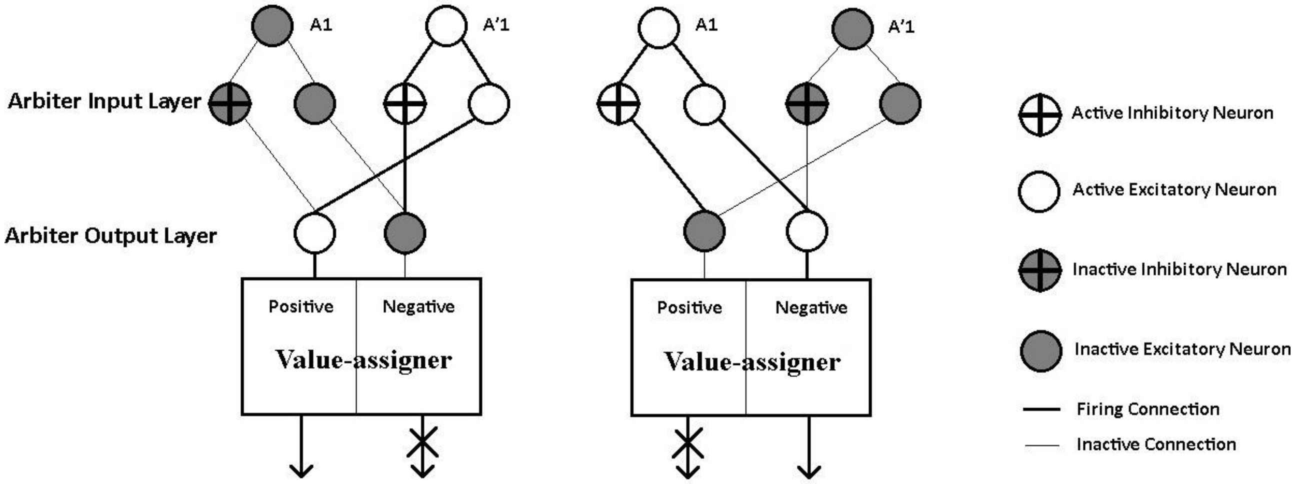

Figure 4.5: Arbiter determining changes in homeostatic value across time interval, part 1. The A1 Neuron represents the status of homeostatic values at the current moment, the A'1 Neuron represents a previous moment. In the left situation, a firing homeostasis Neuron no longer fires, meaning an improvement of homeostasis. The Arbiter-output layer sends a signal to the Value-assigner to release a positive reward signal. In the situation depicted on the right, the current homeostatic condition has worsened (A1 now fires) and the Arbiter prompts the Value-assigner to send a negative reward signal. The depicted network connections flow from top to bottom.

Neuron firing flows from the homeostatic Neuron A1 (or the representative of its past behaviour, A'1), into the Arbiter Input Layer. An active A1 or A'1 will activate the corresponding inhibiting and excitatory Neurons in this layer. An inactive A1 or A'1 will not activate the Arbiter Input Layer. The excitatory and inhibiting Neurons are cross-connected: the Inhibitory Neuron of A1 inhibits, if activated, the Excitatory Neuron of A'1 and vice versa. This allows for the extinction displayed in Figure 4.6, while activation in Figure 4.5 proceeds unimpeded (the active inhibitory Neurons inhibit inactive excitatory Neurons, resulting in no net effect on activation output). Because activation of homeostatic Neurons is a bad sign, a change from activation to no activation (Figure 4.5, on the left) is positive, while the converse (Figure 4.5, on the right) is negative. An unchanged firing pattern

(Figure 4.6, both networks) is neither a positive change nor a negative change.

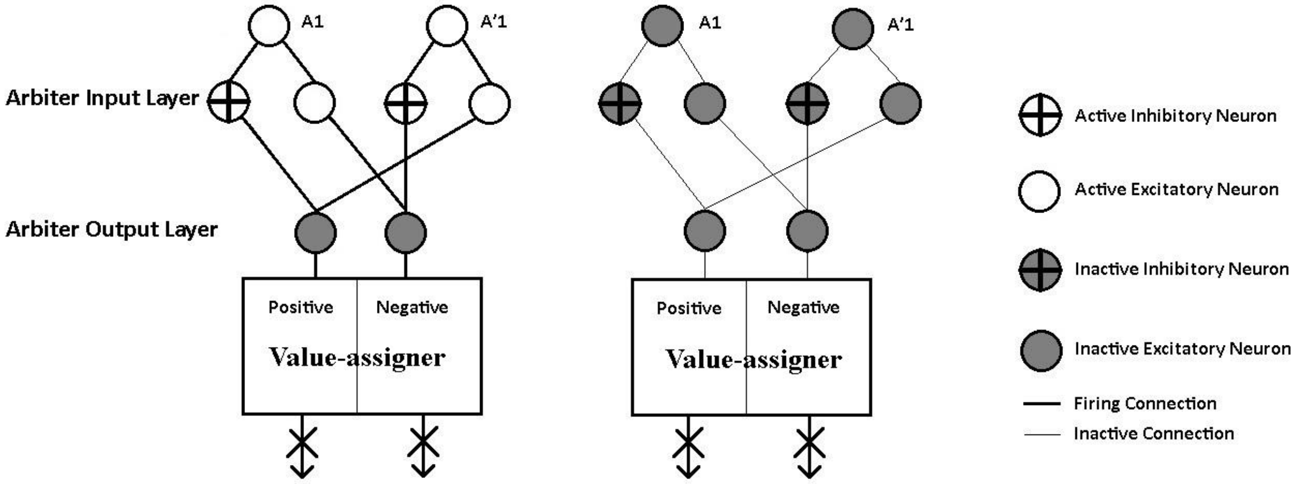

Figure 4.6: Arbiter determining changes in homeostatic value across time interval, part 2. The A1 Neuron represents the status of homeostatic values at the current moment, the A'1 Neuron represents a previous moment. In the left situation, the previously firing homeostasis A1 Neuron is still firing, meaning no improvement or worsening of homeostasis. The Arbiter sends no signal to the Value-assigner (the inhibitory Neuron prevents it). In the situation depicted on the right, the current homeostatic condition has also stayed identical and again the Arbiter sends no signal to the Value-assigner. In reality, the left situation is exceedingly rare, due to the great amount of homeostatic Neurons and the given that homeostasis is never completely stable. The depicted network connections flow from top to bottom.

From Figures 4.5 and 4.6, it becomes apparent that a relatively simple cross-connected mechanism will allow the proper output Neuron to be activated. In the first situation in Figure 4.5, the current situation as displayed by A1 is an improvement over the old situation in A'1, and the Arbiter informs the Value-assigner accordingly, prompting a positive reward. In the second situation in Figure 4.5, the reverse is true and this time the Arbiter informs the Value-assigner to release a negative reward. In the second situation in Figure 4.6, no homeostatic neurons have given off a warning, either in the present, or at the previous measuring moment. The Arbiter takes no action and no reward is given by the Value-assigner. In the first situation depicted in Figure 4.6, something strange occurs that is extremely rare in biology: homeostasis is disrupted and has stayed *exactly* the same. The Arbiter has therefore given no rewarding signal. In all likelihood, this problem is due to insufficient homeostatic sensors (compare Table 4.1 and Table 4.2 and consider expanding the Table 4.2 range of A-Neurons to allow for even smaller increments). Thanks to the Automatic Homeostatic Disruption mechanism (see Figure 4.4 and the section on MAI requirements), as well as the tendency for actions to produce homeostatic disruption due to the expenditure of resources, a negative reward signal is likely to be produced sooner or later, allowing for comparative evaluation over time.[173]

The mechanics for homeostatic comparison between current homeostatic values and the ideal homeostatic values is similarly processed (see Figure 4.7). In Figure 4.7, H1 replaces A'1 as the comparative target, while the Action Selector replaces the Value-assigner as the target module. Because in ideal homeostasis Homeostatic Neurons are inactive, H1 (which stands in for the range of H1, H2, H3, etc.) is only displayed in its inactive state. Only the situations that feature inactive ideal homeostatic H1 Neurons are currently displayed, but if active H1 Neurons were somehow involved, the networks displayed on the left in both Figure 4.5 and Figure 4.6 would function in the same manner with the exception that the target module is not the Value-assigner but the Action Selector.

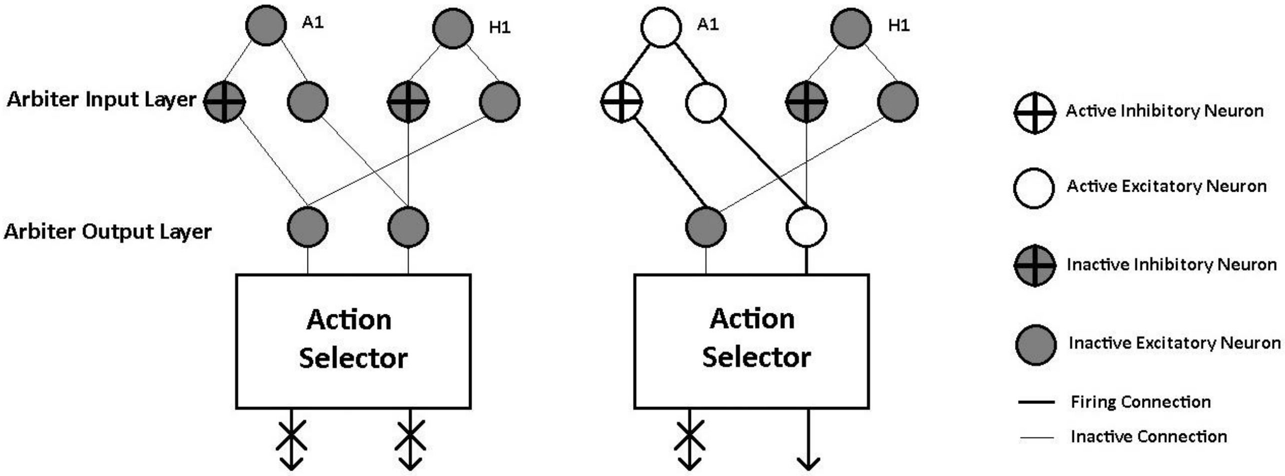

Figure 4.7: Arbiter-mechanism for prompting action-selection on the basis of homeostatic disruption. The A1 Neuron represents the status of homeostatic values at the current moment, the H1 Neuron represents ideal homeostasis. It is in many ways identical to the mechanism that feeds into the Value-assigner. The main difference is that the H1 Neuron and its derivatives do not fire and seem largely superfluous. However, the mechanism as presented is capable of handling a change in homeostatic preferences: if the required homeostatic value changes so that an H-Neuron fires when homeostasis is preserved, the Arbiter would still be able to assess homeostatic conditions correctly to prevent unnecessary Action Selector activation. It seems that adaptability is therefore served by conserving the "superfluous" structure. The depicted network connections flow from top

---

[173] To improve action selection when an action does not have the desired effect, another Arbiter module could be added that compares actual effect to predicted effect. Such a structure is too complex to be covered in this thesis and may be a desirable addition in further experimentation.

This may lead to the impression that H1 and all its derivatives are superfluous: they seem to have no effect on the displayed outcome, because they are always inactive (see Figure 4.7). However, if a change were to occur that alters homeostatic preferences without a change in the sensitivity of the homeostasis-sensing A-Neurons, H1 and its group members can adjust by becoming active instead. Imagine for instance a change in the physical hardware that results in a new homeostatic ideal of 6C to 7C. As established earlier this chapter, the unadjusted Neuron A1 will fire if the value of "C" falls below 7C. This would result in constant signals to the Action Selector prompting action as well as output to the negative Value-assigner, which would give wrongly valued reinforcement to any further actions. By turning the default setting for the "ideal homeostasis" Neuron H1 from "non-firing" to "firing", this can be prevented (an active H-Neuron alongside an active A-Neuron will inhibit action-prompting and signifying, while an active H-Neuron alongside an inactive A-Neuron will cause Action Selector and Value-assigner activity). The Arbiter-mechanic will then still work in maintaining Neuron A activity at the same level as Neuron H activity, preventing unnecessary and detrimental Action Selector activation while promoting it when it is required. Due to the adaptive options keeping the more complicated structure provides (allowing for homeostatic reorientation without changing the sensor array), as well as the occurrence of (limited) homeostatic reorientation in biology, I have decided to maintain the inert H-Neurons and their derivatives even though no mechanism to reorient H-Neurons will be presented in this thesis.

It is important to note that, although the Figures 4.5 – 4.7 only display one homeostatic-sensing Neuron (A1), a similar little network will exist for every A-Neuron (be there six, or six hundred), as well as B Neurons and any other sensory kind. The connections from the Arbiter to the Value-assigner and to the Action Selector are similarly expandable and intended to be cumulative. So six A-Neurons signalling the Value-assigner or Action Selector through the Arbiter mechanism would respectively prompt a higher reward, or more encouragement for action. The more little networks are added, the more precise the mechanism will function and the more adaptable it will be. We must now continue with the interaction between the reward-signal provided by the Value-assigner and the memory-storage that the Action Selector depends upon. The inner workings of the Action Selector itself will not be elaborated upon, because it falls outside of the scope of this thesis which is concerned with the role of motivation in adaptability and not with the mechanism of action selection itself. After modelling the interaction of reward with action memory (see Figure 4.8), I will provide three methods through which the Value-assigner can bring its rewarding message to the

Neuronal connections that form the memory in Neural Nets.

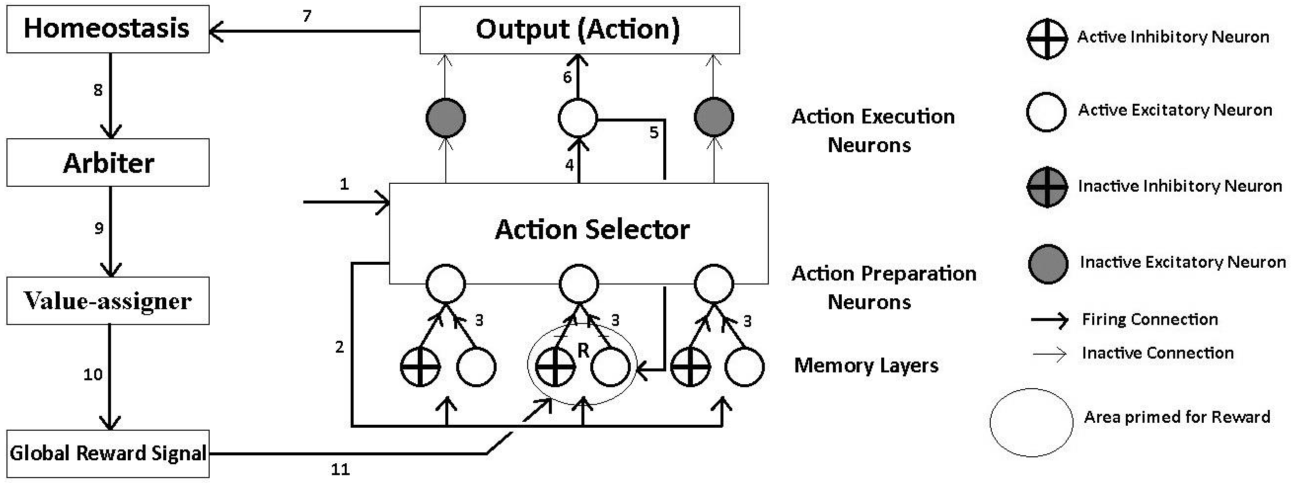

Figure 4.8: Value-assigner influence on action-selection. An action prompt reaches the Action Selector from the Arbiter (1). The Action Selector triggers the action memory for relevant actions (2). The Memory Neurons excite and inhibit Action Preparation Neurons at the Action Selector (3). The strongest signal (excitatory minus inhibitory) is selected by the Action Selector, triggering the execution of the action by activating the Action Execution Neurons (4). These not only execute the action, but also provide a feedback loop back to the relevant memory Neurons (5) and prime them for Reward (area R). Action execution leads to Output (6), which affects Homeostasis (7). This effect is judged by the Arbiter (8) and passed on to the Value-assigner (9), which gives off the corresponding Global Reward Signal (10). This signal is then picked up by the Neurons that are primed for reward (11). If the reward was positive, the excitatory Neurons are reinforced, if it was negative, the inhibitory Neurons are enforced. This impacts the signal strength associated with the action by either increasing it (through enhanced excitation), or lowering it (through enhanced inhibition), affecting the chance that the action will be selected again in these circumstances.

