Natural Information

Master's thesis in Cognitive Artificial Intelligence

Faculty of Humanities
Utrecht University

Julian Gonggrijp

May 16, 2014

Supervised by
dr. Janneke van Lith and Prof. dr. Vincent van Oostrom

Contents

Chapter 1

Introduction

Information, both as a concept and as a substance, has become increasingly important in daily life as well as in academia. Literacy has improved dramatically over the past centuries and people are spending increasing amounts of time in education. At the same time, the means to create, store and distribute information have improved, with the introduction of the internet causing what may be called an information *revolution*. The efficiency of information exchange (communication) for all possible purposes has increased explosively.

The internet has given an unmistakable boost to the sciences. However, information has also gained importance as a *topic* in various studies. In social sciences such as psychology and linguistics the relation between form, meaning and thought is of course an inherently important topic, but the cognitive revolution in the 1950s has further increased the attention for information and in particular mental information *processing*.

Artificial Intelligence is arguably the climax of this development. As a field of study, it revolves around the idea that information processing is a goal in itself that may be independent of any particular biological architecture such as the human brain. As a technology, information processing is both the means and the purpose of its operation. This focus on information processing as an intrinsic value evokes the question, what principles of nature provide the necessary and sufficient conditions for information processing.

Meanwhile, results from the natural sciences, for example thermodynamics, quantum dynamics and genetics, have promoted the idea that information may occur in nature without any form of human or otherwise "mental" or "conscious" intervention. Given a solid set of informational concepts, this would offer great potential for answering the question about the natural conditions for information processing. Unfortunately, until now there has been no comprehensive definition of information that could capture all relevant aspects at the same time.

Claude Shannon first published his *Mathematical Theory of Communication* in 1948. [1] Shannon's model has the important virtues of being *quantitative* and of being widely accepted. For this reason it has become highly influential in the mathematical field of information theory. Consequently, it found practical applications in modern communication techniques as well as theoretical applications in the natural sciences. However, it does not answer any of the more profound questions about information, such as what constitutes "meaning", or what are the necessary and sufficient conditions for one thing to represent another.

There exist several philosophical studies and definitions of information which do answer such questions. An extreme example is *semiotics*, a field wholly concerned with signs and meaning. While semiotics and philosophy offer handles to discuss the profound aspects of information, such handles have drawbacks of their own. In the first place, they tend to be purely qualitative and to rely on other concepts that are hard to define by themselves, such as "mind". This limits their use in the sciences. In the second place, they tend to be controversial. There is no single semiotic theory that is near-universally accepted as an approach to representation and meaning, unlike Shannon's theory which is near-universally accepted as a quantitative approach to communication.

The ambitious aim of the present study is to provide a sound definition of information, that can capture both its quantitative aspects and its profound, qualitative aspects even in a naturalistic setting. Key to our approach is the realisation that Shannon's model does have tacit implications for the qualitative aspects of information. The concepts of *signal*, *signal space* and *transducer* are of central importance.

We start by postulating a coherent set of information-related definitions that build directly on Shannon's mathematical framework. We then analyse the properties of our own framework and provide theoretical arguments why it should meet our goal to provide a sound, profound and naturalistic analysis of information. We demonstrate that our framework can firmly account for representation as well as computation.

We then continue to test the implications of our new framework on three current discussions in the natural sciences: the informational connotation of entropy in thermodynamics and statistical mechanics, the suggestion that information may be the fundamental substance in quantum physics, and the debate in biology on the extent to which genes may be considered carriers of information. In each case, we show that our framework is able to describe all aspects of the discussion and that strict application of the framework will lead to a single definitive answer. The results are coherent across the different scientific disciplines.

Concluding, the transducer- and signal-based framework truly appears to hold the promise of tackling both quantitative and qualitative aspects of information and information processing at the same time. Finally, we discuss some consequences and possible further applications of our framework for Artificial Intelligence and other fields.

Chapter 2

A framework

2.1 Shannon's perspective on information

Claude Shannon's seminal work, *A Mathematical Theory of Communication* [1], still makes for a very inspiring read more than sixty years after its first publication. Although Shannon seems to be specifically interested in the mathematical foundations of *human* communication, the presented framework is general enough to be relevant to *any* form of information exchange. Much of the definitions in the next section will rely heavily on this work.

