Compressionism: A Theory of Mind Based on Data Compression

Phil Maguire (pmaguire@cs.nuim.ie)
Oisín Mulhall (oisinmulhall@imagine.ie)
Department of Computer Science
National University of Ireland, Maynooth

Rebecca Maguire (rebecca.maguire@ncirl.ie)
School of Business, National College of Ireland
IFSC, Dublin 1, Ireland

Jessica Taylor (jessica.liu.taylor@gmail.com)
Department of Computer Science,
Stanford University, California

Abstract

The theory of functionalism holds that mental states are constituted solely by the functional organisation of an implementing system. We propose that the specific mechanism supporting both intelligence and consciousness is data compression. Recent approaches to cognition and artificial intelligence, based on a branch of theoretical computer science known as algorithmic information theory (AIT), have proposed a computational view of induction, prediction, understanding and consciousness which is founded on this concept. Building on existing literature we propose the term 'compressionism' as a label to unite theories which recognise intelligent, cognisant systems as sophisticated data compressors. We show how data compression can shed light on information binding and offer a novel perspective on the hard problem of consciousness.

Keywords: Consciousness, data compression, artificial intelligence, integrated information, combination problem.

Introduction

Throughout history people have speculated whether the human mind might be viewed as a sophisticated automaton. The theory of functionalism holds that mental states are constituted solely by their functional role (e.g. Putnam, 1960). From this perspective it doesn't matter how a given system is implemented functionally, whether it uses electronic circuits or mechanical gears. What matters when attributing intelligence and cognisance is solely the computational processes being implemented, a stance which opens the door for consciousness to be realised in many different forms.

While functionalism tells us that intelligence and consciousness may be implemented in a number of ways, it does not bring us any closer to understanding the specific form of information processing that gives rise to these phenomena. Here, we explore the idea that the key function of the brain which supports both intelligence and consciousness is data compression. In the following sections we elucidate the nature of data compression and bring together existing theories which has highlighted its connection with intelligent thought and subjective experience.

Data Compression

At first blush the concept 'data compression' might seem esoteric, a niche idea related to information storage in computer science. However, the concept runs much deeper. Data compression occurs when information is bound together through the identification of shared patterns. For example the sequence 4, 6, 8, 12, 14, 18, 20, 24... can be simplified as the description "odd prime numbers +1". The latter representation is shorter, hence we can say that it has been 'compressed'. The elegance and concision of this representation suggests that it is the true underlying pattern which governs the sequence. Accordingly, someone who manages to identify this pattern might claim to have understood the sequence, because they appreciate its relationship with the causal mechanism that produced it.

The above example highlights the close link between data compression and prediction. As per Occam's razor, concise models make fewer assumptions and are thus more likely to be correct. Levin's (1974) Coding Theorem demonstrates that, with high probability, the most likely model that explains a set of observations is the most compressed one. In addition, for any predictable sequence of data, the optimal prediction of the next item converges quickly with the prediction made by the model which has the simplest description (Solomonoff, 1964). All successful predictive systems, including machines and humans, are approximations of an ideal data compressor.

Because of its connection to prediction, data compression can be viewed as providing reliable proof of understanding. According to Chaitin (2006), "A useful theory is a compression of the data; compression is comprehension" (p. 77). The more compression is achieved, the greater the extent to which a system can be said to understand a set of data. Based on this principle, an enhanced version of the Turing test for machine intelligence has been established in which the challenge is to compress, to the greatest extent possible, 100 megabytes of textual information drawn from Wikipedia (see Hutter prize; Legg & Hutter, 2007). While people are very good at identifying patterns that link words and sentences together, computer programs struggle to identify the underlying meaning of the text and hence require longer descriptions to encode it.