Reward and connection strength

The interaction of reward and action-memory starts when the Arbiter detects homeostatic disruption and prompts the Action Selector to undertake action. The Action Selector then activates the Neurons that compose its memory. This memory is created by connecting two kinds of Neurons: inhibitory Neurons and excitatory Neurons. These are connected to Action Preparation Neurons that correspond with actions that can be taken. These Action Preparation Neurons spike towards the Action Selector with a connection strength that is increased by excitatory Neurons and decreased by inhibitory Neurons in their corresponding memory. The Action Selector then selects the strongest signal and allows it to continue to the corresponding Action Execution Neurons, which not only perform the action, but also loop back to their corresponding memory Neurons, priming them for modulation by the Value-assigner's signal. Through the Output, Homeostasis is (presumably) affected, after which the Arbiter will judge the effects and prompt the Value-assigner to release a Global Reward Signal. Any Neuron that has been primed to accept this signal, will now have its signal strengths modulated on incoming and outgoing connections. In the posited MAI-model, negative signals strengthen inhibitory mechanisms, while positive signals strengthen excitatory mechanisms. The first will result in a lower peak of the corresponding Action Preparation Neuron, reducing the chance of that particular action being selected in this context, while the second results in a higher peak of the corresponding Action Preparation Neuron, further increasing the chance that the action will be selected above other actions given this context. To add to the versatility of this mechanic, the Reward, and with it the amount of signal modulation, can be dependent on the size of the homeostatic change detected by the Arbiter. As displayed in Figure 4.4, the memory function can be made more adjustable to changing circumstances by allowing external input, i.e. from external senses, to plug into the networked memory, activating excitatory or inhibitory Neurons. Activation or inactivation of particular senses can then be coupled to a higher or lower final Action Preparation strength of each particular action. Through the Global Reward Signal, interaction between action memory, other memory and senses can achieve new adaptable heights. I suspect that by connecting these modular Neural Nets to each other and enforcing the beneficial connections through the Global Reward Signal, a wide range of versatility and adaptability can be achieved by this self-teaching AI. This is, however, the topic of practical experiment, which means I must now abandon the subject. Not, however, before I posit how the Global Reward Signal can be executed.

Transmitting the Global Reward Signal

In the MAI model, information is stored in the connections between individual Neurons, of which there are two types: inhibitory and excitatory. It is adamant that the Global Reward Signal is able to travel to the proper Neurons in order to perform its function. In order to prevent the wrong Neuronal connections from being enforced, Neurons need to be primed for Reward first. This priming should only last for a predetermined amount of time, allowing the Value-assigner to release an unguided Global Reward Signal as only the primed Neurons will pick it up and reinforce their connections. In combination with this priming mechanic, I see three ways for the practical implementation of the Global Reward Signal, each inspired by biology.

The first is to follow the simplest biological route: the dumping of reward chemicals into the neuronal environment. By this method, all Neurons will receive the signal, but only the primed ones will pick it up. There are some practical limitations in the nature of physical Neural Nets. A physical Neural Net lacks the complex biological apparatus that allows individual neuron cells to detect chemical signals in their environment. As a matter of fact, physical Artificial Neural Nets don't really have a shared environment other than the air kept in the server room. It is likely that a mechanical solution to this problem can be devised (perhaps something involving the mounting structure), but for now this is not a serious problem because we are working with a simulated Neural Net. In a simulated Neural Net, the updating of primed Neurons could be solved more easily by including a check that verifies if a Reward signal is currently active in the system and, if so, what strength and value it has. This would have the benefit that only primed Neurons require extra attention when a Reward signal is in play.

A second solution works for both simulated as well as physical Neural Nets. It involves making connections from the Value-assigner to every Neuron that would ever need to be signified. If a Reward signal is to be transmitted, the Value-assigner sends it across all of its connections (perhaps repeatedly if the reward is larger), allowing primed Neurons to receive this input directly. An obvious downside to this would be the massive amount of connections the Value-assigner would have to make as complexity increases. The transmission of a Global Reward Signal would mean checking every single Neuron, a task which can make any system unwieldy and which makes this method unlikely.

A third solution works with the connections that are already present. The Neurons that are part of MAI all contain an extra channel that is independent of their normal connection strengths that allows them to pass rewarding signals “down the line”. The prerequisite of being primed makes sure that the signal is only implemented by those Neurons meant to be affected by it. Again, the downside is that every Neuron would need to be checked when a Global Reward Signal is in play, but perhaps the workload could be reduced by combining it with already occurring activations.

All three methods have their upsides and downsides. Which method is most efficient and/or effective will probably require some practical testing, as will the optimal configuration of the interlocking modules that implement MAI. It is now time for some directives for future work and to posit one final, uncomfortable question:

Is MAI a true Artificial Intelligence as described in Chapter 1?

No, not really. MAI as it has been presented is adaptable: it can interact with the environment, evaluate these interactions based on internal consequences and with it form meaningful memories. It can then adjust its future interactions accordingly. However, it is still far away from planning, reasoning and other skills entailed in the definition of massive adaptability, or "intelligence". Without an overseeing architecture, it is also quite likely that MAI doesn't "experience" any feelings associated with its mechanisms, in spite of its inherent motivation (hence the name "Motivated AI, rather than "Feeling AI"). One likely cause of these problems is that MAI is still extremely simplistic. It has a very limited array of homeostatic values to monitor, includes only the most limited Arbiter mechanic and the Neurons with which it has been built are also still incredibly simple compared to their naturally occurring counterparts. MAI does, however, have a level of bare-bone intelligence and this intelligence is grounded in reality as well as internal consequences through its learning mechanisms. MAI could therefore function as a grounding module implemented in other AI, in order to give internal meaning to external signifiers and to facilitate reasoning already present. Not only would this enable that AI to learn and behave in a more intelligent fashion, but it would also be more easily recognised as intelligent. MAI is currently little more a small behavioural decision-making unit that could for instance be fit into the Physically Grounded Systems approach of Nouvelle AI such as proposed by Brooks in 1990,[174] to replace preprogrammed automated responses with self-directed learning that can establish *which* preprogrammed automated actions should be implemented in which situation.

Future work

An actual experiment involving MAI could show any shortcomings in its current design and could determine if it will actually work as envisioned. If MAI successfully passes that test, there are a number of regions in which its functions could be expanded. I will name some of the possible elaborations that could further improve MAI's adaptability.

Expansion of the number of different homeostatic variables (currently only one) and the introduction of more corresponding arbiter-modules can make MAI significantly more adaptable. A first field of expansion would be to add external sensors of varying types. For the sake of simplicity,

---

[174] Brooks, R.A. (1990). Elephants don't play chess. *Robotics and autonomous systems 6* (1) 3-15.

MAI as designed in the current chapter does not feature any explicit external sensors in spite of them being listed in the section on MAI Requirements. Instead, they have been abstracted away into the “External Input”→”Memory flow” of information in Figure 4.4, while in Figures 4.8 they are simply a potential part of the “memory layers” that are activated. Adding ranges of sensors which input into the Memory layers can allow MAI to take external variables into account when its Action Selector is operating. Feedback provided through the Value-assigner can modulate the impact of sensory information on the signal spike of Action Preparation Neurons through inhibition or excitation of memory layer Neurons, which would make the selection process significantly more adaptable. Another important field of expansion lies in the homeostatic variables. Requiring MAI to keep track of multiple variables will allow it to attach internal value to more encounters it has with the outside world, as well as giving those encounters more dimensions: a positive encounter on one homeostatic range may also represent a negative encounter on a different range. This could allow for complex behaviour to emerge as multiple arbiters vie for Action Selector control. A third, more complicated expansion would be to integrate an Arbiter-like function that compares predicted outcomes of an action to actual outcomes as reflected in homeostatic disruption. This evaluative function could improve the speed of the learning process by quickly eliminating actions that do not produce the wanted result, but will probably require considerably more work to implement. Finally, the largest improvement to MAI would be its implementation in a robot body, which allows for embodiment that doesn’t have to be predefined in a software environment. In the real world, such a MAI-robot could try to find its own way to energy, learning on its own through homeostatic monitoring which interactions are beneficial in achieving that goal.

It is my belief that increasing the amount, range and functions of networks founded on MAI will eventually produce something that can be called massively adaptable, or in other words, intelligent. In the meanwhile, MAI is an improvement on the traditional Neural Net which needs to be trained by an external computer that tweaks its connections on the basis of a large set of training samples and countless trials. MAI tweaks itself and every disruption it encounters is its own, self-evaluated training sample. By internalising the training, based on internal consequences, MAI even challenges some widespread preconceived notions in the Philosophy of AI. It is to these notions that we must now turn our attention.

Literature

Brooks, R.A. (1990). Elephants don't play chess. *Robotics and autonomous systems 6* (1) 3-15.

Brahic, C. (16 July 2014). Meet the electric life forms that live on pure energy. http://www.newscientist.com/article/dn25894-meet-the-electric-life-forms-that-live-on-pure-energy.html?page=1#.U8fh_LGvTX4. *New Scientist 2978* (retrieved 21 July 2014).

Chapter 5: Motivation and the Philosophy of Artificial Intelligence

Now that we have a first model of Motivated Artificial Intelligence (MAI), it is time to evaluate the impact of integrating motivational systems in AI on the matters of AI Philosophy. The philosophy of Artificial Intelligence knows a variety of problems, many of which have an overlap with Philosophy of Mind and other areas of Philosophy. I have selected three problems (although we will later find out that they are in fact only two problems) for further inspection, because they are connected so closely to the purpose behind MAI: providing AI with the tools to assign value to its interaction with the outside world. These three objections to the tenability of the quest for Artificial Intelligence, as described in chapter 1, are:

1. The Chinese Room thought-experiment,
2. The Simulation Objection, and
3. The Symbol Grounding Problem.

For the first two problems, I will mostly argue my case by presenting it against the claims made by John R. Searle, the inventor of the Chinese Room thought-experiment and a fervent proponent of both the Chinese Room and the Simulation Objection as viable and definitive “proofs” against Artificial Intelligence. According to Searle, a computer Artificial Intelligence that is intelligent in the same sense that we are, including “consciousness”, “feelings” and other attributes we associate with exercising our intelligence, is nothing but a fool's dream. His Chinese Room thought-experiment is supposed to demonstrate that computers cannot possibly be able to understand the *meaning* of the symbols they are operating. Since we humans are able to understand what the symbols we operate mean, Searle concludes that a computer could never be intelligent in the same sense that we are. He often compounds this by bringing in the second issue, called the Simulation Objection. This objection entails the claim that a simulation of the human mind on a computer cannot be said to be conscious or intelligent in the same sense that we are, because a simulation of a thing is not the same thing as the original. In my defence of the quest for Artificial Intelligence, I will demonstrate some serious problems with the examples Searle chooses to convince people of the intuitive validity of his Chinese Room thought-experiment as well as the Simulation Objection he so often touts. I hope to demonstrate that the Chinese Room and Simulation Objection are not only based on faulty comparisons, but also generally invalid.

The third issue, the Symbol Grounding Problem, lies at the core of philosophical problems in the field of Artificial Intelligence. Unlike the Chinese Room, which is really just an example of the Symbol Grounding Problem taken to its extreme, the Symbol Grounding Problem merely problematizes the relation between the "Intelligence" of AI and the symbols it manipulates without immediately concluding that no "grounding" is possible. The Symbol Grounding Problem is a philosophical problem that not only affects Artificial Intelligence, but the cognitive "computationalism" theory of the Mind as well.[175] I will argue that the introduction of internal consequences, on which the evaluation of interactions can take place, allows for the "grounding" of programs on computers as long as the program is of a suitable type. Artificial Intelligence, if created from adaptable behaviour based on the four component parts of adaptability (interaction, evaluation, storage and adjustment) will be able to assign internal consequences to external signals. Connections and consequences established through valued adaptability will ground the symbols the AI is constructed with in reality and give it the ability to assign meaning to its interactions. First things first, let us deal with the Chinese Room.

The Chinese Room: semantics and meaningful representation

One of the most famous thought-experiments in the Philosophy of Artificial Intelligence is called "the Chinese Room", devised by John Searle. In this thought-experiment he proposes to lock a man with only an understanding of the English language and no understanding of Chinese, neither in writing nor in speech, into a room. In this room, the man finds instructions to manipulate symbols presented to him through an input slot. After he manipulates the symbols via the instruction manuals filled with rules, he outputs a string of Chinese symbols to be read by Chinese people on the outside. Thanks to the brilliant design of the symbol manipulating manuals, these Chinese readers mistake the output for an intelligent reply to whatever they asked the "machine". Searle then calls out that the man inside the Chinese room doesn't actually understand any Chinese. He is just manipulating symbols by following the manual, which completely robs him of any semantic understanding of what is going on. Of course, the man in this thought-experiment is a replacement for tasks normally left to a computer and the manuals of instructions are simply programmed rules

---

[175] The computationalism theory of Mind assumes that human intelligence and thought is built on symbol manipulation by the brain. I do not agree with this theory, but will not discuss it as it falls outside the scope of this thesis.

that get the computer from the input to the output without any actual understanding being involved. Searle concludes that, since the man is taking over the computer's functions and doesn't understand the symbols manipulated, the computer doesn't have an understanding of the meaning of the symbols it manipulates either.[176,177] Searle even makes a stronger statement, that a computer isn't even really capable of performing syntax or computation: both syntax and computation are solely projected unto the computer by its creators and users: humans.[178] Searle's attack is two-fold. Firstly and perhaps unintentionally, it is an attack on the Turing Test. The computer running the Chinese Room-program should be able to pass the Test and therefore will be consistently mistaken for something intelligent. Searle, however, never actually calls out the test as defective (which it is) although he comes close in his 1990 paper “Is the brain's mind a computer program?”,[179] presumably because he believes that the fault lies not in the Test but in computers themselves. The second attack is that a computer will not actually understand anything, because it is unable to attach any meaning to the symbols it operates, which runs contrary to the claims of those who adhere to the Symbol System Hypothesis.[180]

Anyone who has any knowledge of the inside workings of the computer running the Chinese Room-program will conclude that it is indeed nothing more than a machine shuffling symbols about. According to Searle this means that a symbol manipulating machine will never understand what it is asked, nor will it understand what responses it gives, because its programming is entirely confined to syntax (form) and has no concept of semantics (meaning).[181] After all, the poor English fellow[182] who has been so crudely and cruelly locked away, has no way of understanding what any of the Chinese symbols mean that he's manipulating. He only knows that if he receives one particular set of symbols, he must output another particular set of symbols. Searle seeks to prove that no symbol manipulator will ever be able to gain understanding and therefore no symbol manipulator can ever be considered “intelligent” putting the “strong AI hypothesis” to the grave. He even flat-out states that since Neural Nets can be run on an old fashioned symbol manipulating

---

[176] Searle, J.R. (2004). *Mind: A brief introduction* (New York, 2004) 62-63.

[177] Searle, J.R. (1990). Is the brain's mind a computer program? *Scientific American 262* (1) 26-31.

[178] Searle, J.R. (2004). *Mind: A brief introduction* (New York, 2004) 64.

[179] Searle, J.R. (1990). Is the brain's mind a computer program? *Scientific American 262* (1) 26-31.

[180] The Symbol System Hypothesis entails that true intelligence can be created using a digital computer that uses symbol-manipulation, conventional memory storage and other attributes of non-Neural Net AI.
Copeland, B.J. (1993). *Artificial Intelligence: A philosophical introduction* (Malden, 1993) 80.

[181] Searle, J.R. (1980). Minds, brains and programs. *Behavioral and Brain Sciences 3* (3) 417-24.

[182] To be fair, Searle proposes himself to be the test subject.

computer, no Neural Net qualifies for meaningful intelligence either.[183] Searle's thought experiment paints such a powerful picture that it has drawn, and still draws, a lot of attention in the field of Philosophy of Artificial Intelligence.

Luckily for the proponents of true Artificial Intelligence, there is a myriad of things wrong with Searle's thought-experiment that go beyond the ethically dubious act of locking someone in a closed off space with no means of communication outside of meaningless symbols. Quite a few potential flaws have been brought to light concerning Searle's reasoning. Possible objections have included the “systems reply”-argument that states that Searle only discusses a part of the system (only the computer itself), while leaving out the programming. Taken as a whole, hardware and software, the objection goes, the system may still understand. The reverse holds true as well: just because the system as a whole doesn't comprehend something, this does not preclude that parts of the system may still understand. This riposte is forwarded for instance by Jack Copeland[184] and I will forward a variation on this theme further in this Chapter (see: the code of life).