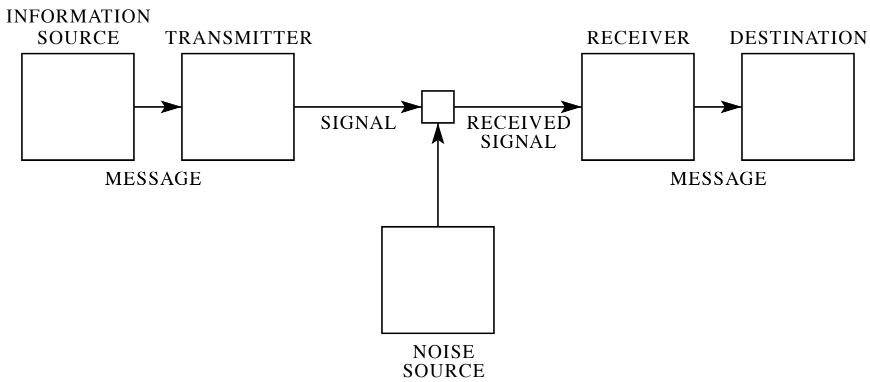

Figure 2.1: Shannon's outline of a communication system from 1948.

After some general considerations on the problem of communication and the best way to measure the amount of information in a message (a logarithmic function of the number of possible messages), Shannon starts with a general outline of the components that may be found in a typical system of human communication, here reproduced in Figure 2.1. The same outline can however also be reasonably applied to any other communication system. As we will see later on, the components of this system can be collapsed into a single class to achieve even higher generality. A short summary of the components is repeated here:

- the *information source* produces a *message*, which may in short be (nearly?) any discrete or continuous function of time;
- the *transmitter* encodes the message into a *signal* suitable for the channel;
- the *channel* is whatever medium is used to transmit the signal, e.g. a wire cable;
- the *noise source* may operate on the channel to introduce noise into the signal;

- the *receiver* decodes the signal from the other end of the channel into a new message, typically by performing the reverse operation of the transmitter;
- the *destination* is the person or device that consumes the final message from the receiver.

Note that in this outline, "messages" and "signals" belong to the same general class, i.e. functions of time. The channel and the signal are only inserted in the middle because the original message often isn't suitable for direct transmission over a large distance. For example, (amplified) speech through air carries no further than a few kilometers in the most favourable conditions, while an amplified modulated electric current over a phone cable can easily span thousands of kilometers.

Using this general outline, Shannon first provides a rigorous mathematical analysis of discrete noiseless communication systems. In such systems, messages and signals are sequences of *symbols* chosen from a finite set. Each of these symbols is assumed to have a certain duration in time. It is then shown how the same analysis can be extended to systems with noise, continuous systems and mixed systems. One key ingredient in the generalization to continuous and mixed systems is that continuous functions can be represented to arbitrary level of detail by discrete functions, a result which is nowadays best known as the sampling theorem [2]. From here on we'll restrict ourselves to the discrete systems, because these provide all abstractions we need.

Two important measures are defined, the capacity C of the channel and the entropy H of the information source. C is the amount of information (for example measured in bits) that can be transmitted per unit of time. H is the average amount of information produced per symbol. An alternative measure of entropy H' is also provided, which represents the average amount of information produced per unit of time. H and H' can in principle be used in similar ways (keeping in mind that they have different dimensions and probably different values); H' is easier to compare with C. Shannon finally uses these units (and others) to derive his fundamental theorem for a noiseless channel, which basically states that a message from the information source can always be compressed enough to transfer the produced information at a rate arbitrarily close to the rate permitted by the channel[1].

However, it's not the fundamental theorem that we're most interested in here. Rather, it's one of the steps in its derivation: the analysis of what happens when a message is translated to a signal or vice versa. The transmitter and the receiver are two sides of the same coin in this regard, and Shannon unites them under the term *transducer*. A discrete transducer takes a sequence of symbols as its input, and its output is also a sequence of symbols. In addition it may have internal memory so that its output may depend on the

[1]Of course, when $H' < C$ an initial delay will be required to make this possible.

past history of input rather than just the last input symbol. Concluding, the transducer can be described with two functions:

$$y_n = f(x_n, \alpha_n)$$
$$\alpha_{n+1} = g(x_n, \alpha_n)$$

where x_n is the n^{th} input symbol, α_n is the state of the transducer at the time of the n^{th} input symbol, and y_n is the n^{th} sequence of output symbols. Note that y_n need not have the same size for every n and may be empty; the number of symbols per unit of time, or even the type of symbols, need not be the same at both sides of the transducer.

Shannon points out that transducers are composable. If the output symbols of one transducer are the input symbols to another, the combination of those two transducers together is also a transducer. If such a composed transducer exists, such that the second transducer recovers the original input to the first, the first is called *non-singular* and the second is its *inverse*. Also, the combination of an information source with a transducer is again an information source. Now, there are three exciting properties of transducers that together allow for an even more general framework.

1. All other components of the communication system are transducers as well.

Information source A transducer of which f and g either ignore x_n, or where x_n is drawn from a set with only a single element. In both cases a transducer is reduced to a *generator*, which is exactly how we understand the information source.

Destination A transducer in which y_n is the same sequence for every n, probably empty. Alternatively, a transducer where y_n and f are completely absent. Both possibilities reduce a transducer to a *consumer*, which is again how we understand the destination.