Shannon (1951) carried out experiments with human participants to establish the entropy of English. He presented masked text and asked participants to guess the subsequent character, finding that, on average, each one required roughly two guesses (an entropy of 1 bit per character). For a computer program to match this rate of compression it would have to make predictions as accurately as humans, meaning it would need to identify the same deep patterns embedded in the text, which relate to real world structures and concepts. This form of deep understanding is what we refer to as appreciating the 'meaning' of the text.

Take a simple drawing as another example. Without background knowledge and prior embodied experience, a computer program cannot associate the drawing with previously experienced objects. To encode the image it must individually register every pixel on the screen, resulting in a large data file, even when compression algorithms are applied. However, a human can recognise an even more complex pattern: for example, a drawing of a house. The pixels are not coloured randomly. Instead there is a clear pattern that explains their ordering and relations, e.g. the presence of a roof, windows and a door. These regularities can be summarised concisely without having to spell them out pixel by pixel.

The greater the extent of the identified patterns, the shorter the resulting representation and the better predictions it supports. Imagine that pixel (254, 56) is corrupted by noise. The uncompressed representation that stores each pixel individually cannot correct this error. However, if we can identify patterns in the data, we understand how it is connected. By appreciating this redundancy we can repair and filter errors. For instance, we realise that pixel (254, 56) is part of the roof of the house and thus should be coloured in the same way.

A weakness of the Turing test (Turing, 1950) is that a program might pass the test simply by exploiting weaknesses in human psychology. If a given system passes the test we cannot be sure if it was because of the quality of the responses or the gullibility of the judge. In contrast, Hutter's compression test is more reliable. The more that data is compressed, the harder it becomes to compress it further (Chaitin, 2006). Because there is no way to cheat by using a simple heuristic, data compression presents a reliably hard standard. We argue that this process of identifying deep patterns through compression is what people mean when they attribute both 'intelligence' and 'consciousness'.

Compression and Intelligence

Research in the area of artificial intelligence and cognitive science is increasingly identifying data compression as a key organisational principle. Wolff (1993) originally identified a link between computing, cognition and information compression. He developed the idea that the storage and processing of information in computers and brains, from the recognition of objects to the use of natural language, can be understood in terms of information compression. Chater and Vitányi (2003) have proposed data compression as a unifying principle in cognitive science. They point out that much of perception, learning and high-level cognition involves finding 'sensible' patterns in data. It is the simplicity, or concision, of these patterns which supports their predictive power.

Schmidhuber (2009) proposes data compression as the simple principle which explains essential aspects of subjective beauty, novelty, surprise, interestingness, attention, curiosity, creativity, art, science, music and jokes. He argues that data becomes temporarily interesting once an observer learns to predict (i.e. compress) it in a better way, making it subjectively simpler and more 'beautiful'. From this perspective, curiosity can be viewed as the desire to create and discover patterns that allow for compression progress, with the level of interestingness being related to the effort required. According to Schmidhuber, this drive for compression motivates exploring infants, mathematicians, composers, artists, dancers and comedians, as well as artificial systems.

In a similar vein, both Maguire et al.'s (2013) theory of subjective information and Dessalles' (2011) simplicity theory view data compression as a key explanative construct in the phenomenon of surprise. When people experience a stimulus which is expected to be random, yet is found to be compressible, it triggers a surprise response. Observations of this form suggest the existence of an underlying pattern where none was anticipated, resulting in an urgent representational updating process. Accordingly, Maguire et al. (2013) suggest that people often rely on data compression rather than probability theory to judge likelihood and make decisions in many real world scenarios.

Adopting the perspective of the mind as a compressor, Gauvrit, Zenil and Tegnér (2015) connect AIT to experimental observations in the areas of working memory, probabilistic reasoning and linguistic structures. They argue that the concepts of data compression and algorithmic complexity provide an important normative tool which can shed light on a broad range of cognitive processes, from language use to the interpretation of EEG and fMRI data (e.g. Wang et al., 2014).