There are other objections to be brought forward against Searle, but I must first object to his choice of computer program to convince the reader that the Chinese Room disproves the possibility that computation can provide any kind of semantics. After all, thought-experiments rely on intuition, which means that the example given has a large impact on the way the argument is received. I will state this outright: I do not dispute that the Chinese Room Computer as provided by Searle cannot think. The program that the man in the room is operating is a “trick” program that evades the need for actual intelligence by only shuffling words around, backed by a large database. It obviously has no means of understanding. That the program is so obviously stupid explains in part the popularity of Searle's argument. The machine that he recreates by replacing it with a person may be able to pass the Turing Test, for the sake of argument, but it will in no way be intelligent. It does not understand its output, but I am willing to admit that it may not even be intelligent if it did “understand” its output. Something extremely vital is missing from the Chinese Room Computer that disqualifies it from being intelligent without any regard towards it being a symbol manipulator: it has no ability to learn whatsoever.

Although Searle has created the “perfect symbol manipulator” by blowing up the syntax database to proportions that would put the well-known computer program PARRY to shame, most

---

[183] Searle, J.R. (1990). Is the brain's mind a computer program? *Scientific American 262* (1) 26-31.

[184] Copeland, B.J. (1993). *Artificial Intelligence: A philosophical introduction* (Malden, 1993) 125-130.

every day computer programs are considerably “smarter” than the machine he devised because they at least have a limited learning capacity. Even if the Chinese Room-program may respond in perfect Chinese at the moment of its creation, it is going to produce outdated facts in archaic language as society progresses past its construction. Although Searle's conclusion that his Chinese Room is not intelligent and that it doesn't understand is in my opinion correct, even though there are others who have argued otherwise, the reason that it doesn't understand is not confined to it being a symbol manipulator, but to the fact that it has no means to learn or evaluate connections. In order to understand something, you must first be able to learn something. In other words, you need to be adaptable.

Chatterbots and the Turing Test

Searle's warning against attributing intelligence to conversational computer programs is not completely unjust. Conversational programs, called “chatterbots”, “chatbots”, or simply “bots”, are a very real presence on the internet, where they may entertain idle people, or more maliciously prey on credit card data and other vulnerable information from unsuspecting humans. One such (non-malicious) “chatter bot” has even been able to pass a variant of the Turing Test[185] during the *Turing Test 2014*, according to the University of Reading.[186] This “Eugene Goostman”-program was able to convince 10 out of 30 judges during a five minute conversation, that he was a 13 year old Ukrainian boy with English as his second language. The pretended identity was a very useful trick. It allowed the program to make mistakes that could be attributed to it being “young and foreign”, while it also enabled it to dodge questions it couldn't deal with, a trick that has been around since at least the “paranoid” PARRY-program. Programming it with intentional spelling and grammar errors even helps hide the fakery due to the fact that human typing is often imperfect, while it also masks unintended and otherwise peculiar errors. All things considered, computer programs that are designed to beat the Turing Test are generally designed to cheat the Turing Test. Because this

---

[185] The Turing Test is a language test wherein a computer attempts to convince a human judge that it is, in fact, human by sending typed conversation back and forth. It is often paired with an actual human who is also submitted to the test. The precise criteria used (time-interval, with or without human test-subject, percentage of judges fooled) vary, which has led to earlier claims of Turing Test success. One thing is clear, Turing seems to have underestimated the willingness of programmers to “cheat” the intention of the test when he mentioned it in his famous 1950 paper: Turing, A.M. (1950). Computing Machinery and Intelligence. *Mind 59* (236) 433-460.

[186] University of Reading (8 June 2014). Turing test success marks milestone in computing history. http://www.reading.ac.uk/news-and-events/releases/PR583836.aspx. *Press Releases University of Reading* (retrieved 24 June 2014).

cheating is so obvious when one looks at the way the program works, they make for a compelling example of why symbol manipulators could never understand their symbols. This is an unfair display of strong Artificial Intelligence though and a more adaptable type of program should be put through the Chinese Room treatment to see if the thought-experiment has any merit.

In the MAInese Room

Let's envision what the Chinese Room Thought-Experiment would look like if we instead ran the “program” MAI (see: Chapter 4). As noted before, the program MAI is run on a computer simulation of a networked AI. It has, on top of regular Neural Net design, a subsystem (Arbiter and Value-assigner) that monitors homeostasis and evaluates actions based on their impact on homeostasis. It uses this information to actively up regulate or down regulate the responses of particular Neurons, thereby tweaking the information storage and prompting the AI to adjust its interactions. This program regrettably will not start out on a physical Neural Network, but will instead be simulated on an old fashioned symbol manipulating computer. This, however, means that the comparison to Searle's Chinese Room can be preserved, although Searle believes that it doesn't matter.[187] I can now lock myself into this MAInese Room and pretend to be a computer running a program that is simulating a Neural Network with a homeostasis and reward-system built in.

Ok, so suppose I lock myself into this windowless, soundproofed and radiation shielded room that only has two narrow slits, carefully constructed so that nothing but a piece of paper will fit through. Inside the room I find piles and piles of manuals in a language I understand describing how to manipulate symbols I do not understand. I write them down on stacks of blank paper that function as a memory and repeatedly change these meaningless symbols in response to input-symbols I receive through the Inbox slit. Of course, this input is also provided to me in symbols that I do not understand. I assume them to be Chinese but later on I am informed that they are in fact Japanese writing or another language and script entirely, it does not matter. However thanks to the manual I do know what to do with them. Curiously, as symbols come in through the inbox slit, I find that the manuals don't limit my actions to just producing a piece of paper with meaningless symbols that I then stuff through the output slot: the instructions also require me to change symbols within the instruction set. I find that most of the input not only results in output, but also results in

---

[187] Searle, J.R. (1990). Is the brain's mind a computer program? *Scientific American 262* (1) 26-31.

changes in the very manuals I am working from. I find myself jumping between pages as the references within the manual are constantly changed by my manipulations. I also find that some manuals are used a lot. One of them consists of symbols that are adjusted with every input/output combination I make and I silently refer to this as the Main Manual. Even when the inbox slot goes quiet and the input trickles down to a slow stream, I find that I am constantly required to change the contents of this book. The simple passing of time requires me to change the symbols around and sometimes, when I have changed the symbols again, the manual instructs me to produce output, which interacts with this particular book, and then a lot of input starts to poor in again. The connections that I am manually operating become more and more complex in response to the received input, the established output and their interaction. Clearly something outside has an impact, not just on what happens inside, but also on what *is* inside.

Now. What to make of this experience? Certainly I did not learn the meaning of the symbols I was manipulating while slaving away in the MAInese Room. Frankly, I did not expect to learn the meaning. What I, as the operator, did notice was that the program I was running changed a lot. It changed in response to input, it changed while producing output, its output changed in response to its input and previous inputs, and it even changed when no input was given. It not only reacted to its environment, it actually interacted with the environment. It seemed to be in constant flux as it adapted to the input provided. Although I did not understand any of the symbols in the MAInese Room, I could tell that it was changing and adapting. The very symbols and the relations between them and the manuals they were written in, changed in response to the environment. The MAI program was altered at its core by these very interactions. This seems to me a very clear form of grounding: the symbols inside the room change and adapt in response to the input, the output and the passage of time. This limited form of grounding can be expanded upon by running one more instance of the MAInese Room-program:

Again, I am in the MAInese Room, performing the actions required for the program to run. However, this time something goes wrong. I keep changing the Main Manual as per the instructions in response to the outside input and the internal processes, while the output and input keeps streaming in. It even accelerates as the Main Manual repeatedly needs adjustment. And then it stops. Right there, in the manual, is a line that says "if this particular string of symbols comes about in this location, cease all symbol manipulation, do not provide any further output and do not accept any further input. Please wait for the door to open." A short while later, the door opens and I am let out of the room, only to be informed that after my marathon session, the program I was running failed to maintain its homeostasis with respect to its environment. It suffered too much disruption to continue and shut down instead. It died. Its internal mechanisms tried to keep up with the external world, but something went wrong, its main manual “Homeostasis” reached a definitive value for the ultimate consequence: it ceased functioning. Now MAI's internal workings as well as external actions are not only rooted through internal adjustment, but also in a final consequence found in biological organisms: failure to sufficiently adapt to the environment leads to the cessation of function.

Searle would probably object to this representation of MAI. He’d argue that in spite of this interplay of mechanisms that lies much closer to what we would call intelligence, the computer running the program would still not understand what it was doing. Although Searle initially boasted the claim that “[t]he point of the parable is clear but it is usually neglected. *Syntax by itself is not sufficient for semantic content*”[188], he more than once claims that the Chinese Room proves that a computer can't gain any semantic context.[189] After all, I didn't understand the meaning of the symbols I was manipulating either. The computer is even worse off, having, as far as we know, no mind of its own with which it could try to ground the symbols. It was just manipulating 1's and 0's. In fact, I agree that the computer itself does not understand anything. However, perhaps the requirement that the computer should understand its own programming is a fundamental mistake. Let us examine the connection between the make-up of biological life and the understanding of the intelligence it runs.

The code of life

Searle's objection to computer-based AI is founded on the claim that logical syntax is restricted to formal languages and systems, without regard to any interpretation or meaning given to them. It is also based on the claim that it is somehow the computers *function* to understand the symbols it operates. Computers manipulate this logical syntax to do what they do, but because we as humans impose the meanings upon things we call syntactical. Searle concludes that computer syntax is strictly divorced from meaning without an outside agent or observer.[190,191] No computer is able to

---

[188] Searle, J.R. (2002). *Consciousness and Language* (Cambridge, 2002) 16.
[189] Searle, J.R. (2004). *Mind: A brief introduction* (New York, 2004) 62-64.
[190] Searle, J.R. (2004). *Mind: A brief introduction* (New York, 2004) 62-64.
[191] Searle, J.R. (2002). *Consciousness and Language* (Cambridge, 2002) 16.

form an understanding of its symbols because computation is, according to Searle, the empty shuffling of symbols that provides no means of semantics. I think the computer is entirely the wrong level of the system to explore. In order to back up this claim, I will explore what gives us "meaning" in human intelligence. The following could be perceived as a variant of the system reply, although it goes a bit deeper than that. Contrary to most "System-replies", I will not argue that Searle is looking only at a part of the system. Instead, he is simply looking at entirely the wrong level of organisation.

We know, or perhaps more accurately *feel*, that meaningful content is possible in human intelligence. In other words: biological massive adaptability is capable of generating meaning and attaching it not just to objects, but to our symbolic representations of those objects as well. Somehow, the human brain, which produces our consciousness and our intelligence, enables the production of meaning. Where does this meaning come from? We know that our brain is composed of a great number of individual cells that can be divided into several different cell-types. These cells are interwoven with connections that allow them to excite and inhibit each other through electrical and chemical signalling. Somehow, this conglomerate of interacting cells that compose the brain is capable of producing consciousness, intelligence and meaning. I have speculated in Chapter 2, that these individual cells possess a very rudimentary meaning-mechanic: bad and good, the meaningful representation of death and survival at a cellular level. The reward-systems present in human intelligence are based on this primitive signalling of good actions and bad actions for its implementation of the widespread "good" and "bad" signalling that is so vital for meaning. But this cellular level, or "layer", is not the smallest constituent part of the brain. After all, those brain-cells are composed of smaller parts as well. Is meaning conserved there too? Brain-cells contain DNA and mitochondria, a cellular membrane, enzymes and other structures found in organic cells. It seems strange to posit any conservation of "meaning" at this level of organisation, but perhaps some could still be found in DNA, which not only provides the code from which whole organisms are built, but functions as the storage of eons of experience: death and survival imposed by natural selection. Age-old selection of successful interactions and deselection of unsuccessful interactions reside in the DNA-structure, so maybe there is still some "meaning" here.

To ensure that we have left the organisational levels where meaning still has any sense, we can go down even further: all these structures, including DNA, are built of smaller composite parts. DNA, the very code of life, is nothing more than a molecule, or rather two molecules paired together. This DNA-molecule consists of the sequencing of four smaller molecules (guanine, adenine, thymine and cytosine), which make up the “code”. It is the sequence in which these molecules appear that holds the key to the massive and versatile amount of information stored in DNA. These sub-molecules individually, however, appear to store no relevant information. It is only in concert that they express anything and they only do so in the environment of suitable biomechanics that physically operate on their presence. It seems that at this level, the chemical compounds that make up DNA, all meaning disappears. Still, to be completely sure we have left the domain of “meaning”, we can go even deeper. Again, these smaller molecules can be broken up into smaller parts, the atoms of 5 elements: hydrogen, oxygen, nitrogen, carbon and phosphorus. These same atoms can be used to build constructs that are completely inert and that certainly do not have any interior “meaning” to any “experience” recognised by modern science. Surely at the atomic level all meaning is lost, but even these meaningless atoms can be divided further into subatomic particles that interact with the physical forces that bind our universe according to the Standard Model. I will call the physical forces and subatomic particles “Layer 1” required for biological massive adaptability. “Semantics”, as Searle likes to call meaning, is certainly not present at this lowest layer. Subatomic particles have no “feelings” and interactions with other subatomic particles do not have semantic properties other than that they either occur or do not occur, depending on probability and particle composition.[192] Yet from this meaningless interaction of subatomic particles and physical forces, we *know* that meaning can arise: we are composed of these subatomic particles interacting with physical forces and after several layers of organisation we certainly experience meaning. So how does this apply to Searle’s argument that a digital computer cannot develop meaning?

Any Artificial Intelligence that is constructed as a computer program on a digital computer has *at least* two layers that are important to it:

- **Layer 1**: This layer is comprised of the **Computer** itself, which provides the physical hardware, and the **Symbols**, the 0's and 1's that are manipulated by the computer. This first layer is simply a layer of absolute requirements for an AI built on a computer: there has to

---

[192] Even if someone were to disagree and state that atomic or subatomic particles are somehow responsible for meaning, I refer to the given that our digital computers are, in the end, also composed of subatomic particles. Extending meaning downwards indefinitely makes an even stronger case for potentially “meaningful” computers.

be a machine to operate the program, and there have to be symbols for the program to be expressed in and the computer to manipulate.

- Together these requirements enable **Layer 2**. This layer is the software, or computer program which is constituted of the symbols 0 and 1 arranged in a particular order and moved about by the computer.
- In a Neural Net run on a digital computer, a **Layer 3** is required as well: the Neurons and their interactions are built in a computer program (that may itself also run on another computer program).

Searle's argument is that since the Computer doesn't understand the meaning of the symbols it manipulates, there is no such thing possible as a "Strong AI". This is ridiculous. The computer and its symbols may very well be "Layer 1" for an AI construct. An AI built on the foundations of a Computer and its Symbols is, like biological life and the subatomic particles and forces it's built upon, not necessarily constrained by the computer's lack of semantics in its own establishment of meaning. If it were, we would likewise be kept from meaning by the lifeless compounds that compose us. Atoms and forces simply interact the way they do, because that is what they are, and yet they still allow for meaning at the level of biological massive adaptability. Likewise, computers and symbols simply interact the way they do, because that is what they are, so who is to say that meaning is somehow outside of Strong AI's reach (see Figure 5.1)?

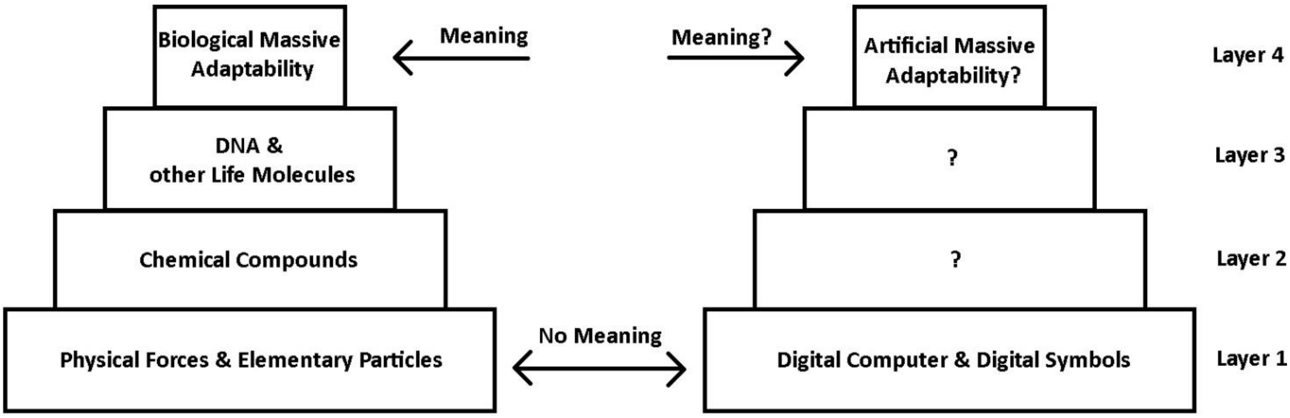

Figure 5.2: Layers of organisation. On the left are some of the layers of organisation that support biological massive adaptability. Although it seems trivial that no meaning is present at the "physical forces & elementary particles" layer, biological massive adaptability is eventually founded on this meaningless layer. Likewise, an AI may also be constructed on initially meaningless layers, the empirical question remaining what programming is necessary to fill in the intermittent layers and

enable a meaningful AI. Searle's Chinese Room "proof" that computers have no semantic insight into the symbols they are manipulating has no impact on the question whether an advanced program constructed on that computer can contain semantic content of its own.