Channel A transducer where x_n and y_n are drawn from the same set. Note that in this case y_n must always be a sequence of size 1. If the channel is faithful it has $y_n = f(x_n, \alpha_n) = x_n$. In a nonfaithful channel, the difference between y_n and x_n may be due to nontrivial effects of α_n and g, in which case the difference is systematic, or it may be induced by an external source, in which case the difference is random noise.[2]

Noise source A generator, just like the information source.

[2] Strictly speaking, in order to obtain noise from an external source the channel needs to take composite symbols as input, consisting of one symbol from the transmitter and any number of symbols from the noise source. Hence y_n is drawn from the same set as the component of x_n that came from the transmitter, rather than from the set of which x_n as a whole is drawn. This does not pose a problem for the formalism, nor does it bear much relevance to the rest of the discussion.

To emphasize this, Figure 2.2 repeats Shannon's outline with all transducers explicitly marked as boxes with the same appearance. Each transducer box has an input end and an output end, with the triangle pointing to the output end, and the lines that connect the transducers are output–input connections. We will use this notation again later on. Thus Shannon's communication system outline turns out to be a special case of a general class of directed graphs, where the nodes are transducers and the edges are output–input connections (as output–input connections behave exactly like faithful channels we will use these concepts interchangeably from here on). Any such graph would be a valid communication system. Indeed, in the real world we usually expect the destination to be a full (though perhaps singular) transducer rather than a consumer; cases where a flow of information terminates completely at the intended destination are extremely rare.

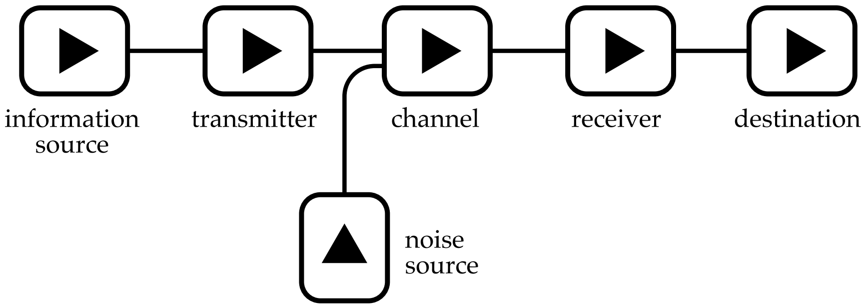

Figure 2.2: New version of Figure 2.1 where the transducers are made explicit.

2. Transducers model our understanding of interacting entities in physics. That is, the state of an entity and its interactions depend on its previous state and its previous interactions. Similarly, the state of a transducer and its interactions with the output channel depend on its previous state and its interactions with the input channel.[3] Together with property 1, this allows us to conclude that any interacting physical entity can be part of a communication system.

3. Transducers are modified Turing machines, with one restriction and one extension compared to the original Turing machine. The extension is that a transducer operates on two tapes, one for the input and one for the output.

[3]The apparent difference is that in physical entities we usually understand interactions to be pairs of opposite mutual influences (such as forces), while in the mathematical description of a transducer the influences seem to propagate only in one way: from the input channel through the state of the transducer to the output channel. This difference is only superficial. In each interaction we are simply abstracting away from one of the opposite influences, without losing information, knowing that each influence has a matching opposite. For example, the interaction in which the input channel influences the transducer at the same time also constitutes an opposite influence from the transducer on the input channel: a symbol is removed from it.

The restriction is that the read and write heads can only move in one direction. Together with properties 1 and 2 this seems to invite for a more formal way of analysing natural processes in terms of computational structure. In Section 2.4 we will see that transducers with feedback loops are Turing-complete.

2.2 Proposal

Building on our generalisations from Shannon's perspective on information, the following definitions are proposed. In the remainder of this work we will use capitalization to indicate that a word adheres to the definition provided in this section.

A **Signal** is any discrete, continuous or mixed function of time.

A **Symbol** is the value of a Signal at a given time slice. An alternative formulation is that a Signal is a sequence of states of some medium; each of those states is a Symbol. Note that "state" and "function" are here taken to mean abstracta.

A **Transducer** is a deterministic finite automaton that converts one Signal into another, characterised by the transition functions f and g as introduced in the previous section. It may have memory, so that its output depends not only on the last input Symbol but on the past history of input. The output of a Transducer may be routed to the input of another Transducer (shortly called an output–input connection) and Transducers are closed under composition by this operation. Note that a sequence of internal states of a Transducer $\alpha = (\alpha_n, \alpha_{n+1}, \ldots, \alpha_{n+x})$ is also a Signal.

Several subclasses of Transducers can be identified. A **Generator** has nonconstant output (constant output would be equivalent to no output). When a Transducer depends on nonconstant input it is a **Consumer**; i.e. it may not be the case that for all α_n, both y_n and α_{n+1} are the same for all possible x_n. A Transducer that is both a Generator and a Consumer is a **Translator**.[4] A **Trivial Translator** has $y_n = x_n$. Output–input connections happen to behave exactly in that way and are therefore also Trivial Translators.