In the field of AI, Hutter (2005) views data compression as the key to an ideal mathematical definition of intelligence. He defines intelligence in terms of an agent's ability to achieve goals and succeed in a wide range of environments, a capacity which depends on being able to identify the simplest model which explains a set of data. Hutter's parameterless theory of universal artificial intelligence, known as AIXI, relies on the most compressed model to predict the future and guide its decisions.

The connection between compression and prediction can provide an elegant explanation for why the brain has evolved to carry out compression so effectively. In order to thrive in an uncertain environment, organisms must be able to anticipate future events. The more efficiently they can compress their experiences, the more accurate their predictions will be. As a result, organisms have evolved brains which are prodigious compressors of information: compressing sensory information provides them with an understanding of their environment. We propose the term 'compressionism' to group together those theories which view data compression as the key organisational principle underlying the structure of intelligent thought. In the following sections we make the case that, as well as providing a grounding for a formal definition of intelligence, compressionism also provides the foundations for a philosophy of mind which can shed light on the combination problem and offer insight into the nature of subjective experience.

The Combination Problem

Intuitively, our subjective experience seems to carry the qualitative characteristic of being united and singular, a property which is at odds with the reducibility of the physical world. How can the brain, whose processing is clearly distributed in space and time, give rise to a consistently integrated perspective? James (1890) vividly articulated this problem: "Take a sentence of a dozen words, and take twelve men and tell to each one word. Then stand the men in a row or jam them in a bunch, and let each think of his word as intently as he will; nowhere will there be a consciousness of the whole sentence" (p. 160).

Here James is highlighting the apparent incoherence of a collection of physical objects sharing a combined experience. We propose that combination is achieved, not by any local physical convergence or supernatural process, but through the process of data compression. When the brain compresses information, it is binding data together through the identification of shared patterns that might originally have been dispersed in space and time. The process of grouping and encoding these patterns yields computational results which reflect a 'coming together' of information. We will now explore the idea that this combination effect might be what underlies the united characteristic of conscious behaviour.

United Through Cooperation

The reproductive success of an organism is dependent on cooperation between all of its constituent components, leading to a form of compression which unites data distributed across space and time. For instance, it does not make sense for an organism's legs to maintain independent agenda. Because the interests of both legs are intimately bound, it is more productive for them to cooperate with each other in achieving a single set of objectives (e.g. walking). Accordingly, the brain sources sensory information from all over the body and binds it through compression, thereby optimising predictive accuracy for the organism as a whole. Tactile information from every limb is compressed alongside visual information from the eyes and audio information from the ears, giving rise to a form of understanding that is centralised and representative of the organism's experiences as a singular unit. The resulting decisions of the organism also appear centralised: to the external observer it seems as if the organism's body is being 'controlled' by a single entity.

The success of an organism also depends on cooperation through time. Accordingly, the response it exhibits to a sensory stimulus depends not just on its immediate processing, but also on its memories. Patterns in a current stimulus are matched against patterns distilled from historical stimuli, leading to a form of understanding that unites not only distributed sensory organs but also an organism's past and present states (see Maguire & Maguire, 2010).

The more memories we store, the greater our ability to identify and compress patterns in novel stimuli. For instance, when we observe an apple we can connect the sensory data with a large number of previously encountered apples, allowing us describe the stimulus concisely in terms of its relationship to an existing set of memories. Rather than needing to encode every detail of the fruit, we can simply note the ways in which it differs from a historical prototype, allowing greater redundancy to be identified, and thus enhancing predictive accuracy. In this way, data compression can support the identification of concepts.

Intuitively, features of the brain which are not linked to optimising behaviour should not have been rigorously preserved by evolution. The question thus arises of why the brain should go to the effort of producing consciousness at all. One advantage of the 'compressionist' perspective is that, rather than requiring an additional unique property of the brain to answer this question, it simply extends an approach which is already used to model intelligence. Given that it manages to link unitary processing to predictive and reproductive success, compressionism can provide an elegant explanation for why and how the brain has evolved to bind information.