My argument is that, depending on the make-up of the computer program, the computer itself could take on the role of any of the supporting layers required for meaning, without it itself gaining understanding or holding any meaning of its own. For instance, the computer and its symbols could be equated to physical forces and the subatomic particles that are moved by them. After all, the computer moves the symbols and guides the interactions between the symbols. On top of this architecture, a sophisticated computer program could be built. The core of this conception is that the hardware is both absolutely vital and completely insignificant for massive adaptability. Likewise, it is not about the symbols. It does not matter what massive adaptability is built on, be it the stringing together of atoms in just the right way, or the compiling of "arbitrary" symbols into a code in just the right manner: if that manner is capable of interacting with its environment, evaluating the outcome, storing the information in a manner that reflects its own agenda and adjusting its behaviour accordingly, it is adaptable and it can possess meaningful representation. If it is very good at these things, and capable of doing so in a sufficiently complex manner, it is massively adaptable.

So does the computer understand? Is it massively adaptable? Yes and no. The atoms that compose our body are not intelligent, nor do they understand. Our individual brain-cells are unlikely to have a concept of understanding equal to that of our integrated experience. They are, however, crucial for our understanding and intelligence. Likewise, the computer and its symbols, if the architecture is that of an old fashioned digital computer, are both far removed and integral to any AI that is created upon it. It may take several layers of code just to enable it to serve as the basis for operating a Neural Net structure. The physical machine will have no conscious understanding of the symbols it operates, it just moves them about as automated based on the interactions between the symbols. But built upon those symbols, interactions and connections could be created that actually reflect the outside world as well as the consequences that these connections have for the computer program. It is these connections and interactions, and not the symbols or atoms that enable them, that matter. Because the computer program is bound to the computer in the same way that we are bound to our physical body, it may be more expedient to say that if a true AI is composed by a particular computer program on a particular computer, that "computer" is intelligent.

That does not mean that any kind of software on any old computer simply qualifies, adaptability and grounding are still required, but AI-software cannot be excluded from meaning *prima facie*. We must now proceed to a second argument is forwarded as “proof” that computers are uncapable of true intelligence, consciousness or meaning. It is the the Simulation Objection

The Simulation Objection: a real storm in a glass of simulated water

Searle, and others, reject that consciousness, mind, or meaning can be formed on a computer. Searle goes so far as to say that while it may be true that computer programs could be able to think, they are nevertheless unable to be conscious.[193] He outright states in one of his books, although similar statements can be found in the others:

> “*[C]onsciousness is a natural, biological phenomenon. It is as much a part of our biological life as digestion, growth, or photosynthesis.*”[194]

Somehow, through some mystical, unnamed law of nature, a machine can be programmed to think and to learn, but it can never be conscious. Searle calls his doctrine biological naturalism,[195] and only gives two arguments for it. The first is his Chinese Room thought-experiment, which I have already rejected above. The other argument seems more fundamental to the physical world. It is the simulation argument, where the general thesis is that a simulation of a mind does not a real mind make.

To sell the simulation argument, Searle makes the comparison between a simulated mind and a simulated rainstorm over London in order to reveal the “mistake that a computational model of consciousness is somehow conscious”.[196] Just as a simulated rainstorm will not make you wet, a simulated mind would not truly be capable of consciousness, or meaningful intelligence. This is a curious statement. First of all, simulations of rainstorms are not intended to be artificial rainstorms, while the creation of true AI certainly aims for the creation of an artificial mind. Different goals will lead to different results: the wetness of a rainstorm is, for the purposes of weather prediction, not a

---

[193] Searle, J.R. (1980). Minds, brains and programs. *Behavioral and Brain Sciences 3* (3) 417-24.

[194] Searle, J.R. (1997). *The Mystery of Consciousness* (New York, 1997) xiii, but see also: Searle, J.R. (2002). *Consciousness and Language* (Cambridge, 2002) 7.

[195] Searle, J.R. (1997). *The Mystery of Consciousness* (New York, 1997) xiv.

[196] Searle, J.R. (2002). *Consciousness and Language* (Cambridge, 2002) 16.

useful variable to implement in any way. There will be no simulation of wetness on London, or even falling water, as the model is designed to predict the weather, not to simulate a wet environment. This is a stark contrast with attempts at simulating a mind, where intelligent thought, and perhaps emotion and consciousness are the entire point. Searle is confusing a simulated rainstorm, which is really only a model, with attempts at creating a real artificial mind, which, as Jack Copeland put it, may be more akin to creating artificial coal.[197]

There are situations imaginable in which simulating wetness might apply. If one were to build a digital world for recreational purposes, such worlds are already plentiful and continually being improved upon,[198] generating convincing wetness in a rainstorm may at one point become a design goal. Perhaps a virtual reality will be created that can trigger the wetness sensation as experienced by the human brain. In a sense wetness would be created within this virtual reality, although it obviously will not leave the simulation. However intelligence and mind are not exactly the same thing as wetness. Wetness has no action at a distance, it cannot communicate and a wetness expressed in code is unlikely to leak into the physical world. Although I agree with claims that there is no dualistic difference between mind and body, they are all part of the same construct for me, intelligence is about adaptability: interaction, evaluation, storage and adjustment. None of these basic factors require a biological, “non-simulated” body. We know that, unlike wetness, the expression of intelligence and mind is not limited by virtual barriers. Through interaction over the internet, countless minds communicate on a daily basis using symbols that are sent through computers and wires and emerge on the other end with the expression of the mind behind it still intact. And it is not just interaction between minds through a virtual gateway: interaction with virtual worlds is frequently and vehemently undertaken by humans and these virtual worlds function very much as real environments for their interactions.[199] A virtual gateway is no limitation for interaction, which means that a virtual mind may be well equipped to communicate and interact with the physical world, provided it has access to means that can affect the physical world. Unless it was Searle's intention to state that it is impossible for the “simulated” Artificial Mind to leave the

---

[197] Copeland, B.J. (1993). *Artificial intelligence: A philosophical introduction* (Malden, 1993) 47.

[198] Coincidentally, it has been suggested that one type of such digital worlds, interactive computer games, can function as a good research field for Human Level AI to explore natural behaviour.
Laird, J.E. & Lent, M. van (2001). Human-level AI’s killer application: Interactive computer games. *AI Magazine 22* (2) 15-26.

[199] Laird, J.E. & Lent, M. van (2001). Human-level AI’s killer application: Interactive computer games. *AI Magazine 22* (2) 15-26.

computer its program is running on and step out into the "real world", any objection that a simulated mind could have no real world consequences is silly. If, however, that was all that Searle meant to state, his statement has become both trivial and strangely hypocritical: the human mind has been unable to exit the body that is "running" it just as well.

An artificial mind should have no problems interacting with the outside world through the digital portals that already allow our minds to interact with the outside world. This artificial mind could be constructed in the flesh, or silicon rather, by creating or simulating a Neural Net of a large number of interconnected Neuron units. As mentioned before, a physical Neural Net is infeasible for now, due to the almost insurmountable difficulties creating a physical Neural Net entails. Simulating this Neural Net on a computer would yield, according to Searle, a sterile mind capable of thought, but somehow incapable of meaning or consciousness. This is absurd as the presence of consciousness and meaning depends entirely on the kind of simulation. I will now argue just how a conscious, grounded mind could exist inside a computer, in the following section on the Symbol Grounding Problem.

The Symbol Grounding Problem

Secretly, the Chinese Room thought-experiment is an extreme illustration of a bigger problem that has plagued the Philosophy of Artificial Intelligence for decades. This very serious problem is called the Symbol Grounding Problem. Whenever we humans use words we get the feeling that we know what we are talking about, beyond the symbols or sounds the word is composed off. Although words are symbols of the things they represent, we seem to *understand what they mean*. It has proven very hard to come up with a satisfying answer as to how an Artificial Intelligence could come about actually understanding the symbols it is manipulating. Searle's example above highlights the trouble for a symbol manipulating computer to get to the meaning of any of the symbols it is manipulating. It is like trying to learn Chinese from a Chinese/Chinese dictionary without any preconceived knowledge of any language. When trying to find out the meaning of one string of symbols, you are led to another string of symbols, which leads you to another ad infinitum.[200]

One of the solutions proposed to this problem is to somehow "ground" the symbols by

---

[200] Harnad, S. (1990). The symbol grounding problem. *Physica D 42* (1-3) 335-346.

providing the AI with sensors with which to detect the environment.[201] This has met with objections that the sensors may indeed detect their surroundings, but the computer still receives nothing but a string of symbols from the sensors attached to it,[202] which are again intrinsically ungrounded according to some. Pain-detection, a kind of sensory input for humans, also relies on sensors that send electrical pulses to the brain, but we also very much *feel* pain. Yet somehow it seems strange that the symbols streaming in to the computer suddenly are experienced as pain, which is “the unpleasant experience that is associated with actual or potential tissue damage”.[203] Simply attaching some sensors to a computer seems insufficient to ground that computer's symbols.

Another proposed solution to the symbol grounding problem is the application of Neural Nets, which, depending on your definition of symbolic, are or are not symbolic in nature.[204] In the section on the Chinese Room, I have already forwarded some of my arguments as to why a properly constructed Neural Net AI should not immediately be considered an ungrounded straightforward symbol manipulator. I have proposed that the bits that digital computers manipulate, the “ones and zeroes” that form its most fundamental “symbols” should be seen as building blocks from which an adaptive organization could be crafted. This bypasses both the “necessity” for the computer itself to understand the symbols it manipulates and the uncanny observance that in the end, everything the computer does is simply a manipulation of bits. Instead, it is the connections and consequences that can be built up from these symbols that need to find meaning if the construct we built is to be regarded as truly intelligent and conscious. I have already suggested a means of grounding the connections in a way that exceeds the simplistic answer of “attaching some sensors to it”. I will now elaborate my view on the grounding problem and its possible resolving.

What does it mean to be “grounded”?

An important issue regarding the symbol grounding problem is that it has not been properly established what “grounding” is supposed to be. In twenty years, the debate appears to not have gotten much further than the vague comment that whenever we humans say something, do

---

[201] Harnad, S. (1990). The symbol grounding problem. *Physica D 42* (1-3) 335-346.

[202] Copeland, B.J. (1993). *Artificial intelligence: A philosophical introduction* (Malden, 1993) 131.

[203] Eisenberger, N.I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience 13* (6) 421-434.

[204] Harnad, S. (1990). The symbol grounding problem. *Physica D 42* (1-3) 335-346.

something, or in any other way express ourselves *we mean something by it*.[205,206] This is somewhat shocking. After all, if we are unable to express grounding in any other way than an intuition, perhaps its existence is a simple illusion? This may perhaps prove true in the future, but at this point it seems clear to me at least that there are some words that I can learn and pronounce without grasping what they mean, or even what their function is. A quick search-request on Google provided me with a Chinese sentence written down in Pinyin notation: *Chē zǐ zài hòu miàn*. I do not know what it means. I actually do not know if it means anything at all, although I trust that it does. However, I can reproduce the letters and words in writing and may even be able to properly pronounce them if they were read to me or if I was handed a pronunciation guide. I will therefore labour under the assumption that it is indeed possible to manipulate symbols without knowing what they mean, which means that there is such a thing as a “meaning” to understand. And if there is such a thing, then the Grounding Problem is one that needs to be dealt with, even if the nature of the symbols (or atoms, see: The code of life) is irrelevant to their function.

Grounding is the accurate and reliable connection of internal representations, for instance a symbolic representations, to the external objects they refer to. Words, symbols or any other way of representing can then find their basis in actual reality. If a computer program is able to accurately and consistently couple its internal representation to external objects, its internal representation has found a basis in reality and can therefore be said to be grounded through this connection.[207] Any utterance produced by an AI on the subject of a “mouse” can now be taken to actually refer to the real world organism we call a “mouse”, allowing for more connected conversation and instruction. (It is clear why the Symbol Grounding Problem applies so easily to the Chinese Room: the Chinese Room has no way to connect its internal representation to the outside world. There are no sensors with which to detect real world objects and there is no learning mechanic to ensure that internal representation stays connected to the outside world.) However, a reliable connection of internal representation to external objects is only part of the story. Even if a computer equipped with sensors is capable of reliably identifying a mouse as a mouse, this identification still seems devoid of the internal meanings and experiences we so strongly associate with words. This leads to the important

---

[205] Harnad, S. (1990). The symbol grounding problem. *Physica D 42* (1-3) 335-346.

[206] Harnad, S. (2009). From sensorimotor categories to grounded symbols. http://eprints.soton.ac.uk/268030/. *University of Southampton.* (Draft version published under Creative Commons Attribution No Derivatives) 5p.

[207] Nilsson, N.J. (2007). The Physical Symbol System Hypothesis: Status and Prospects. *50 Years of artificial intelligence* (Berlin, 2007) 9-17.

question: how can this grounded internal representation of the external world gain *internal meaning*?

Above, I have rejected the notion that creating a grounded intelligence through symbol manipulation is impossible (see: The code of life). Rather than compare the computer-shuffling of symbols to human mental processes, I have instead equated the computer and its operations to our physical, rather than physiological, make-up. The computer's symbols, with the simple 0's and 1's at their core, could function as the universal particles manipulated by the computer to allow interaction between composite parts. Like the physical atoms and forces in our world, it could be quite possible to construct something more complex from the simple symbols of the computer. Even on an ungrounded, straightforward personal computer, this effect is very visible. The interaction of 0's and 1's with the computer screen and loudspeakers produce interesting visual and auditory effects and patterns that go beyond mere 0's and 1's, even though all associated meaning comes from the human observer. It is therefore possible to create something complex from something simple. Perhaps a computer program can in fact function as a basis for a grounded and meaningful intelligence. So what would be required? Let us review the demands for bare-bone adaptability I laid out in Chapter 1, from the perspective of establishing meaning:

5. A computer-program would need to be able to **interact** with its environment and be able to learn from the interaction.
6. For that interaction to have any kind of grounding in reality as well as any internal evaluation, the computer-program needs to form **connections** between internal representation and external reality, these connections can be expressed in memory storage.
7. To ensure that the grounded interactions have actual internal meaning, the connections and interactions need to have external as well as internal **consequences** that promote and allow for the evaluation of interactions and adaptations.

A computer-program that adheres to the criteria posited above stands a good chance of establishing at least a primitive form of meaningful grounding, the key to which is evaluating **internal consequences**. If a computer-program is able to reliably interact with the outside world and alter its own programming in response to what it's learned from the environment, it would be inextricably linked to the environment. The internal code would depend on what the program had encountered, creating a tight connection between the outside world and the computer code representing it. This primitive understanding should be complemented with connections to internal consequences to greatly improve it. The outside world should be linked to the inside world through an internal representation that makes it possible to differentiate between beneficial and detrimental, based on real world properties as demonstrated in Chapter 4. By linking external signals and internal representation to self-oriented evaluation, these internal representations gain an internal meaning. For that, the external environment needs to be capable of affecting the AI itself. In other words, for an AI that is not only grounded but has internal meaning as well, it needs to be embodied in the world it inhabits. External signals (senses) can now be linked to their internal consequences (homeostasis). Everything else, every action and plan, can now be attached to this additional grounding mechanism through reward or punishment (see Figure 5.2).

In other words, the internal representation of the external environment has now gained an additional dimension: an internal value, a value that directly represents its significance, or insignificance, to the AI itself and the situation it is in. The external information now can be signified with an internal value for the AI, a meaning that goes beyond the reliable identification of a “mouse”. This “mouse” can now acquire a variety of internal meanings based on context: for instance a dead mouse can serve as a biofuel for replenishing energy reserves, while a live mouse gnawing on computer cables is a clear danger. Through external context and internal consequences, normal sensory input and internal representation can be grounded in both the environment and homeostasis. This illustrates once more why learning with internal consequences is the key to true intelligence. It allows for a constant update and revision of the connections between the outside world, the internal representations and the internal consequences. Without a learning mechanism in place that ensures this meaningful grounding, any intelligence would become detached from reality by the passage of time and fail to be adaptable in short order. It would cease, for all intents and purposes, to be intelligent.