Since unfaithful channels are irrelevant to the topics of this work, we will use **Channel** as a reference to a directed graph $\langle S, R\rangle$ where S and R are sets of respectively sending and receiving Transducers. See Figure 2.3 for a visualisation and the Appendix for a precise definition. While this definition assumes many-to-many connections, it is general enough to also include singular output–input connections. A Channel is a feedback loop if $S = R$;[5] in that case the Channel graph collapses into an unpartitioned, strongly connected complete directed graph.

For the sake of simplicity, we assume the following temporal properties of Channels and Transducers. All interconnected Transducers are synchronized with a global "clock" and take their nth input at the same time. Symbols are enqueued when the production rate is greater than the consumption rate, such that they arrive in the same order as they were inserted into a Channel.

[4] Transducers that are neither Generator nor Consumer will be uninteresting for any outside observer. If you require a name, however, I suggest "universe" or "isolated world".

[5] A Channel is only *not* a feedback loop if $S \cap R = \varnothing$. I will not discuss the grey area in between, however interesting that may be.

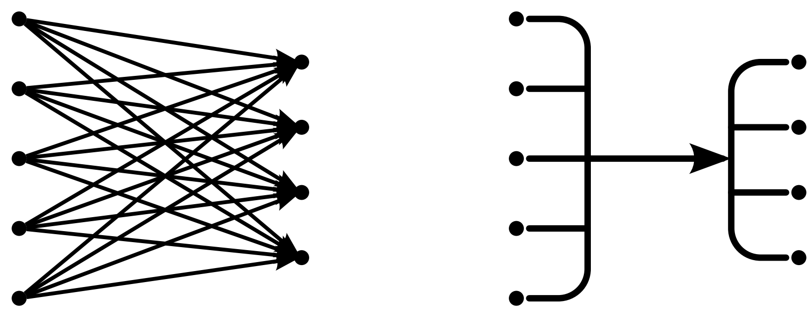

Figure 2.3: Different presentations of the same many-to-many Channel. Left: individual output–input connections between senders on the left and receivers on the right are shown. Right: the Channel as a single entity with multiple inputs and multiple outputs. These presentations are equivalent.

Channels do not fill up or block. If multiple Transducers output to the same Channel, the resulting input to Transducers on the receiving end consists of tuples with one Symbol from each source. Lacking output from one or multiple sources is considered a null Symbol. These assumptions will stay implicit for the remainder of this work. We believe that our proposed definitions could also be made to work under different assumptions.

The output Signal of a Translator is a **Representation** of the input Signal. This definition is chosen because it agrees with the common intuition that, for example, an electrical signal in a microphone wire is a representation of the sound that the microphone is registering, but not the other way round: the sound is not a representation of the electrical signal. If we reverse the process, i.e. if we let a speaker translate an electrical signal into sound, the distinction becomes a bit more cloudy; however, we still agree that the sound from the speaker is a representation of the sound that the microphone recorded and not the other way round. It therefore seems natural to require that the antecedent of a Representation should be causally prior to the Representation itself.[6]

There is a twist regarding Representations, which is that we usually also take "materialised signals" such as a photograph to be a representation. This observation calls for an analysis of such materialised signals. The materialised signal can be described entirely as a Translator; rather than directly producing output, it stores the to-be-materialised input Signal in its altered state. In real world objects such as photographs this altered state manifests as the altered physical structure of the object. This state can then later be retrieved on the output Signal. Hence, strictly speaking a materialised signal is a Translator, not a Signal, but for all practical purposes it is safe to pretend that it is a Signal.

[6]By which we mean singular causality. I.e., the particular input Signal of the Translator that the particular output Signal is a Representation of should be causally prior to that particular Representation itself.

A Signal is associated with a set from which its Symbols are drawn. This set is determined by the Transducer that produces the Signal. The Signal itself is drawn from the space of all sequences of Symbols from the set. The set is finite if and only if the Signal is discrete. For convenience we will henceforth assume discrete Signals and Symbols, but all conclusions derived with that assumption can be extended to continuous and mixed Signals and Symbols as well.

While the Transducer defines the total space that a Signal is drawn from, each particular Signal defines a subspace containing the Signal itself and all of its possible continuations. Shannon defined *the amount of information* carried by a Signal as the logarithm of the factor by which the Signal-defined subspace is a reduction of the Transducer-defined total space; indeed, this is exactly entropy times Signal length. We will adhere to this definition.