Integrated Information Theory

The above account of binding through data compression bears close resemblance to Tononi's (2008) integrated information theory. In line with compressionism, Tononi proposes that consciousness is an information processing phenomenon and can thus be quantified in terms of a system's organisational structure, specifically its capacity to integrate information.

Tononi provides two thought experiments highlighting the need for conscious observations to a) produce information, and b) integrate information. In the first thought experiment he considers the difference between a human and a photodiode viewing a screen. The photodiode can only respond with two outputs, either light or dark. Accordingly, it generates a single bit of information about the stimulus, whereas a person, being able to distinguish between millions of different images, generates a much greater quantity of information. According to Tononi, the ability to discriminate between many different alternatives is an essential ingredient to conscious experience.

In his second thought experiment, Tononi (2008) establishes that information must be integrated to produce consciousness. He considers a digital camera whose sensor chip is a collection of a million binary photodiodes. Taken as a whole, the camera can distinguish among $2^{1,000,000}$ alternative states, corresponding to 1 million bits of information. With this level of precision, the camera is capable of responding differently to every frame from every movie that has ever been produced. However, because the information is not integrated, it's still not conscious. The photodiodes have no way to interact, and thus the camera is no different to a collection of a million independent photodiodes. If we chopped it up into a million pieces, each containing a photodiode, the function of the camera would not change at all.

Tononi (2008) seeks to quantify integration as the amount of information generated by a system as a whole above and beyond the information generated independently by its parts. For integrated information to be high, a system must be connected in such a way that information is generated by causal interactions among, rather than within, its parts, making it difficult to decompose into informationally disjoint parts. For example, people do not consider a digital camera to be conscious because, although it records information, each piece of data is processed independently and thus easily separated. It is straightforward to edit and delete one photo without affecting any of the other digital content stored on the camera.

Intuitively, people's conscious experiences cannot easily be broken down in this way. It would be far more difficult for a neurosurgeon to operate on somebody's brain and alter their memory of a stimulus. Rather than being encoded in a discrete, localised region of the brain, as per the camera, a person's sensory experience is integrated with a vast number of memories, causing widespread changes across all areas of the brain. Inside a digital camera the encoded information lies isolated, detached and dormant. In somebody's brain the same information alters every aspect of how they think, act and process subsequent information.

Tononi's (2008) idea of binding through information integration is closely aligned with the idea of binding through data compression. Indeed, Maguire et al. (2014) have shown that, for information lossless processes, Tononi's quantification of integration is equivalent to data compression. Consider, for example, the case of an uncompressed string of data. Here, every bit carries independent information about the string. In contrast, when the same file is compressed to the limit, each bit in the final representation is fully dependent on every other bit for its significance. No bit carries independent information about the original text file. Damaging the first bit of an uncompressed file leaves you with a 50% chance of getting the first bit right and 100% chance of getting the rest of the bits right. Each bit holds independent significance. For a compressed file, damaging the first bit corrupts everything and leaves you with only a 50% chance of getting all the bits right and a 50% chance of getting them all wrong. The significance of the first bit has been totally integrated with all of the other bits through the process of data compression.

If data is optimally compressed then it becomes extremely difficult to edit in its compressed state. For example, imagine a compressed encoding of a Wikipedia page. You want to edit the first word on the page. But where is this word encoded in the compressed file? There is no easily delineated set of bits which corresponds to the first word and nothing else. Instead, the whole set of data has been integrated, with every bit from the original file depending on all the others. To discern the impact that the first word has had on the compressed encoding you have to understand the compression scheme. There are no shortcuts. Accordingly, Maguire et al. (2014) have developed a formulation of a memory's 'edit distance' as a reliable measure of its integration.