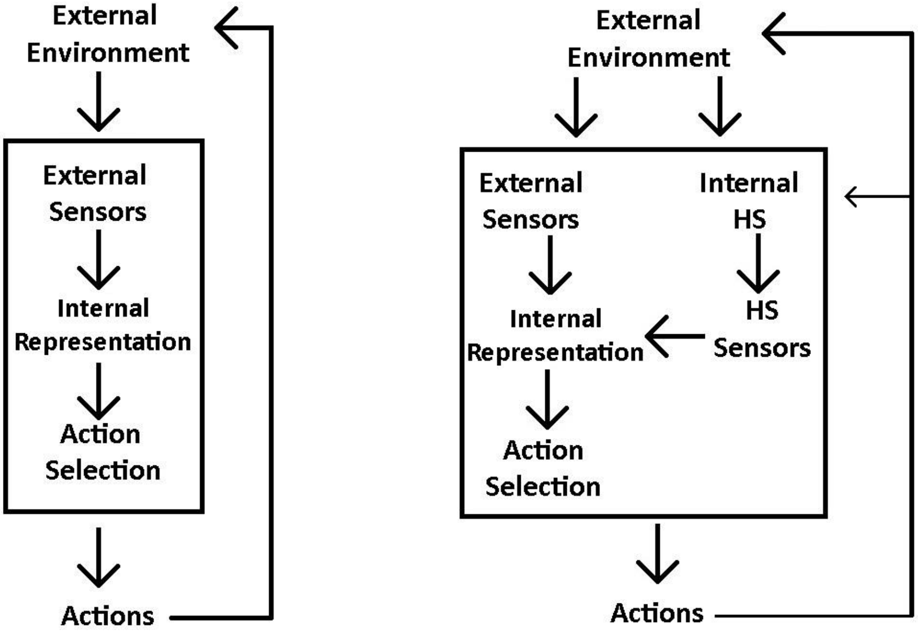

Figure 5.3: Dual grounding in MAI. On the left an AI that is "grounded" through external sensors and reliable identification of its environment, while on the right grounding as approached by MAI is depicted. (The boxed areas denote components of the AI itself.) Grounding in the regular AI (left) only takes a single path which allows for successful external identification through internal representation, but allows no way of determining internal meaning. In MAI (right), external sensors convey their information on the external environment to internal representation, while homeostatic (HS) sensors report on the effects of the external environment and MAI's actions on internal homeostasis. This allows MAI to not only successfully connect internal representation to external sensory information, but also allows it to determine what that external information means to its own survival. Internal representation in MAI is grounded through dual pathways, rather than just one, adding an extra dimension to grounding its internal representation. Furthermore, because internal homeostasis is coupled to the very real consequence of cessation of function ("death"), it is once again firmly rooted in reality.

The grounding of AI through these measures may be further enhanced by the widening of sensory and homeostatic components. Expanding the external sensory array to more frequencies of light and sound, and broader detection of scent and other environmental clues, will allow for a more defined and more accurate connection between internal representation and the outside world it represents. Expanding the homeostatic array to include the monitoring of more kinds of survival/death-conundrums, such as excessive heat or cold, physical impact trauma, radiation and energy-reserve monitoring, allows for a more varied valued input to give meaning to internal representation and context. Perhaps more steps are required to fill out AI-grounding, but the two prongs of external sensors and internal homeostasis-monitoring can already allow AI to reliably refer to the outside world and place value on the external environment based on its own learning experiences.

Now that we have established that meaningful grounding is the result of successfully connecting internal consequences to interactions with the external world, it is time to review one final, as of yet unnamed objection to programming any kind of intelligence, including MAI.

Pre-programming homeostasis

The first steps of programming MAI may have felt quite artificial. We assigned it a range of values for homeostasis to keep track off and then gave it “environmental options” that are preprogrammed to affect that homeostasis. Some may object that this is hardly fair, or even real at all: we are still “predetermining” consequences and therefore meaning! This objection falls apart into two fundamental objections: the first objection is the objection against pre-programming preferences, the second objection is against pre-programming interaction with the outside world. To determine their worth, let us once more turn to the biological example of a grounded, meaningful intelligence: humans.

Humans do not come into this world as a *tabula rasa*. Set in the brain of new-borns are a wide variety of “preprogrammed” reference frames, which include requirements for homeostatic safety and the tools to monitor these. These reference frames exist in order to kick-start the learning curve of the organism and prevent a lot of unnecessary death. These systems are encoded in the DNA code of the organism, which has been modified by millions if not billions of years of evolution to increase the chances of survival and decrease the chances of untimely death. It is in the very code of life that these prerequisite systems are captured and through the early stages of embryonic growth they are expressed as integral parts of the organism and its adaptive circuitry. A child is not an empty vessel, expecting an AI to learn and “be intelligent” from scratch is unreasonable given the hundreds of millions of years genetics and the brain (and the body) have had to develop. MAI does not have millions of years of evolution to pre-program its core instruction set, so these instructions must be composed and delivered in another way. This is essentially what we have done when we have built MAI to observe certain homeostatic values and attempt to stay within them. This is also what we have done when we program MAI to be rewarded when it succeeds in correcting a disturbed homeostasis and to associate negative values with disturbing homeostasis or exacerbating homeostatic disturbance. We have taken the role of millions of years of evolution to instruct MAI in the basic necessities for survival. This is not a problem unless those basic necessities do not actually provide survival consequences. As long as the consequences are real, pre-programming the means to detect those disturbances in order to be able to learn how to avoid the consequences is not problematic.

“But the consequences are fake!” objectors could now exclaim. After all, the consequences detected by the MAI-program are *also* preprogrammed as long as MAI is not implemented in an actual robot body to interact with the actual outside world. The reply to this is twofold: first, if MAI can be instituted in an actual “body”, which can gather actual energy and store it, and which will stop functioning when energy reserves are depleted, MAI's consequences are very much real. A physical body that suffers from the same physical dependencies as biological organisms do (such as being prone to destruction, requiring energy to function, etcetera) will immediately reveal the homeostatic requirements for MAI to be very real. If it does not adhere to the homeostatic requirements for its body, nothing about its demise will be fake. Secondly, even if MAI is not provided with a robot body, but is instead given a simulated body, this does not change the fact that the consequences are still very real. Even though they are instituted by human programmers in the code that governs MAI's survival or death, the consequences are still the same: failure to adhere causes death, success prolongs survival. For those still not convinced: we can see this in biological organisms as well. Milk-producing Dutch cows have been selected to produce absolutely staggering amounts of milk, much more than would be helpful in the wild. In fact, milk cows would most likely quickly go extinct if all human involvement was ceased. However, human involvement isn't stopping. The evolutionary pressure that up-regulated the survival value of producing large quantities of milk is due directly to human involvement. Humans have caused the untimely death of cows that did not produce large quantities of milk (by either deselecting them for procreation, or leading them to actual slaughter) while greatly promoting the survival for the cows that did by instituting breeding programs. Of course, milk-production is a purely physiological characteristic: it has little to do with the brain or behaviour (or maybe it does, hormones secreted by the brain play an important role in milk-production), but the analogy can be easily extended to other domesticated animals such as dogs. Dogs have, since interaction with humans started, evolved to develop, among other things, a more submissive stance towards humans, this is definitely a brain function for which humans have applied selective evolutionary pressure. If we decide to apply survival values and death conditions upon MAI, this is in no way “fake”. If we decide that MAI should greatly enjoy finding more expedient ways to find the solutions to mathematical problems, we can certainly decide that this should be represented in its survival conditions. We would be, in essence, doing nothing fundamentally different from out impact on domesticated animals.

In short, it is okay to pre-program the tools required to prevent untimely death and make intelligent inferences. Even if these are limited to being simulated within a computer, the consequences and the connections made on them will still be very real. The MAI-program, even one run in a simulated environment on an old fashioned computer, would still be grounded in its environment and have internal meaning for its interactions with it.

The meaningful mind

We are our body, we are our mind. The old philosophical school of Dualism is wrong to separate the two, nor are our minds and conscious experiences illusions. Mind and body are one and the same. Our physical and physiological make-up is what enables our mind to exist. More accurately, our mind is an aspect, a consequence even, of our physical and physiological make-up. This does not mean that a mind can only be created from biological matter. Our mind has evolved as part of the selective pressure for greater and greater adaptability. Its purpose is to make valued connections between our interactions with the environment and the consequences these have for our well-being. Any construction that can make those required connections can function as the “hardware” on which a mind is built. That “hardware” may even be built out of the “software” that runs on a computer.

When we have a conscious experience, such as that very basal experience of pain triggered by a pinch in the arm, it is our body that has that experience. Our brain, which is an integral part of that body, registers that pinch through the nerve system, where our mind becomes aware of the sensation. No single cell in the brain coordinates this registration, it is the brain as a whole that registers the experience, although some brain regions do experience more excitement than others. There is no impartial observer in the mind that reads out the minds' computations and is the true “person”, the true interpreter. The neurological evidence suggests that consciousness, mind and intelligence are an integrated network experience.[208] Our “qualitative” experiences, seen by some as irreducible to material causes, are the valued connections in our mind based on the consequences of our interactions. We **are** our bodies and we **are** our brains. Every essence of our being, including

---

[208] Thioux, M., Keysers, C. (2010). Empathy: shared circuits and their dysfunctions. *Dialogues in Clinical Neuroscience 12* (4) 546-552.

consciousness, is completely integrated in our body. This means that when the neurons that govern pain and its affective component fire, we "feel" pain because we **are** in pain. The registration of those firing neurons, informing us of the consequences of our actions and allowing us to make the proper connections, is what pain is. That is what those neurons are for. They signify the consequences of our actions for our survival and provide meaningful connections, so that we may better adapt and prevent negative occurrences from happening in the future. The function of our brain is to enhance adaptability through the establishment of consequences and valued connections, it is this function that, in the end, creates the Meaningful Mind.

It is now time to turn to some examples of AI in practise, to see if the homeostasis-evaluation MAI can be of practical use.

Literature

Copeland, B.J. (1993). *Artificial intelligence: A philosophical introduction* (Malden, 1993).

Eisenberger, N.I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience 13* (6) 421-434.

Harnad, S. (2009). From sensorimotor categories to grounded symbols. http://eprints.soton.ac.uk/268030/. *University of Southampton* (Draft version published under Creative Commons Attribution No Derivatives) 5p.

Harnad, S. (1990). The symbol grounding problem. *Physica D 42* (1-3) 335-346.

Laird, J.E. & Lent, M. van (2001). Human-level AI’s killer application: Interactive computer games. *AI Magazine 22* (2) 15-26.

Nilsson, N.J. (2007). The Physical Symbol System Hypothesis: Status and prospects. *50 Years of artificial intelligence* (Berlin, 2007) 9-17.

Searle, J.R. (2002). *Consciousness and Language* (Cambridge, 2002).

Searle, J.R. (1990). Is the brain's mind a computer program? *Scientific American 262* (1) 26-31.

Searle, J.R. (2004). *Mind: a brief introduction* (New York, 2004).

Searle, J.R. (1980). Minds, brains and programs. *Behavioral and Brain Sciences 3* (3) 417-24.

Searle, J.R. (1997). *The Mystery of Consciousness* (New York, 1997).

Thioux, M., Keysers, C. (2010). Empathy: shared circuits and their dysfunctions. *Dialogues in Clinical Neuroscience 12* (4) 546-552.

Turing, A.M. (1950). Computing Machinery and Intelligence. *Mind 59* (236) 433-460.

University of Reading (8 June 2014). Turing test success marks milestone in computing history. http://www.reading.ac.uk/news-and-events/releases/PR583836.aspx. *Press Releases University of Reading* (retrieved 24 June 2014).

Chapter 6: Reward in Artificial Intelligence Practice

Human-level AI and other humanlike AI projects

At the end of the last century and during the first decade of the current century, some influential founders of AI have reformulated the desire that AI researchers try harder at creating the original goal of Artificial Intelligence: creating an intelligence that is comparable to humans across the board. Two of the approaches in this field are known as human-level AI (HLAI) and Artificial General Intelligence (AGI).[209] In Chapter 1 I defined the type of Artificial Intelligence under investigation in my thesis as:

> *"(...) [T]he quest to manufacture a machine which is characterized by its massive adaptability. [An AI] is a machine that is capable of interacting with its environment, capable of storing these interactions in a meaningful way and able to adjust its future interactions based on these learned experiences."*

As "massive adaptability" is based on human intelligence, it seems obvious that my goals overlap with those of HLAI and AGI in that I am more interested in learning the roots of human intelligence and building an actual all-round intelligent machine, than building specialised one-trick programs. Human intelligence is obviously too complicated to create from scratch. The inescapable founding father of AI, Alan Turing, already speculated on simplifying the quest by creating a childlike intelligence that could be educated to the intelligence level of a human adult, rather than trying to create an adult intelligence from the get-go.[210] Learning, therefore, seems paramount in creating such an educable intelligence. Turing speculated on the implementation of reward and punishment in these "Child-machines", a kind of proto-HLAI, as a possible engine for education. One of the high-profile proponents of HLAI, Nils J. Nilsson, joins Turing in his claim that reinforcement learning is important. Surprisingly, perhaps, he then proceeds to give no further direction other than that the field has important research areas concerning the effective combining of reinforcement learning with perceptual, representation and action hierarchies as well as relational state descriptions.[211] In a more recent paper on learning, not written by Nilsson but based on the TRP-type programming that Nilsson developed, very little attention is given to reinforcement learning

---

[209] Russell, S. & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach; Third Edition* (New Jersey 2010) 27.
[210] Nilsson, N.J. (2005). Human-level Artificial Intelligence? Be serious! *AI Magazine 26* (4) 68-75.
[211] Nilsson, N.J. (2005). Human-level Artificial Intelligence? Be serious! *AI Magazine 26* (4) 68-75.

and no attention at all is given to reward and punishment mechanisms.[212] It seems that the inclusion of reinforcement learning has been somewhat forgotten. I, however, do think that MAI can function as the basis of reward learning as well as offering a basis for sensory-motor system processing, another important part of Nilsson's theoretical AI. It can function especially well as the evaluative basis for prediction, the precondition for planning.

In the following pages I will explore three contributions to the field of AI that have an explicit focus on modelling the way humans think. I will try to apply the functions and structures that went into MAI to these discussed approaches, to see if any beneficial interaction can be established. Due to demands of time and space, the following selection is in no way meant to be a cross-section of the vast field involved with creating humanlike intelligence. They are just three examples of current AI research into humanlike AI. I will present one symbolic-AI project where the focus lies heavily on explicit rule-driven reasoning that is only supported in the background by reward learning, while the other two papers present possible Neural Net learning solutions that strongly involve reward modulation.

Laird et al: Soar

In a 2001 paper, Laird and Van Lent propose that AI-research concerning human-level AI should use "moddable" games as a basis for Artificial Intelligence research. The increasingly complex and realistic pre-existing universe provided by interactive videogames would provide the required and increasingly complex environmental information for the AI to process, while the interaction with the humans playing the games would force game-creators into more and more sophisticated AI that would prove more and more humanlike. Hoping that demand from commercial game-creators for better AI behaviour would fuel human-level AI research, Laird and Van Lent go as far as calling interactive videogames the prospective "killer-app" of HLAI. Thanks to the nature of the videogame environment and the low cost of the pre-existing software architecture in which AI programs and programmers could run amok, videogames were deemed very suitable for incremental research into human-level AI.[213] A paper published in 2002 went into a few details regarding Laird's own experiment with the Unreal Tournament Engine, called Haunt 2,[214] but after another article published in 2004 that project seems to have been put on hold.[215] Later exploration

---

[212] Vargas, B. & Morales, E.F. (2009). Learning navigation Teleo-Reactive Programs using behavioural cloning. *IEEE International Conference on Mechatronics* (ICM 2009) 6pp.

[213] Laird, J.E. & Lent, M. Van (2001). Human-level AI's killer application: Interactive computer games. *AI Magazine 22* (2) 15-26.

[214] Laird, J.E. (2002). Research in human-level AI using computer games. *Communications of the ACM 45* (1) 32-35.

[215] Laird, J.E. (2011). Artificial Intelligence & computer games research. http://ai.eecs.umich.edu/people/laird/games_research.html. *University of Michigan* (retrieved 28 August 2014).

of reinforcement learning in a videogame environment is also found in a 2011 paper on Infinite Mario.[216] Indeed Laird's main project, Soar, uses videogames as well as the real world as a testing-ground for its performance.[217,218]

Soar is a general cognitive architecture that is designed with the ultimate design goal to have the same cognitive abilities as humans. As it is, it is both a program and a theory of mind as the cognitive architecture of Soar is supposed to be a representation of human cognition.[219] Although Soar is intended as a Unified Cognitive Theory, it greatly differs from the biological approach featured in MAI: it is a top-down, explicit, rules-governed, decision-making hierarchy. This stands in stark contrast to MAI's bottom-up approach, which may (and is intended to) produce behaviour that could be described by rules, but does not follow any explicit rules. To further complicate matters, Soar is a symbolic system, whereas MAI is based on Neural Nets. Nonetheless, some implementation of MAI-functionality may still be able to improve Soar functioning.