Shannon did not, however, explicate what *information itself* is. We would like to have such a definition because information is usually understood not only to come in various amount (quantity), but also in various content (quality). Fortunately the definition of the amount of information allows us to derive the definition of information itself: the **Information** inherent to a Signal is the reduction of the Transducer-defined total space to the Signal-defined subspace. In less formal terms we can make this result very intuitive: as the amount of Information is the amount of reduction of a Signal space, Information itself is the reduction itself. We can now also identify the qualitative aspect of Information: it is the direction of the reduction.[7] This interpretation of information is visualised in Figure 2.4.

Shannon also defined the amount of **Mutual Information**, which symmetrically relates two Signals such that each Signal reduces the Transducer-defined space of the other. Apart from the fact that Mutual Information involves two Signals rather than one, it's completely analogous to Information proper in the sense that it's reduction to a subspace with corresponding direction and magnitude. As a Transducer principally involves three related Signals (input, output and state), Mutual Information naturally lends itself to describe the degree of interdependence between these Signals, but it may be applied to any pair of Signals.

Often we are not interested in the Information that is inherent to a Signal, but only in the part that is "useful", "accessible" or *relevant* for the receiver. A rigorous way to decide to what extent a Signal is affecting a Transducer, is to assess to what extent it is influencing its state. **Relevant Information** is the Mutual Information between a Transducer's input Signal and its state Signal.

While Transducers are closed under composition we may sometimes want to emphasize that multiple Transducers are interconnected. Because of that

[7] This might remind you of a vector. A vector has a direction and a magnitude. Together with an origin, these properties together define the location of a point. Information also has a direction and a magnitude. Together with a Signal space, these properties define a subspace.

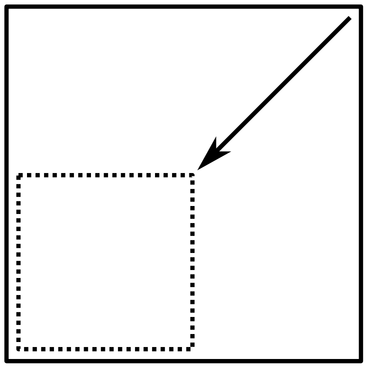

Figure 2.4: Information as reduction of a Signal space.
The outer box represents the total space of possible Signals for a given Symbol set. The inner box is the remaining part of the space given a Signal prefix. The arrow, which encodes both the scale and the direction of the reduction from the total space to the remaining space, is the Information in the prefix.

Communication Network will be used as a synonym for *composite Transducer*. The act of Transduction by a composite Transducer may also be called Communication.

Communication is Turing-complete, as will be shown in Section 2.4. We may also speak of Computation.

2.3 On Representation

We have seen that Shannon was only concerned with the quantitative aspects of the transmission of Information over a Channel. Intuitively this seems to neglect two important qualitative aspects of Information: its *content* and its *aboutness*, i.e. the fact that Information is always "about" something, the antecedent. Closely related is the concept of representation, where the antecedent seems to take a similar role as the content of Information. Since we based our definition of Representation on Shannon's quantitative framework, an exploration of these concepts is in order.

In the following discussion we will build on Cummins and Poirier (2004) [4] to argue that our definition of Representation is indeed properly named and would not be more aptly described as *indication*. In fact we will argue that a Representation is always composed of indications, i.e. the Symbols. After that we show that there are actually three modes of "aboutness", which explains how *misrepresentations* can exist. We will show how this avoids the problems that haunt Dretske and Millikan when they attempt to define representation in naturalistic terms. The three modes of aboutness also explain how Cummins and Poirier are right that representation is in some ways intransitive while our framework rightly predicts that it is transitive. Finally we briefly consider the aboutness of sentences, arguing that not all parts of a Representation may need to structurally correlate with the antecedent.

2.3.1 Representation versus indication

According to Cummins and Poirier, a representation refers to its antecedent by structural correlation while an indicator is arbitrary and refers to its antecedent purely by virtue of the context. In addition, Cummins and Poirier claim that indication is transitive while representation is not—more on that further below. It is important to note that all of their examples of indication refer to properties while all of their examples of representation appear to refer to entities. In order to clarify these distinctions it is imperative that we first determine the ontological category of each concept.

Indicator The examples provided in [4] include (a) the shape of the bimetallic element in a thermostat, (b) the firing intensity of an edge detector cell in the primary visual cortex and (c) the burning (or not burning) state of an "idiot light" of the kind that you find on the dashboard of a modern car. All of these examples are simple properties, or equivalently *states of entities*, corresponding exactly to our definition of Symbols.

Indicator antecedent The aforementioned examples were taken to indicate, respectively (a) the ambient temperature, (b) the (degree of) presence of an edge under a certain angle within a certain area of the visual field and (c) engine state parameters such as fuel level, oil pressure or coolant temperature. Again these are all simple properties, hence Symbols. In conclusion, indication is a relation where the state of one entity encodes the state of another, potentially very different entity. In all these examples it seems justified to say that the relation between the indication and what is indicated is the result of a causal process.