We propose that this unavoidable difficulty of breaking down people's behaviour into a set of discrete, independent components is precisely what people mean when they apply the term 'conscious'. From this integration or 'compressionist' perspective consciousness is not a tangible property, but rather a heuristic that people adopt in modelling the behaviour of a system. Specifically, we attribute consciousness to a system when the data compression it carries out is so sophisticated that we are forced to model it as unitary. In the case of the brain, its extraordinary capacity for compressing information leads us to attribute consciousness to the behaviour produced.

Dennett's Multiple Drafts Model

While compressionist theories of consciousness emphasise integration as the defining feature of consciousness (e.g. Tononi, 2008), Dennett (1991) provides a theory which focuses specifically on the disintegrated nature of consciousness. Similar to the compressionist perspective, his multiple drafts model views the mind in terms of information processing, and consciousness as an explanative model for behaviour. However, instead of relying on information theory to resolve the combination problem, he proposes that there actually is no problem, because there is no strong combination. He criticises the idea of what he calls the "Cartesian theatre", a point where all of the information processing in the brain is integrated. Instead, Dennett (1991) presents consciousness as a succession of multiple drafts, a process in flux without central organisation.

Intuitively, consciousness does not appear 'drafty' (i.e. easily decomposable into disjoint, unintegrated components). While physical objects are clearly drafty, being reducible into discrete atoms, the standout feature of conscious systems is how undrafty they are. According to the compressionist perspective, it is the undraftiness of a system's information processing we have in mind when we describe it as being 'conscious'. Although Dennett (1991) provides examples where consciousness seems to be less strictly integrated (e.g. the phi phenomenon, the cutaneous rabbit illusion), what is striking about these examples is how few and far between they are. We simply don't notice disintegration effects in everyday life. For this reason, the burden of explanation on theories of consciousness should be to explain how such a level of extreme undraftiness (i.e. integration) is achieved by the brain. Dennett's (1991) efforts to illustrate that consciousness is somewhat drafty fail to tackle the burning question of how and why people's behaviour appears so consistently and convincingly unitary. The advantage of compressionism is that it can offer a computational account for how such strong information binding can occur, without needing to invoke any special non-physical properties.

In the following sections we demonstrate the potential of compressionism as a theory of mind by detailing how it can account for other mysterious features of consciousness, such as qualia and self-awareness.

Addressing the Hard Problem of Consciousness

It seems possible to conceive of an artificial compressor which compresses large amounts of current and historical data in parallel, though without experiencing the same form of awareness that humans are familiar with. For example, critics of compressionism might question whether off-the-shelf image or video compressors understand the images they are compressing. We argue that compression carried out by the brain is likely to have two additional ingredients which set it apart from simpler compression systems by supporting self-awareness: first, it is embodied, and second, it compresses observations of its own behaviour.

Compression enables the binding of information that is distributed in both time and *space*. The embodied approach in cognitive science emphasises the vital role an organism's body plays in determining how and what it thinks. According to Clark (1999): "It is increasingly clear that, in a wide variety of cases, the individual brain should not be the sole locus of cognitive scientific interest. Cognition is not a phenomenon that can be successfully studied while marginalizing the roles of body, world and action". A computer program that compresses files on a computer lacks the embodiment that would provide it with a situated perspective on the world. It has no means of engaging physically with its environment, hence its compression does not instantiate the singular embodied agent we associate with consciousness.

The second ingredient concerns the social aspect of consciousness. According to Dunbar and Shultz (2007), intelligence was selected for, not by the need for technical competence, but by the computational demands of living in large, complex societies. When we watch other individuals, we realise that there is a great deal of redundancy in their activity. Rather than simply cataloguing and memorising every action they perform, we can instead posit the more succinct hypothesis of a concise 'self' which motivates these actions. By maintaining this theory of selfhood we can compress the behaviour of others and thus make accurate predictions as to how they will behave in different contexts. But the behaviour of other humans also has another component, namely that they react to you, the observer. In order to best predict and manipulate the behaviour of others, it pays to maintain a model of one's own self, a process which is achieved by compressing our own observations of past behaviour (see Friston & Frith, 2015).