Although it is only one of at least 4 major cognitive functions utilised by the Soar architecture, some reward-reinforcement learning has been integrated into its design since 2004.[220] The four predominant types of cognitive operations in Soar are chunking, reinforcement learning, episodic memory and semantic memory. Chunking occurs when the current rule-set is insufficient to make a decision on what to do. After consulting its memory banks for relevant further information, Soar will create a new rule on what to do if this particular situation comes up again, so that it doesn't have to go through the decision-making process over and over again in the future. Chunking ensures that Soar has a rule for every situation, otherwise it will create one. The second cognitive function is Reinforcement learning, which is based on external feedback or the achievement of goals and sub-goals. It aids Soar by tweaking the value of operators: actions that Soar can undertake in particular situations, based on the rules provided in the architecture. Reinforcement learning therefore only plays a role in the background of Soar's processing. The third cognitive function, Episodic Memory, stores the raw, uninterpreted memories of past events and goals so that they can be searched through or even repeated again step by step if a similar set of

---

[216] Mohan, S. & Laird, J.E. (2011). An object-oriented approach to reinforcement learning in an action game. *Proceedings of the Seventh AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*. 164-169.

[217] Laird, J.E., Derbinsky, N. & Tinkerhess, M. (2012). Online determination of value-function structure and action-value estimates for reinforcement learning in a cognitive architecture. *Advances in Cognitive Systems 2*. 221-238.

[218] Mohan, S., Mininger, A.H., Kirk, J.R. & Laird, J.E. (2012). Acquiring grounded representations of words with situated interactive instruction. *Advances in Cognitive Systems 2*. 113-130.

[219] Lehman, J. F., Laird, J. & Rosenbloom, P. (2006). A gentle introduction to Soar, an architecture for human cognition: 2006 update. http://ai.eecs.umich.edu/soar/sitemaker/docs/misc/GentleIntroduction-2006.pdf. *University of Michigan* (retrieved 1 September 2014) 1-37.

[220] Nason, S. & Laird, J.E. (2004). Integrating reinforcement learning with Soar. *International Conference on Cognitive Modeling 2004*. 208-213.

circumstances and goals occurs. The final form of Soar's cognitive functions is semantic memory, which stores abstracted structures that are absolute knowledge, such as the rules to baseball. This memory is small and fast, but may occasionally lack detail.[221]

Soar functions by trying to fulfil externally set and defined goals. It will create sub-goals on its own in the pursuit of these set goals, but it has no original drive or motivation. Instead, it must be given tasks by the outside "environment": a human controller. This seems at odds with the official objective to provide a Unified Cognitive Theory: humans have intrinsic goals pursued for other reasons than explicit demands by other humans, yet these are entirely lacking from Soar. Therefore the first contribution MAI could make to the Soar project is a renewed focus on the single goal that is shared by all healthy humans: maintaining homeostasis. This homeostasis could even be monitored by a Neural Net that plugs its rewarding-signals into the Soar mainframe as the basis for the reward-learning that already takes place. As is, reward in Soar is usually given by an external input variable, although it "can also be modified by internal knowledge (additional rules) that generates a reward for sub-goal achievement."[222] This reward, too, is eventually encoded in the environment, namely in the final assigned goal.[223] In contrast, in MAI reward is encoded within the Neural Net architecture: it is the Arbiter that determines when it is produced, it is the Value-assigner that distributes it and it is the individual connections between memory Neurons that store its effects. By internalising reward, Soar could abandon explicitly defined reward in the external environment and instead have its human controller "plug into" the now already internal reward-matrix by connecting homeostatic means to requested goals. Even teacher-student learning can be promoted through this method: positive language from the teacher could tap directly into Soar's rewarding circuits. As Laird et al have proposed, a dedicated and straightforward reward learning mechanism is to be preferred.[224] Nothing seems to me more dedicated and straightforward than a full integration of reward into the system.

It seems the only reason the chunking learning mechanism is necessary is the basic architecture that Soar is set up in: it is a rule-driven system so it requires explicit rules that are to be evaluated. In a 2010 paper by Wang and Laird, Soar is combined with reward learning and

---

[221] Nason, S. & Laird, J.E. (2004). Integrating reinforcement learning with Soar. *International Conference on Cognitive Modeling 2004*. 208-213.

[222] Nason, S. & Laird, J.E. (2004). Integrating reinforcement learning with Soar. *International Conference on Cognitive Modeling 2004*. 208-213.

[223] Nason, S. & Laird, J.E. (2004). Integrating reinforcement learning with Soar. *International Conference on Cognitive Modeling 2004*. 208-213.

[224] Lehman, J. F., Laird, J. & Rosenbloom, P. (2006). A gentle introduction to Soar, an architecture for human cognition: 2006 update. http://ai.eecs.umich.edu/soar/sitemaker/docs/misc/GentleIntroduction-2006.pdf. *University of Michigan* (retrieved 1 September 2014) 1-37.

unsupervised hierarchical categorization to produce its own reward-based learning in a "hunting game". Although this paper efficiently uses reward in conjunction with object identification hierarchies, in my opinion a more flexible system would be achieved if Soar's "reward" was instead founded on a homeostatic monitoring system, which can allow for both goal-selection and a broader applicability of reward. It could then even function as goal-creator. In order to still get the program to play the "hunter" game, Soar's homeostatic "energy" values could be set to "too low" at the start of the trials, prompting it to gather new energy (hunt) and allowing it to find reward in success and punishment in failure.[225]

Soar also has trouble assigning reward in cases where background knowledge is insufficient. Background knowledge can influence reinforcement learning in Soar, but can suffer from the problem that all relevant features of the situation need to be included, but including irrelevant features will slow down learning because two similar instances will be valued as two entirely different instances.[226] MAI does not suffer this problem because the reward signal is global: all active and signified connections are enforced when a reward is triggered, requiring no active pre-selection of reward targets. The overlap across multiple trials as discussed in Chapter 3 takes care of differentiating the actually relevant signals from the static. This process can be even more effective when negative reward signals are also involved and can be further refined by implementing a reward-prediction-error Arbiter (see Chapter 4).

Unlike Soar, MAI does not work through a database of explicit rules. Instead its rule-like behaviour emerges from the system's composition. The Chunking mechanism as presented in Soar is superfluous in MAI (which does not follow rules and therefore doesn't suffer from rules-conflict impasse), any decision MAI makes is automatically weighted and stored in its memory for future retrieval. It seems that MAI may provide a more elegant starting point for humanlike intelligence than the symbolic cognitive architecture of Soar. It is now time to look at an architecture closer to MAI's own, to see if MAI can make a greater contribution there.

Rvachev: neurons as a reward-modulated combinatorial switch

In a 2013 paper, Rvachev presents the application of pyramidal Neurons to act as a "combinatorial switch" in determining connection activation and connection strength.[227] As has been done in the

---

[225] Wang, Y. & Laird, J.E. (2010). Efficient value function approximation with unsupervised hierarchical categorization for a reinforcement learning agent. *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on Web Intelligence (WI-10) and Intelligent Agent Technology (IAT-10)* (2) 197-204.

[226] Laird, J.E., Derbinsky, N. & Tinkerhess, M. (2012). Online determination of value-function structure and action-value estimates for reinforcement learning in a cognitive architecture. *Advances in Cognitive Systems 2*. 221-238.

[227] Rvachev, M.M. (2013). Neuron as a reward-modulated combinatorial switch and a model of learning behavior.

current thesis, Rvachev approaches the learning problem of individual neurons in biological organisms from a reward and punishment perspective, where reward signals lie at the basis of a feedback pathway that informs neurons whether their spiking was beneficial or detrimental for a particular input combination. As also proposed in this thesis, Rvachev postulates that biological reward has a modulating effect of neuronal connection strengths. It works at the neuronal level, directly influencing information storage in neuron-connections. In Rvachev too, a global reward signal modulates action selection through a relatively simple memory-storage process at the neuronal level that tweaks connection strength. This global reward signal triggers long-term enhancement or weakening of the neuron's spiking response to the preceding neuronal input firing pattern, which enables reward-learning on the basis of trial-and-error. In Rvachev, artificial pyramidal Neurons manipulate the signal strength of individual action Neurons by forming synaptic clusters on the dendrites of these Neurons. It is the strength of these clustered connections that is affected by the global reward signal. The summed values of these clusters are then used by the Neuron to determine whether it should fire or not in response to a particular input pattern.[228]

The pyramidal Neuron approach posited by Rvachev could be a possible implementation of how to modulate signal strengths at the multi-Neuronal level, a discussion largely omitted in Chapter 4, where all memory properties were represented (and therefore simplified) by a single set of excitatory and inhibiting Neurons. There are, however, some important differences between Rvachev's Neurons and those featured in MAI. Rvachev's action Neuron's operate on an "on/off" principle,[229] while the Neurons in MAI's action preparation layer are all on after being prompted by the action-selector, it is only their signal strengths that vary. Rvachev's Neurons also don't seem to feature a separate weighted signal for inhibition. Instead the program elaborated upon in the Discussion features a blanket reset of Neuronal synaptic cluster value when a negative reward signal is released. I will return to this in a moment, but let us first consider another important design difference: the lack of a reward-determining mechanism.

Rvachev largely excludes the source of the global reward signals from the discussion. There is a short reference to a possible origin of the signal, which he calls '"elementary" reward generators' (such as those reflecting pain or satisfaction of hunger).[230] This seems to refer to an un-modelled structure similar to the homeostasis-Arbiter-Value-assigner relationship described in

---

Neural Networks 46 (2013) 62-74.

[228] Rvachev, M.M. (2013). Neuron as a reward-modulated combinatorial switch and a model of learning behavior. *Neural Networks 46* (2013) 62-74.

[229] Rvachev, M.M. (2013). Neuron as a reward-modulated combinatorial switch and a model of learning behavior. *Neural Networks 46* (2013) 62-74.

[230] Rvachev, M.M. (2013). Neuron as a reward-modulated combinatorial switch and a model of learning behavior. *Neural Networks 46* (2013) 62-74.

Chapter 4 of this thesis. He also names reward prediction error-mechanisms,[231] which are the inspiration for the more advanced Arbiter-mechanic briefly discussed but not expanded upon in Chapter 4, as possible sources of a global reward signal. Due to the fact that the origin of the reward-signal is not really addressed, Rvachev misses an obvious possible solution to some of the learning problems of the reward learning program that he elaborates on in his discussion.

His Neural Net-program, founded on the pyramidal neuron approach, was intended to learn the general detection of apples to eat and stones to discard using reward. To make matters complicated, the apples and stones has a total of 12 possible features, some of which could only occur when others were absent (e.g. three different sizes). Of these features, multiple overlapped (both a stone and an apple could be red and round, for instance), while others were unique to one particular type of object (e.g. having a stem). A large global reward signal (an addition of 0.25 “weight” to active clusters) would be transmitted if the program ate an apple, while a much smaller global reward signal (an addition of 0.01 “weight” to active clusters) would be transmitted if the program discarded an object. A cluster’s weight would be reset to 0 if it had been excited while an associated action neuron fired and triggered a negative global reward signal. This meant that any clusters active during a decision that resulted in a negative signal, would be reset to 0.[232]

This Neural Net could run into trouble if it decided to try pushing apples off the table considerably more often than it decided to try to eat them: pushing the apple off would also give a reward, it was just much smaller than eating it would give. If the Network threw away apples often enough, it would learn this as the standard reaction to all objects and no longer try the eating response. Rvachev addressed this problem (which led to the majority of the 4.5% failures of the network to pass the tests correctly) by replacing the random action selection mechanism that prompted exploration when not enough valued information was present in the network, with a cycling “round-robin” mechanism that would ensure that each option was selected evenly in the case of under-excitation. This reduced failure to pass the tests to 1.7%.[233]

The apple-throwing problem could in my opinion also have been addressed through a more elegant Network design, taking into consideration motivation as it occurs in biological organisms. Although the network was characterised as “hungry” (as a justification for the eating of apples giving a large positive reward) this hungriness was not actually embodied in the network. Nowhere

---

[231] Rvachev, M.M. (2013). Neuron as a reward-modulated combinatorial switch and a model of learning behavior. *Neural Networks 46* (2013) 62-74.

[232] Rvachev, M.M. (2013). Neuron as a reward-modulated combinatorial switch and a model of learning behavior. *Neural Networks 46* (2013) 62-74.

[233] Rvachev, M.M. (2013). Neuron as a reward-modulated combinatorial switch and a model of learning behavior. *Neural Networks 46* (2013) 62-74.

in the Network was there any sign of why the Network would experience reward if an apple was eaten or a rock was thrown. Another problem lies in the fact that punishment (or negative reward) has not been modelled as a separate, opposed, vector, but instead has been implemented as a reset of the rewarding value in erroneous clusters to 0. The obvious solution to Rvachev's apple-throwing problem lies in embodying this "hungry" aspect through homeostasis, as well as permitting a weighted modifying inhibiting signal rather than a resetting signal based on negative reward signals, such as done in MAI. If that were done, throwing away an object would be marked by the Arbiter as the expenditure of a (small) amount of energy, rather than a small benefit as in Rvachev. Thanks to the weighted negative reward signal, clusters would be weakened in their response, rather than reset, leading to a preference not to throw away objects. If attempting to eat a stone was then coupled by the Arbiter to a stronger weighted negative reward signal, throwing away an object would still be preferable if the alternative was something worse, such as eating a stone, but would otherwise be something to avoid when hungry. Any network that set out on throwing things away would, if following MAI's design philosophy, establish that throwing away things that may be potential food is harmful to homeostasis and would instead try out the other options. Actually embodying homeostasis would, when hungry, favour trying to eat things over not trying to eat things, doing away with the apple-throwing problem.

All things considered, the paper by Rvachev demonstrates that a global reward signal can indeed be used to modulate Neural Net signal strengths. This mechanic can, as demonstrated, function as the basis for generalised learning as also proposed in this thesis. Enforcing Neuronal strengths through reward signals can give the outside impression of "intelligent deduction" of aspects that signal an apple. According to Rvachev's report, the network would indeed learn to pass the tests correctly in 95.5% of the unmodified cases and 98.3% of the round-robin cases. This result can be reformulated to say that the Network "learned" to generalise the characteristics of an apple versus a stone on the basis of rewards and without any explicit rule-formulating or rule-following.[234] To see if this successful result of global reward signal-driven learning holds up, a review of another reward-learning Neural Net is in order.

Rombouts et al: Neurally plausible reinforcement learning of working memory tasks.

In their 2012 paper, Rombouts et al also tackle the problem of reinforcement learning through a global reward signal and its effects at the neuronal level.[235] Their model, called AuGMEnT

---

[234] Rvachev, M.M. (2013). Neuron as a reward-modulated combinatorial switch and a model of learning behavior. *Neural Networks 46* (2013) 62-74.

[235] Rombouts, J.O., Roelfsema, P.R. & Bohte, S.M. (2012). Neurally plausible reinforcement learning of working

(Attention-Gated MEmory Tagging), is also inspired by neuroscience and uses the modulation of synaptic weights as the mechanism to modulate connection strengths. These synaptic weights are updated through a mechanism strongly reminiscent of the one proposed in MAI: the winning memory unit (the one that actually gets to express its action), not only expresses the action but also loops back to the synapses that have been responsible for its (feedforward) activation. This attentional feedback process, as Rombouts et al call it, informs the rest of the network about which action was taken and in the process it creates synaptic “tags” (akin to the more abstract "priming” mentioned in Chapter 4 of the current thesis) on the involved synapses. These tags expose the involved neuronal connections to a global “neuromodulatory signal” (a global reward signal also posited in Chapter 4 and Rvachev) that has been transmitted by a prediction error detection mechanism (the more advanced Arbiter-type shortly mentioned in Chapter 4). This global reward signal interacts with the tags and permits learning, even if some time passes between synaptic activity and the animal’s choice. Of course, the strength of the modulation fades over time.[236] This allows for the memory modulation effects mentioned in Chapter 3, as well as enabling more time-differentiated reward connections to be made, up to a point.[237]

However, some important differences between AuGMEnT and MAI are also noticeable. First of all, Rombouts et al do not discuss the method through which the rewarding signal is delivered to the tagged Neurons, leaving the actual distribution of the signal unidentified. Secondly, AuGMEnT does not seem to make use of specialised inhibiting Neurons, but instead passes inhibitory signals down through the same (unknown) pathway it transmits excitatory signals (the inhibitory variant is simply a negatively signed version of the positive reward value). This means that there is no inherently different mechanic for negative feedback, unlike the mechanic proposed in MAI. This has consequences for both the delivery of the signal as well as the structure of the network. The third important difference is that AuGMEnT is entirely fixed on reward prediction error evaluation: it has no homeostasis as a foundation for reward signals and it does not monitor actual effects of the environment on itself. It only evaluates the quality of its predictions. This sadly robs AuGMEnT of one of the grounding pathways, as well as leaving little leeway for the option of enabling different types of reward signals for different types of events. This possibility is preserved with multiple homeostatic parameters to be monitored, such as is possible in MAI.