Representation No explicit list of examples is provided as for indication, but the following examples can be found throughout the chapter: (a) a digital photograph (p. 23, 33), (b) a written sentence (p. 36), (c) a map and (d) a scale model (p. 33). These are all "materialised Representations" as discussed when we introduced the definition of a Representation. Indeed, each of these examples "produces" a Signal when it is "read"; we could take either the object itself or the Signal that it produces to be the *actual* representation as per Cummins and Poirier. The most consistent interpretation is to equate the representation to the Signal, both for coherence with our own framework and for coherence with the notion that a representation is composed of indications—see below.

Representation antecedent Either explicitly or implicitly, the previous examples presumably represent (a) a person, (b) a spoken sentence, (c) the street plan of a city and (d) a three-dimensional object such as a car. At first sight these are all objects except for (b), which is a Signal. However, we argue that the true antecedent is a Signal in *each* example. For example, the structure of a digital photograph doesn't really correlate with the person it portrays; rather, it correlates with a two-dimensional visual projection of the person. This is the case exactly because in order to create the photograph, light originating from the person was captured when it crossed a planar section of the camera. So while the ontological category of the representation itself may be ambiguous, its antecedent is a Signal without doubt.

Discussion

It seems self-evident that a representation is composed of indicators, each encoding a property of the structurally correlating part of the representee. To stay with the digital photograph example: each pixel indicates the color of one coordinate in the planar visual projection of the person at the time when the picture was taken. Cummins and Poirier seem to agree with the idea that an image is composed of indicators. This analysis of a representation aligns perfectly with our Signal-based definition of a Representation.

Cummins and Poirier suggest that a representation is source-independent in a way that indicators are not. In particular a representation may refer to an antecedent that “isn’t there”; a photograph can represent a person long after the person has gone. However, this is not something that distinguishes a representation from an indication because of the structural correlation, as the authors suggest. Instead, it’s simply because they disambiguate representations as being the object that produces the Signal and they chose to treat indications differently; when we realise that a pixel in a photograph may still indicate the color of a coordinate in a planar visual projection long after the latter has disappeared, the distinction is revealed to be false.

For now we can conclude that the Representation from our framework is really a representation rather than an indication. Symbols, which every Representation is composed of, are indications of the corresponding Symbols in the antecedent Signal.

2.3.2 Intent versus interpretation versus causal antecedent

A recurring issue in the study of representation is the possibility for a representation to “miss its target” [5, 6]. We argue that this possibility is introduced by an important ambiguity about what it means for a representation to be “about” an antecedent.

Cummins and Poirier distinguish what they call the target from the content of a representation, or the “target content” and the “actual content”. The former denotes the intent of the person producing the representation while the latter denotes a type of content that is somehow inherent to the representation itself. The example brought forth by the authors is a situation in which two people inspect the collection of photographs of opera singers owned by one of them. When the other person asks what Anna Moffo looked like, the owner inadvertently hands a picture of Maria Callas. The target is Anna Moffo while the “actual content” is Maria Callas.

Obviously, in this case the mismatch between intent and “actual content” is due to an error on the part of the sender of the representation. It is easy to envision a similar mismatch on the end of the receiver. Imagine, for example, that the owner correctly hands a picture of Anna Moffo, but the other person—due to a peculiar frame of reference—complains that he or she was given a picture of Roberta Peters. Here, the intent and the “actual content” align, but the *interpretation* does not match with either.

In fact, we can imagine that both errors would happen at the same time: the owner intends to show Anna Moffo but hands a picture of Maria Callas, which the other person mistakes for Roberta Peters. In this case, we would still intuitively insist that Maria Callas is the “actual content” of the picture, even though neither of the participants in the exchange is aware of it. So what determines this “actual content”?

As far as Cummins and Poirier are concerned, the answer is structural correlation: the very fact that the picture is isomorphic to (a planar visual projection of) Maria Callas makes it a picture of Maria Callas. This answer however is problematic because it doesn't match our intuitions about accidental images. For example, consider the kind of image that may be formed from the thin film of debris on an office window on a hot sunny day. While such images are vague, human spectators are apt at recognising various themes such as the weeping Mary or Santa Claus. Even when the similarity is very convincing an accidental image is usually still ambiguous, so how do we decide which interpretation is the "actual content"? A numerical approach to objectively identify the strongest similarity seems inappropriate here. Our intuition tells us that the apparent content of such images is purely in the eyes of the beholder; indeed, there need not be any relation between the "actual content" of the image—assuming such a thing exists—and its interpretation.