We propose that this 'understanding of the self' is a requirement for the accurate modelling of the behaviour of others. If an individual lived in complete isolation within a simple environment, there would be no motivation for maintaining a complex model of a self that has experiences. It is only when people are embedded in a complex social environment that the goal of interacting with others requires them to maintain a detailed model of their own actions. When you observe others, they observe you observing them observing you. This recursive modelling, achieved through data compression, gives rise to a rich understanding of selfhood, an understanding of what it feels like to be conscious.

This recursive self-modelling may also explain another aspect of consciousness that seems to evade logical explanation, namely its subjective flavour, or qualia. Some examples of qualia include the pain of a toothache, the taste of sweetness or the perceived redness of an apple. These kinds of experiences seem to defy objective, reducible description.

Dennett (1991) points out that even the most subjective of qualia are closely intertwined with the concept of self-modelling. He argues that an observer's perception of an experience is nothing other than an understanding of how the observer is *affected* by that experience: there is nothing left over that could be considered the essence of a quale. For example, your personal experience of the colour red is nothing greater than an understanding of the implications of encountering a red stimulus, a model which you maintain to assist in compressing the behaviour of others. According to this view, all subjective feelings are, at their root, based on the identification of patterns which connect and unite events distributed in space and time.

We propose that the unique flavour of consciousness results from individual differences in how stimuli are compressed. Conscious perception is not a passive process whereby incoming sensory information is simply recorded in its raw format. Instead, people gaze through the lens of understanding provided by data compression. Different people, with different memories, will extract different patterns in a given stimulus, leading to different subjective experiences. In particular, people will have an intimate personal understanding of how they themselves experience or 'compress' a stimulus, one which remains beyond the remit of objective science.

Philosophical Zombies

Chalmers (1995) has previously identified a distinction between what he views as the 'easy' problems of consciousness (e.g. explaining object discrimination) and the much harder problem of explaining why subjective feelings accompany cognitive information processing. He explores the idea of an 'explanatory gap' between the objective and subjective, arguing that physical explanations cannot account for mental experience. He maintains that mental states are ontologically distinct from and are not reducible to physical, computable systems.

In support of this position, Chalmers entertains the possibility of philosophical zombies, entities which act just like a conscious individual, but which lack qualitative experience. He argues that because such zombies seem conceivable, they must be logically possible, hence raising questions about the soundness of physicalism. Compressionism, on the other hand, takes the opposite view, seeking to bridge the explanatory gap. If consciousness is equivalent to data compression then any two systems which carry out the same compression should be viewed as equally conscious.

Chalmers' (1995) acceptance of philosophical zombies paves the way for a double-aspect theory of information, leading him to speculate on the possibility of alien 'psychophysical laws' that govern qualia, supporting widespread 'panprotopsychic' consciousness across the universe. Compressionism, in contrast, offers a more parsimonious account of the link between consciousness and the physical world, relying on computational complexity rather than supernatural phenomena. By explaining how straight-forward information processing can bind information and lead to subjective, self-aware experiences, it might seem that compressionism can dissolve the hard problem.

Nevertheless, proponents of compressionism should not be so quick to dismiss Chalmers' views. For a start, optimal data compression is not computable (see Chaitin, 2006). What this means in practice is that the problem of consciousness is guaranteed to remain hard: there is no computable scenario that would allow it to be resolved. Indeed, given this intrinsic intractability, Maguire et al. (2014) have shown that there is no potential of a breakthrough theory in neuroscience which would explain how the brain carries out its compression. Thus, what compressionism offers is not the solution to the puzzle of consciousness, but merely a framework which can serve to point us in the right direction. To be clear, identifying a structural parallel between information processing and phenomenal experience does not eliminate the mysteriousness of that experience. It simply allows us to express the problem in a more systematic fashion, potentially supporting further insight.