There are also potential upsides to the AuGMEnT approach. The synaptic weights from

---

memory tasks. *Advances in Neural Information Processing Systems 25* (NIPS 2012) 1871-1879.

[236] Rombouts, J.O., Roelfsema, P.R. & Bohte, S.M. (2012). Neurally plausible reinforcement learning of working memory tasks. *Advances in Neural Information Processing Systems 25* (NIPS 2012) 1871-1879.

[237] Rombouts, J.O., Roelfsema, P.R. & Bohte, S.M. (2012). Neurally plausible reinforcement learning of working memory tasks. *Advances in Neural Information Processing Systems 25* (NIPS 2012) 1871-1879.

input Neurons to memory cells indeed became strongly correlated to the true weights of the represented symbols. This indicates that the global reward feedback mechanism indeed is successful in integrating sensory information into memory connection strengths. AuGMEnT, like Rvachev's method, demonstrates that implementing reward feedback at the neuronal level is a viable way to strengthen or weaken connection-strengths and thereby training Neural Nets. Rombouts et al have also demonstrated that this technique will also function in greatly up-scaled Neural Networks, although a corresponding decrease in the modulatory strength of the reward signals is required.[238] A property not considered by me in Chapter 4 but a consequence of the AuGMEnT-method is that association units with a strong feedforward connection also have a strong feedback connection. This entails that synapses on association units that provided a strong input for the actually selected action, will also respond stronger to a following reward signal. Those connections that have a large effect on the outcome of the firing pattern therefore are also most strongly adjusted with the following reward.[239] Although this functionality is currently not present in MAI architecture, it can in all likelihood be easily implemented. All things considered, the paper by Rombouts et al also illustrates the validity of using a global reward signal to train Neural Networks on the basis of a global reward signal.

The lack of homeostasis

From a philosophical point of view, all three discussed AI architectures have something to gain from implementing homeostasis and homeostasis monitoring as the basis for their reward signals. In doing so, the programs designed can more freely interact with the real world, because they are now susceptible to un-programmed valued input: real world occurrences that hurt or help their homeostatic maintenance can form the basis of valued interaction with the world without the necessity for direct human intervention. By integrating the reward value in a way that is inherently connected to the outside world, these AI get access to a second path of grounding. Not just one that ties outside observations to inside representations, but also one that ties outside observations to internal consequences. Now, if the AI is confronted with an apple it is not limited to simply correctly identifying it as an apple, but it may also have a meaningful response grounded on internal valence. It may not just describe the apple as round, red and smooth, but may also be motivated to eat it when its hungry, or to leave it where it is when it is not without being specifically trained by

---

[238] Rombouts, J.O., Roelfsema, P.R. & Bohte, S.M. (2012). Neurally plausible reinforcement learning of working memory tasks. *Advances in Neural Information Processing Systems 25* (NIPS 2012) 1871-1879.

[239] Rombouts, J.O., Roelfsema, P.R. & Bohte, S.M. (2012). Neurally plausible reinforcement learning of working memory tasks. *Advances in Neural Information Processing Systems 25* (NIPS 2012) 1871-1879.

humans or another computer to do so. The importance of the apple is no longer limited to external context, but has acquired an internal context, an internal *meaning* as well.

From a practical point of view, the lack of a homeostatic monitoring system may also be regrettable. The advantage of including grounded reward systems in learning AI is that it can be "repurposed" to apply to all kinds of learning task. Rather than encode new reward attachments to the newest learning task, the AI can now be enticed to learn by using its already present preference for (say) apples as a reward-mechanism. This makes adding new learning topics much more natural, using an intrinsic motivation of the construct to drive its behaviour. Direct access to the Signifier may even allow for spoken commands, or other methods, to give the AI actually valued and grounded feedback. An AI equipped with homeostatic-grounded reward-systems can be shown any new task and be given rewards for doing it well or, if necessary, punishment for doing poorly. This learning mechanic, so integral to the learning of humans, would add much to AI's *general* learning capabilities and would put it one important step closer to massive adaptability.

Literature

Laird, J.E. (2011). Artificial Intelligence & computer games research. http://ai.eecs.umich.edu/people/laird/games_research.html. University of Michigan (retrieved 28 August 2014).

Laird, J.E. (2002). Research in human-level AI using computer games. *Communications of the ACM 45* (1) 32-35.

Laird, J.E., Derbinsky, N. & Tinkerhess, M. (2012). Online determination of value-function structure and action-value estimates for reinforcement learning in a cognitive architecture. Advances in Cognitive Systems 2. 221-238.

Laird, J.E. & Lent, M. Van (2001). Human-level AI's killer application: Interactive computer games. *AI Magazine 22* (2) 15-26.

Lehman, J. F., Laird, J. & Rosenbloom, P. (2006). A gentle introduction to Soar, an architecture for human cognition: 2006 update. http://ai.eecs.umich.edu/soar/sitemaker/docs/misc/GentleIntroduction-2006.pdf. *University of Michigan* (retrieved 1 September 2014) 1-37.

Mohan, S. & Laird, J.E. (2011). An object-oriented approach to reinforcement learning in an action game. *Proceedings of the Seventh AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*. 164-169.

Mohan, S., Mininger, A.H., Kirk, J.R. & Laird, J.E. (2012). Acquiring grounded representations of words with situated interactive instruction. *Advances in Cognitive Systems 2*. 113-130.

Nason, S. & Laird, J.E. (2004). Integrating reinforcement learning with Soar. *International Conference on Cognitive Modeling 2004*. 208-213.

Nilsson, N.J. (2005). Human-level Artificial Intelligence? Be serious! *AI Magazine 26* (4) 68-75.

Rombouts, J.O., Roelfsema, P.R. & Bohte, S.M. (2012). Neurally plausible reinforcement learning of working memory tasks. *Advances in Neural Information Processing Systems 25* (NIPS 2012) 1871-1879.

Russell, S. & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach; Third Edition* (New Jersey 2010).

Rvachev, M.M. (2013). Neuron as a reward-modulated combinatorial switch and a model of learning behavior. *Neural Networks 46* (2013) 62-74.

Vargas, B. & Morales, E.F. (2009). Learning navigation Teleo-Reactive Programs using behavioural cloning. *IEEE International Conference on Mechatronics* (ICM 2009) 6pp.

Wang, Y. & Laird, J.E. (2010). Efficient value function approximation with unsupervised hierarchical categorization for a reinforcement learning agent. *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on Web Intelligence (WI-10) and Intelligent Agent Technology (IAT-10)* (2) 197-204.

Concluding Remarks

At the start of this thesis I formulated the following research question:

> "*In what manner can biological reward and punishment systems be integrated in the Neural Net-approach to creating Artificial Intelligence with humanlike learning and recognisable intelligence? What are the consequences of such a Natural AI for the field of Philosophy of AI?*"

It is now time to answer this question.

In Chapter 1 I argued for the replacement of the term "intelligence" with the term "massive adaptability" as forwarded by Jack Copeland. I reduced this massive adaptability to something I called "bare-bone adaptability", where only the essentials of adaptability are conserved. I posited that massive adaptability, and with it intelligence, is built from a multitude of small mechanisms that each allow for bare-bone adaptability.

The essentials for bare-bone adaptability are:

- A "being" must be capable of **interaction** with its environment (requiring some form of perception and some means of altering itself or the environment),
- A "being" must be capable of **evaluating** its interactions with the environment,
- A "being" must be capable of **storing** these valued interactions, (more commonly known as having a "memory"),
- A "being" must be capable of **adjusting** its interactions **based on these values** attained through previous interactions/perceptions (more colloquially known as "learning").

When I explored the relation between biological reward and punishment systems and this bare-bone adaptability, I discovered the massive importance of reward and punishment for biological evaluation. When exploring the evaluating structures in biological organisms, a few mechanisms became apparent as crucial components of any Motivated AI (MAI). The first discovery was that internal homeostasis, the ideal parameters of a variety of physical attributes such as PH-value and energy levels, is firmly connected to biological adaptability. It is through the consequences of actions for internal homeostasis that organisms are able to attach meaningful connections to their internal representation of the outside world. In other words, monitoring homeostatic disruption plays a major grounding role in biological learning, driving it forward. Biological adaptability, and with it biological intelligence, is at the fundamental level embodied in homeostatic consequences, because maintaining homeostasis is directly coupled to natural selection. Failure to maintain homeostasis results in death, while restoration of homeostasis promotes survival. Monitoring homeostasis therefore allows organisms to make valued judgments without unnecessary death.

The second noteworthy discovery was that developing a brain is only one way of providing adaptability. It became clear that even organisms as deceptively simple and limited as bacteria had some learning capacity that allowed them to adapt during their lifetime rather than just through brute-force natural selection. In the exploration of that learning-mechanism, it became apparent that bacteria are capable of detecting good from bad, learning new good and bad signs, and acting accordingly. Thanks to the grounding of their experiences in the consequences they have for internal homeostasis, bacteria are capable of learning to distinguish good from bad and attach value to previously valueless external signifiers.

Since bacteria are single-celled life forms, this suggests that it is already possible at the cellular level to distinguish between good and bad, allowing for valued judgement. This opened the door to a third important part of MAI: valued communication. A fundamental grounding mechanism for cellular valued judgement lies in the maintenance of homeostasis, a requirement for any organism that seeks to survive. On the basis of this homeostasis, bacteria are able to attach positive or negative value to chemical signals they pick up from their environment, which among other things allow them to communicate among each other in a meaningful fashion. I posited that cellular evaluation and this chemical communication between cells plays an important role in complex multicellular life as well.

When exploring detection of good and bad in life with brains, it became apparent that brains are quite invested in the subject. The abilities discovered in bacteria were present in the brain as well, but with additional complexity. The brain has many important regions devoted to the detection of homeostatic disturbance, as well as regions preoccupied with distributing valued information to the other brain cells in other to promote or inhibit their activity. This intercellular communication is likely experienced at the macro-level as “reward and punishment”. A combination of macroscopic homeostatic evaluation and chemical signalling allows brain cells to strengthen beneficial connections, while weakening detrimental ones, thereby manipulating each other’s behaviour. This process is so potent that it even allows for connecting already valued signals to non-valued signals, allowing the non-valued signals to become valued as well. This process, called conditioning, still plays a major role in the most massively adaptable organism we know of: the human. Disruption of these value-signalling mechanics leads to the breakdown of all kinds of aspects of human intelligence, such as insight, planning and decision-making. Because motivation also plays a large role in conscious learning, by driving action for instance, brains that do not process reward[240] properly will often display dysfunctional learning and/or lack of action. Reward signals, positive or negative, are also known to have a more general impact on learning. Learning occurs not only through direct motivation, but through memory modulation as well: the mere transmission of a rewarding signal is enough to improve memory storage, presumably because the rewarding signals strengthen the connections of active neurons. This makes sense if the transmission of rewarding chemicals indeed enforces active connections. At the macro-level this mechanic, although potentially a side effect of the other signalling function, also benefits the organism: the release of reward signals tends to be a good indicator that something important has happened that is worth remembering.

Extrapolating from this information on reward in the brain, I abstracted a few functionalities important for motivated learning in biological organisms and adjusted them for the implementation in a new AI model I called Motivational AI (MAI). This MAI has the following modules to improve adaptation to its environment:

- The first I called the Arbiter, which monitors homeostasis for disruptions and restorations. Although I refer to the Arbiter as a singular structure, multiple Arbiter-like mechanisms are likely in place to monitor a variety of homeostatic values. Furthermore, the detection of reward-prediction error is also likely done by a mechanism similar to, if not the same as, the Arbiter.
- The second structure I called the Reward-assigner, which receives input from the Arbiter and releases the relevant reward signals to the rest of the network. It may be possible that the Value-assigner services multiple Arbiters, but there may also be multiple Value-assigners to correspond with different Arbiter tasks.
- The third structure is the Action Selector, which is prompted by the Arbiter when homeostasis is disrupted and which relies on Memory modulated by the Value-assigner to facilitate action selection.

---

[240] The word “reward” when used in a sentence without “punishment”, implies both positive and negative rewards unless it has been specifically noted otherwise.

These three functionalities are connected to each other and integrated in a Neural Net which they provide with an intrinsic valued learning ability grounded in internal consequences to external reality.

At the Neuron level, the Artificial Neurons used in MAI has two characteristics that are worth noting. The first is that the Neurons used by the MAI architecture are either inhibitory or excitatory, not both. This inhibitory or excitatory characteristic is directly linked to respectively the negative or the positive global reward signal released by the Value-assigner. The second Neuron-characteristic is that they can be primed to accept reward signals (in their default mode of operation, they do not). This allows MAI to only enforce Neurons that actually contributed to taken actions. The emerging mechanics of internally valued adaptability can be increased in complexity by adding more Arbiter modules, opening MAI to more homeostatic sensitivities and adding modules that compare expected outcomes to actual outcomes. MAI can also give internal value to other Neural Net memory and sensory modules that are plugged into it, allowing for better self-evaluation and therefore better valued behaviour. Due to its evaluation of information on the basis of recognisable internal homeostasis, MAI has a much more recognisable potential intelligence. However, until an actual complex MAI has been constructed and tested, it is difficult to judge its actual potential.

On a philosophical level, too, MAI may have a significant impact. The homeostatic approach to bare-bone adaptability calls into question the validity of the mantra “Minds are to Brains what Software is to Hardware”. If adaptability through connections to homeostatic consequences is the measure of bare-bone intelligence and intelligence is indeed the massive adaptability as Copeland forwards, the type of underlying physical matter is, perhaps ironically, of no consequence for embodied intelligences or “minds”. This even means that some of the underlying “physical” matter can be a simulation itself. Following this logic, MAI should not be evaluated as a software program running on computer hardware, but as a bunch of variable connections between outside consequences and internal actions valued on internal consequences. This immediately invalidates both the Chinese Room argument and the Simulation-objection: it is not the computer that MAI is stored on that is capable of understanding its symbols, but rather the various connections and consequences to the real world entailed in the MAI-simulating program that directly attach it to an understanding of the outside world. The adaptable connections in MAI are grounded in both external reality and internal consequences, effectively representing both. The symbols that can be used to build these connections are merely the physical constituents, or “building blocks”, of an intelligence that is grounded in external consequences and that attains internal meaning through the relation between external and internal consequences. The Symbol

Grounding Problem and Simulation Objection both don't seem to apply to an adaptable system that has internal representation and evaluates consequences of external events. It is not about the meaning of the symbols used, but their effect. The fact that the homeostasis upon which all of this depends has been preprogrammed by humans is irrelevant: humans are preprogrammed by hundreds of millions of years of evolution and the consequences for MAI are still very much real. If MAI fails to respond to internal consequences in an adequate manner, it will cease functioning. Likewise, any real power projected by MAI into the world due to access to external components, be they arms or a text-screen, is very real as well. This does not mean, however, that MAI as it has been modelled so far, is "massively adaptable" and should be considered intelligent in the same degree that we are: it is by far not complex and versatile enough to qualify for such strong statements. MAI is, however, definitely adaptable and it is well grounded in the external environment. It may function as an important stepping stone in achieving AI capable of evaluating their environment.

The consequences of MAI for already existing AI, such as Laird's Soar or Rombouts et al's AuGMEnT are visible at both the philosophical and the practical level. At the philosophical level, these AI miss out on a main pathway to the grounding of their internal knowledge representations. As long as they are without homeostatic consequences that are connected to their global reward signals, they can only achieve grounding through reliably connecting internal representations to outside events. With the implementation of a MAI-like architecture, this single grounding root can be joined by a second root that provides internal consequences based on the outside world. Internal representation can now not only reliably recognise external objects, but can also attach an internal meaning to them as well. At the practical level, MAI allows for additional versatility in the application of reward learning. The two reviewed Global Reward Signal approaches to training Neural Nets show that a Global Reward Signal as also featured in MAI can indeed be used to train a Neural Net AI. Making reward-learning directly connected to homeostatic values, which can again be directly affected by a large variety of real world events, allows for great repurposing options in teaching AI information. It offers both a more accessible handle for human teachers to express what is important to an implementation of MAI and the opportunity for the AI to explore and learn things for itself. A reliable and multi-functional reward-matrix allows for much greater versatility in learning, making any AI, at least potentially, significantly more adaptable and less dependent on human guidance.