Correlation thus seems insufficient. What we've been missing so far is a stronger necessary condition: a causal relation between the representation and what is represented. Consider the picture of Maria Callas: its "actual content" is Maria Callas exactly because Maria Callas posed for the photograph. We may simply say that (a planar visual projection of) Maria Callas is the *causal antecedent* of the picture. Causal antecedent seems to be sufficient for something to be what Cummins and Poirier call the "actual content" of a representation. This agrees with our definition of Representation requiring Translation, which implies a causal relation.[8]

By accounting for the "actual content" of a representation however, we have not yet explained how there can be a mismatch with its "target". In previous attempts to naturalize representation, both Dretske and Millikan therefore introduced the notion of a *function* of representation [6, 7]. Function, by itself, is potentially problematic to naturalize, but both authors appeal to evolution (and some other causal processes) to solve this.

Dretske defines function on the side of the producer of the representation. For example, beavers have evolved the behaviour to splash with their tails, with the function to signal danger to other beavers. The content of tail splashing therefore is danger. Dretske states that when the behaviour fulfills its function, i.e. when there is actually danger, it is a representation; otherwise, i.e. in case of false alarm, it is a misrepresentation. Unfortunately this account does not work for mismatches on the side of the consumer. In the example above where a picture of Anna Moffo is mistaken for Roberta Peters, Dretske would conclude that both the function and the content of the representation is Anna Moffo, because Anna Moffo ultimately caused the photograph. In other words, in this case the way in which the representation appears to the consumer does not bear any consequence to its target or content. On the other

[8]For an additional justification of this requirement, consider the following: images that causally originate from their antecedent tend to bear much stronger similarity to the interpreted antecedent than accidental images. After all, it's so vastly unlikely for a chance event to produce an accurate depiction of anything, that we may just as well call it an impossibility.

hand, according to Dretske a hallucination of a unicorn does not have a cause external to the hallucinator, but its target and content are the properties of a unicorn, so the appearance to the consumer *does* determine the target and content of the misrepresentation. This is inconsistent. [6]

Millikan defines function on the side of the consumer of the representation. For example, beavers have evolved the function to interpret tail splashing as a warning signal and flee, because this contributes to their survival. The content of a representation is determined by the kind of situation where the function is appropriate: danger in this case. Unfortunately this is circular reasoning: it was danger, i.e. the content of the representation, that made the function to flee appropriate in the first place. In this way Millikan effectively loses the distinction between target and content. In fact, when beavers flee because of tail splashing while there is no danger, or when they flee because of imaginary tail splashing, she refuses to attribute any aboutness to the interpretation–as–a–warning-signal and the consecutive fleeing behaviour. Misrepresentation is simply equated to the absence of representation. [6, 7] This is unsatisfying.

It seems unsurprising that Dretske and Millikan fail to fully account for misrepresentation, given that both identify only two sources of reference that need to agree in order for a representation to succeed, while our photographs example made plausible that we may need three sources. In addition, when one of the sources is identified to be a *function*, the challenge to naturalize representation has essentially been *expanded* by the need to link function to reference, rather than simplified. Instead, we will now define the three modes of "aboutness" of a Representation purely in terms of Transducers and Signals.

The ambiguous statement that X "represents" Y may either mean that

causal antecedent X is the output Signal of a Translator that took Y as the input Signal;[9]

intention X is the output Signal of a Translator which has the same state as when it would Translate Y;

interpretation X is the input Signal of a Consumer which has the same state as when it would read a Translation of Y.

[9] When X is a photograph of Anna Moffo, the Translator under consideration is the camera and the causal antecedent is the light passed through the shutter when the picture was taken. Obviously the weather, the breakfast of the photographer and many other factors influenced the appearance of the particular picture as well. These factors were however already encoded in the *state* of the camera when the picture was taken. The light passing the shutter was the only actual input Signal at the time of Translation. So, while *a* causal account for the contents of a Signal will tend to be fuzzy and complex, *the* causal antecedent of said Signal is sharply defined. (Because of the transitivity of Translation, Anna Moffo herself is still an indirect causal antecedent of the picture in the "materialised signal" sense.)

When these three modes do not all agree about Y, we may speak of a "misrepresentation", "misinterpretation", "miscommunication", etcetera.[10] Note that disagreement requires discriminative power on the part of the involved Transducers, by having distinct states for non-matching Y signals. For example, when a Transducer interprets a signal as Y_1 while its causal antecedent was Y_2, this is only a disagreement if that Transducer has distinct states for reading Translations of Y_1 and for reading Translations of Y_2.

We can now show how these modes of aboutness allow us to tackle the examples. We already saw how the triple distinction naturally fits to the photographs example. X is the photograph. The picture may causally originate from Maria Callas, while the sender is in the state of handing a picture of Anna Moffo and the receiver ends up in the state of accepting a picture of Roberta Peters. Since the sender is able to distinguish Anna Moffo from other opera singers and the receiver is able to distinguish Roberta Peters from other opera singers, causal antecedent, intention and interpretation all disagree, so this is a misrepresentation in all respects. On the other hand, if the sender intends to show Anna Moffo and also hands a picture that causally originates from Anna Moffo, while the receiver interprets it unspecifically as the portrait of some unknown opera singer, there might not be a disagreement because the receiver might simply be unable to distinguish Anna Moffo from other opera singers. In that case there is no misrepresentation.