Conclusion

Converging sources of evidence suggest a fundamental explanatory role for data compression in the evolved function of the brain. We have explored some of this research and sought to connect it together by proposing the term 'compressionism'. The concept of data compression is already pervasive in the history of psychology and cognitive science. From neural networks, to classical ACT-R and SOAR architectures, to Bayesian predictive models, information is processed in ways that end up compressing events spread out over space and time. Thus, what compressionism offers is not a radically new idea but rather a framework for reconciling these varied approaches, one which opens up the use of powerful mathematical tools developed in AIT by Solomonoff, Chaitin, Hutter and others. Although it cannot and does not seek to eliminate the hard problem, compressionism can serve to put a name and structure to that hardness.

References

Chaitin, G. (2006). The limits of reason. *Scientific American*, *294*(3), 74–81.

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consc. Studies*, *2*(3), 200–219.

Chater, N., & Vitányi, P. (2003). Simplicity: a unifying principle in cognitive science? *Trends in Cognitive Sciences*, *7*(1), 19–22.

Clark, A. (1999). An embodied cognitive science? *Trends in cognitive sciences*, *3*(9), 345–351.

Dennett, D. C. (1991). *Consciousness explained*. Little, Brown.

Dessalles, J.-L. (2011). Coincidences and the encounter problem: A formal account. *arXiv preprint arXiv:1106.3932*.

Dunbar, R. I., & Shultz, S. (2007). Evolution in the social brain. *Science*, *317*(5843), 1344–1347.

Friston, K., & Frith, C. (2015). A duet for one. *Consciousness and cognition*.

Gauvrit, N., Zenil, H., & Tegnér, J. (2015). The information-theoretic and algorithmic approach to human, animal and artificial cognition. *arXiv preprint arXiv:1501.04242*.

Hutter, M. (2005). *Universal artificial intelligence: Sequential decisions based on algorithmic probability*. Springer.

James, W. (1890). *The principles of psychology*. MacMillan.

Legg, S., & Hutter, M. (2007). Tests of machine intelligence. In *50 years of ai* (pp. 232–242). Springer.

Levin, L. A. (1974). Laws of information conservation (non-growth) and aspects of the foundation of probability theory. *Problems Information Transmission*, *10*(3), 206–210.

Maguire, P., & Maguire, R. (2010). Consciousness is data compression. In *Proceedings of the thirty-second conference of the cognitive science society* (pp. 748–753).

Maguire, P., Moser, P., Maguire, R., & Griffith, V. (2014). Is consciousness computable? Quantifying integrated information using algorithmic information theory. In *Proceedings of the thirty-sixth annual conference of the cognitive science society.*

Maguire, P., Moser, P., Maguire, R., & Keane, M. (2013). A computational theory of subjective probability [Featuring a proof that the conjunction effect is not a fallacy]. In *Proceedings of the thirty-fifth annual conference of the cognitive science society* (pp. 960–965).

Putnam, H. (1960). *Minds and machines*.

Schmidhuber, J. (2009). Driven by compression progress: A simple principle explains essential aspects of subjective beauty, novelty, surprise, interestingness, attention, curiosity, creativity, art, science, music, jokes. In *Anticipatory behavior in adaptive learning systems* (pp. 48–76). Springer.

Shannon, C. E. (1951). Prediction and entropy of printed English. *Bell System Technical Journal*, *30*(1), 50–64.

Solomonoff, R. J. (1964). A formal theory of inductive inference. Part I. *Information and control*, *7*(1), 1–22.

Tononi, G. (2008). Consciousness as integrated information: a provisional manifesto. *The Biological Bulletin*, *215*(3), 216–242.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*(4), 293.

Wang, Z., Li, Y., Childress, A. R., & Detre, J. A. (2014). Brain entropy mapping using fMRI. *PloS one*, *9*(3), e89948.

Wolff, J. G. (1993). Computing, cognition and information compression. *AI Communications*, *6*(2), 107–127.