In short, it seems that biological reward and punishment systems have found a preliminary integration into the Neural Net approach to creating Artificial Intelligence through the MAI-model. Founded on the biological mechanisms of adaptability and homeostasis, MAI provides new learning methods for Natural AI and deals with some of the more tenacious problems in the Philosophy of AI as well. Further exploration of the subject as well as actual experimentation will undoubtedly shed further light on the characteristics of human intelligence and the potential construction of a true AI.

Bibliography

Agah, A. (ed.) (2014). *Medical Applications of Artificial Intelligence* (Boca Raton, 2014).

American leprosy missions (2014). Leprosy frequently asked questions. http://www.leprosy.org/leprosy-faqs/. *American leprosy missions* (retrieved 6 March 2014).

Armstrong, J.S. (2012). Natural learning in higher education. In: Seel, N. M. (ed.). *Encyclopedia of the sciences of learning* (2012) 10p (page numbers unknown).

Baker M.D., Wolanin, P.M. & Stock, J.B. (2005). Signal transduction in bacterial chemotaxis. *BioEssays 28* (1) 9-22.

Barbano, M.F. & Cador, M. (2007). Opioids for hedonic experience and dopamine to get ready for it. *Psychopharmacology 191* (3) 497-506.

Bechara, A. (2004). The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage. *Brain and Cognition 55* (2004) 30-40.

Berridge, K.C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology 191* (3) 391-431.

Bos, P.A., Panksepp, J., Bluthé, R.M. & Van Honk, J. (2012). Acute effects of steroid hormones and neuropeptides on human social-emotional behavior: A review of single administration studies. *Frontiers in Neuroendocrinology 33* (1) 17-35.

Brahic, C. (16 July 2014). Meet the electric life forms that live on pure energy. http://www.newscientist.com/article/dn25894-meet-the-electric-life-forms-that-live-on-pure-energy.html?page=1#.U8fh_LGvTX4. *New Scientist 2978* (retrieved 21 July 2014).

Brooks, R.A. (1990). Elephants don't play chess. *Robotics and autonomous systems 6* (1) 3-15.

Bush, G., Vogt, B.A., Holmes, J., Dale, A.M., Greve, D., Jenike, M.A. & Rosen, B.R. (2002).

Dorsal anterior cingulate cortex: a role in reward-based decision making. *Proceedings of the National Academy of Sciences 99* (1) 523-528.

Camalier, C.R. & Kaas, J.H. (2011). Chapter 9: Sound. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

Chatterjee, S., da Silveiram R.A. & Kafri, Y. (2011). Chemotaxis when bacteria remember: Drift versus diffusion. *PloS Computational Biology 7* (12) Special section 5. 1-8.

Chikazoe, J., Lee, D.H., Kriegeskorte, N. & Anderson A.K. (2014). Population coding of affect across stimuli, modalities and individuals. *Nature Neuroscience 17* (8) 1114-1122.

Cook, N.D. (2008). The neuron-level phenomena underlying cognition and consciousness: synaptic activity and the action potential. *Neuroscience 153* (3) 556-570.

Copeland, B.J. (1993). *Artificial Intelligence: A philosophical introduction* (Oxford, 1993).

Craig, A.D. (2003). A new view of pain as a homeostatic emotion. *Trends in neuroscience 26* (6) 303-307.

Damasio A. & Carvalho G.B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience 14* (2) 143-152.

Danchin, E., Charmantier, A., Champagne, F.A., Mesoudi, A., Pujol, B. & Blanchet, S. (2011). Beyond DNA: integrating inclusive inheritance into an extended theory of evolution. *Nature Reviews: Genetics 12* (7) 475-486.

Decety, J. & Svetlova, M. (2012). Putting together phylogenetic and ontogenetic perspectives on empathy. *Developmental Cognitive Neuroscience 2* (1) 1-24.

Dictionary.com (2014). http://dictionary.reference.com/browse/artificial?s=t. *Dictionary.com* (retrieved 14 January 2014).

Dobrindt, U., Chowdary, M.G., Krumbholz G. & Hacker, J. (2010). Genome dynamics and its impact on evolution of *Escheria coli*. *Medical Microbiology and Immunology 199* (3) 145-154.

Eccleston, C. & Crombez, G. (1999). Pain demands attention: a cognitive-affective model of the interruptive function of pain. *Psychological Bulletin 125* (3) 356-366.

Eisenberger, N.I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience 13* (6) 421-434.

Eisenberger, N.I., Jarcho, J.M., Lieberman, M.D. & Naliboff, B.D. (2006). An experimental study of shared sensitivity to physical pain and social rejection. *Pain 126* (1-3) 132-138.

Federle, M.J. & Bassler, B.L. (2012). Interspecies communication in bacteria. *The Journal of Clinical Investigation 112* (9) 1291-1299.

Fellows, L.K. (2011). Chapter 16: The neurology of value. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

Fields, R.G. & Stevens-Graham, B. (2002). New insights into neuron-glia communication. *Science 298* (5593) 556-562.

Fine, J. (1997). Deep Blue wins in final game of match; Chess computer beats Kasparov, world's best human player. http://faculty.georgetown.edu/bassr/511/projects/letham/final/chess.htm. MSNBC (retrieved 6 June 2014).

Gottfried, J.A. (2011). Preface. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Chicago, 2011).

Gottfried, J.A. & Wilson, D.A. (2011). Chapter 5: Smell. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

Gray, P. (2002). *Psychology; Fourth Edition* (New York, 2002).

Gregory, F. (2008). Chapter 17: New ideas about life and its past. *Natural science in western history* (Boston, 2008).

Grosberg, R.K. & Strathmann, R.R. (2007). The evolution of multicellularity: a minor major transition? *Annual Review of Ecology, Evolution, and Systematics 38*. 621-654.

Gyles, C. & Boerlin P. (2014). Horizontally transferred genetic elements and their role in pathogenesis of bacterial disease. *Veterinary Pathology 51* (2) 328-340.

Harnad, S. (2009). From sensorimotor categories to grounded symbols. http://eprints.soton.ac.uk/268030/. *University of Southampton* (Draft version published under Creative Commons Attribution No Derivatives) 5p.

Harnad, S. (1990). The symbol grounding problem. *Physica D 42* (1-3) 335-346.

Hathcock, J. N. (1989). High nutrient intakes – the toxicologist's view. *Symposium Upper limits of nutrients in infant formulas* (November 7-8, 1988) 1779-1784.

Hoebel, B.G., Avena, N.M. & Rada, P. (2007). Accumbens dopamine-acetylcholine balance in approach and avoidance. *Current Opinion in Pharmacology 7* (2007) 617-627.

IBM 100 (2011). Icons of Progress; Deep Blue Overview. http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/. *IBM* (retrieved 16 May 2014).

Jackson, J. (16 February 2011). IBM Watson vanguishes human Jeopardy foes. http://www.pcworld.com/article/219893/ibm_watson_vanquishes_human_jeopardy_foes.html. *PCWorld* (retrieved 15 May 2014).

Jacobs, D.K, Nakanishi, N., Yuan, D., Camara, A., Nichols, S.A. & Hartenstein V. (2007). Evolution of sensory structures in basal metazoa. *Integrative and Comparative Biology 47* (5) 712-723.

Kalat, J.W. (2004). *Biological Psychology; 8th Edition* (Belmont, 2004).

Kaneshiro, N.K. (12 April 2013). Infant reflexes. http://www.nlm.nih.gov/medlineplus/ency/article/003292.htm. *Medline Plus* (retrieved 6 March 2014).

Laird, J.E. (2011). Artificial Intelligence & computer games research. http://ai.eecs.umich.edu/people/laird/games_research.html. University of Michigan (retrieved 28 August 2014).

Laird, J.E. (2002). Research in human-level AI using computer games. *Communications of the ACM 45* (1) 32-35.

Laird, J.E., Derbinsky, N. & Tinkerhess, M. (2012). Online determination of value-function structure and action-value estimates for reinforcement learning in a cognitive architecture. Advances in Cognitive Systems 2. 221-238.

Laird, J.E. & Lent, M. van (2001). Human-level AI's killer application: Interactive computer games. *AI Magazine 22* (2) 15-26.

Laurenti-Lions, L., Gallego, J., Chambille, B., Vardon, G. & Jacquemin, C. (1985). Control of myoelectrical responses through reinforcement. *Journal of the Experimental Analysis of Behavior 44* (2) 185-193.

Lawrence, J.G. (2005). Horizontal and vertical gene transfer: The life history of pathogens. In: Russell, W. & Herald, H. (ed.). *Concepts in Bacterial Virulence; Contributions to Microbiology 12* (Basel, 2005) 255-271.

Lehman, J. F., Laird, J. & Rosenbloom, P. (2006). A gentle introduction to Soar, an architecture for human cognition: 2006 update. http://ai.eecs.umich.edu/soar/sitemaker/docs/misc/GentleIntroduction-2006.pdf. *University of Michigan* (retrieved 1 September 2014) 1-37.

Lewis, James L. III (October 2013). Acid-base regulation. http://www.merckmanuals.com/professional/endocrine_and_metabolic_disorders/acid-base_regulation_and_disorders/acid-base_regulation.html?qt=&sc=&alt. *The Merck manual* (retrieved 2 June 2014).

Lucantonio, F., Takahashi, Y.K., Hoffman, A.F., Chang, C.Y., Bali-Chaudhari, S., Shaham, Y., Lupica, C.R. & Schoenbaum, G. (2014). Orbitofrontal activation restores insight lost after cocaine use. *Nature Neuroscience 17* (8) 1092-1099.

MacDonald, G. & Leary, M.R. (2005). Why does social exclusion hurt? The relationship between social and physical pain. *Psychological Bulletin 131* (2) 202-223.

Magnab, R.M. & Koshland, D.E. Jr. (1972). The gradient-sensing mechanism in bacterial chemotaxis (temporal gradient apparatus/stopped-flow/S. Typhimurium/motility tracks/memory). *Proceedings of the National Academy of Sciences of the United States of America 69* (9) 2509-2512.

McCarthy, J. (2007). From here to human-level AI. *Artificial Intelligence 171* (18) 1174-1182.

Mee, S., Bunney, B.G., Reist, C., Potkin, S.G. & Bunney, W.E. (2006). Psychological pain: a review of evidence. *Journal of Psychiatric Research 40* (8) 680-690.

Mikolov, T., Sutskever I. & Quoc, L. (15 August 2013). *Learning the meaning behind words.* http://google-opensource.blogspot.nl/2013/08/learning-meaning-behind-words.html. *Google Knowledge* (retrieved 24 March 2014).

Miller, M.B., Bassler, B.L. (2001). Quorum sensing in bacteria. *Annual Review of Microbiology 55*. 165-199.

Mitchell, T.M. (1997). *Machine Learning* (Boston, 1997) xv.

Mohan, S. & Laird, J.E. (2011). An object-oriented approach to reinforcement learning in an action game. *Proceedings of the Seventh AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*. 164-169.

Mohan, S., Mininger, A.H., Kirk, J.R. & Laird, J.E. (2012). Acquiring grounded representations of words with situated interactive instruction. *Advances in Cognitive Systems 2*. 113-130.

Moscarello, J.M. & LeDoux, J.E. (2013). The contribution of the amygdala to aversive and appetitive Pavlovian processes. *Emotion Review 5* (3) 248-253.

Murray, E., Wise, S. & Rhodes, S. (2011). Chapter 4: What can different brains do with reward? In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

Nason, S. & Laird, J.E. (2004). Integrating reinforcement learning with Soar. *International Conference on Cognitive Modeling 2004*. 208-213.

Nilsson, N.J. (2005). Human-level Artificial Intelligence? Be serious! *AI Magazine 26* (4) 68-75.

Nilsson, N.J. (2007). The Physical Symbol System Hypothesis: Status and prospects. *50 Years of artificial intelligence* (Berlin, 2007) 9-17.

Nisbett, R.E. & DeCamp Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review 84* (3) 231-259.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions* (New York, 1998).

Park, S., Wolanin, P.M., Yuzbashyan, E.A., Silberzan P., Stock, J.B. & Austin R.H. (2003). Motion to form a quorum. *Science 301* (5630) 188.

Quartz, S.R. (2009). Reason, emotion and decision-making: risk and reward computation with feeling. *Trends in Cognitive Sciences 13* (5) 209-215.

Rilling, J.K. (2013). The neural and hormonal bases of human parental care. *Neuropsychologia 51* (4) 731-747.

Rilling, J.K. & Sanfey, A.G. (2011). The neuroscience of social decision-making. *Annual Review of*

Psychology 62. 23-48.

Robinson, H. (2012). Dualism. http://plato.stanford.edu/archives/win2012/entries/dualism. In: Zalta, E.N. (ed.). *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition) (retrieved 24 March 2014).

Rombouts, J.O., Roelfsema, P.R. & Bohte, S.M. (2012). Neurally plausible reinforcement learning of working memory tasks. *Advances in Neural Information Processing Systems 25* (NIPS 2012) 1871-1879.

Russell, S. & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach; Third Edition* (New Jersey, 2010).

Rvachev, M.M. (2013). Neuron as a reward-modulated combinatorial switch and a model of learning behavior. *Neural Networks 46* (2013) 62-74.

Saunders, B.T., Yager, L.M. & Robinson T.E. (2013). Cue-evoked cocaine “craving”: role of dopamine in the accumbens core. *The Journal of Neuroscience 33* (35) 13989-14000.

Schoenbaum, G., Roesch, M.R., Stalnaker, T.A. & Takahashi, Y.K. (2011). Chapter 15: Orbitofrontal Cortex and Outcome Expectancies: Optimizing Behavior and Sensory Perception. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

Searle, J.R. (2002). *Consciousness and Language* (Cambridge, 2002).

Searle, J.R. (1990). Is the brain's mind a computer program? *Scientific American 262* (1) 26-31.

Searle, J. (2004). *Mind: A brief introduction* (Oxford, 2004).

Searle, J.R. (1980). Minds, brains and programs. *Behavioral and Brain Sciences 3* (3) 417-24.

Searle, J.R. (1997). *The Mystery of Consciousness* (New York, 1997).

Tang, N.K.Y. & Crane, C. (2006). Suicidality in chronic pain: a review of the prevalence, risk factors and psychological links. *Psychological Medicine 36* (5) 575-586.

Thioux, M., Keysers, C. (2010). Empathy: shared circuits and their dysfunctions. *Dialogues in Clinical Neuroscience 12* (4) 546-552.

Turing, A.M. (1950). Computing Machinery and Intelligence. *Mind 59* (236) 433-460.

University of Reading (8 June 2014). Turing test success marks milestone in computing history. http://www.reading.ac.uk/news-and-events/releases/PR583836.aspx. *Press Releases University of Reading* (retrieved 24 June 2014).

Upbin, B. (8 February 2013). IBM's Watson Gets Its First Piece Of Business In Healthcare. http://www.forbes.com/sites/bruceupbin/2013/02/08/ibms-watson-gets-its-first-piece-of-business-in-healthcare/. *Forbes* (retrieved 27 August 2014).

Vargas, B. & Morales, E.F. (2009). Learning navigation Teleo-Reactive Programs using behavioural cloning. *IEEE International Conference on Mechatronics* (ICM 2009) 6pp.

Wang, Y. & Laird, J.E. (2010). Efficient value function approximation with unsupervised hierarchical categorization for a reinforcement learning agent. *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on Web Intelligence (WI-10) and Intelligent Agent Technology (IAT-10)* (2) 197-204.

Way, B.M., Taylor, S.E. & Eisenberger N.I. (2009). Variation in the mu-opioid receptor gene (OPRM1) is associated with dispositional and neural sensitivity to social rejection. *Proceedings of the National Academy of Sciences of the United States of America 106* (35) 15079-15084.

Weinberger, N.M. & Bieszczad K.M. (2011). Chapter 1: Introduction: From traditional fixed cortical sensationism to contemporary plasticity of primary sensory cortical representations. In: Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

White, N.M. (2011). Chapter 3: Reward: What is it? How can it be inferred from behaviour? In:

Gottfried, J.A. (ed.). *Neurobiology of Sensation and Reward* (Boca Raton, 2011).

Williams, J. (2007). Thinking as natural: Another look at human exemptionalism. *Society for human ecology 12* (2) 130-139.

Illustrations

Cover illustration:
“Punishment and Reward”
Custom image provided by Izak Perdijk
http://pinkoliphant.deviantart.com/

All remaining illustrations have been created by the author.

List of Figures and Tables