In the case of the beavers, X is the tail splashing. The causal antecedent may be danger detected by the owner of the tail, or it might just be a nervous twitch while the beaver is otherwise relaxed. A disagreement between causal antecedent and interpretation occurs when another beaver enters the warned state because of the tail splashing while the cause of the splashing was a nervous twitch, or if it enters a different state while the cause of the splashing was danger. A disagreement between causal antecedent and intention may occur when the tail owner reacts to danger in a reflex by splashing but subsequently remains in a non-alarmed state, or if it enters the alarmed state while there isn't any danger. An example of disagreement between intention and interpretation would be that the tail splasher is in the alarmed state while the observer of the tail splashing does not assume the warned state. Note that we are able to capture misrepresentation in beaver tail splashing without any appeal to the notion of purpose; we can, however, do some justice to the intuition that there is a purpose by observing that beavers have dedicated states corresponding to danger detection and being warned.

Finally, representations of nonexistent things such as an image of a unicorn. This is a misrepresentation, which is hidden by the fact that intention and interpretation agree. A unicorn cannot be the causal antecedent of an image, but we may just as well say that the image "represents" a white horse with

[10]Of course, all three modes are ultimately causal; the state of a Transducer is a result of its past history.

a narwhal tusk mounted on its forehead. After all, when we follow the chain of events that led to the invention of the unicorn back in time, we will probably find that the horse and the narwhal tusk are the indirect causal antecedents. Thus, representations of nonexistent things will always have a disagreement between the causal antecedent and either the intention or the interpretation. Hence, such representations are always misrepresentations.

2.3.3 Transitivity

According to Cummins and Poirier representation is intransitive, because a representation of the pixel structure of a photograph of a person in itself is not a representation of the person's visual appearance. We now know that the truth of this claim depends on the ambiguity of "is a representation of": a representation of a digital image of a visual projection is still isomorphic with and causally originated from the visual projection, but the intent has changed. Let's explore this in more detail.

While upon reproduction the chain of events is likely to go from numeric representation to digital image, let's assume for the sake of argument that the digital image is first created from Maria Callas' visual appearance and that a numeric representation of the pixel structure of said image is created after that. For each stage in this chain of events, we will identify the intent, the interpretation and causal antecedent of the representation, as well as the manner of structural correlation with its antecedent.

In the digital image, the intent and interpretation are both likely to be the person or object of which the causal antecedent is a planar visual projection. To be very precise, the causal antecedent is the (reversed) image in the plane of focus of the lens of the camera corresponding to the field of view.[11] The pixels each indicate the average colour of one rectangular section of that causal antecedent.

In the numeric representation, intent has almost disappeared (and therefore changed): the state of the device that converted the image to numbers does not distinguish in much detail between different antecedents. The interpretation can be anything, depending on who or what is confronted with the numbers. The causal antecedent however is very clear: it is the digital image, and thence the projection in the plane of focus of the camera. The manner of correlation has also been partly preserved; each number indicates (an aspect of) either the position or the color of a pixel in the digital image, and therefore indirectly of a rectangular section of the original projection in the camera.

In conclusion, this example shows that intent and interpretation are not transitive, but causal antecedent is. This is good news because it agrees with our definition of Representation, which is the output of Translation, which is a transitive operation.

[11] Hopefully the plane of focus aligns with the plane of the sensor.

2.3.4 Sentences

An image can be accurate to a varying degree, for example depending on the skill of the artist or the state of degradation of the image itself. The same is true of maps and scale models. Intuitively, one might define the accuracy of a representation as the degree of structural correlation with its causal antecedent. However, Cummins and Poirier choose to define accuracy as the degree of similarity with the *target*, i.e. the intent of the sender. This leads to the observation that a sentence cannot more or less accurately represent a proposition; it either hits or misses.[12] Since a proposition cannot be represented more or less accurately, they then conclude that images, maps and scale models cannot represent propositions.

This line of reasoning seems problematic. It begs the question whether the target of a sentence is always a proposition, or even most of the time. In addition, a proposition is an abstractum relating possible worlds, rather than a signal or an entity, so it cannot be the causal antecedent of a representation. The causal antecedent of a sentence, on the other hand, may very well be more or less accurately represented. Finally, there is no theoretical reason why the target of an image couldn't be a proposition.

So while the reasoning seems to hold some truth, the choice to define accuracy as the degree of similarity between representation and target seems to introduce unnecessary complications. If we use the similarity to the causal antecedent instead we can treat sentences and images uniformly, and save the interaction between targets and propositions for another study.

This leads us to another question regarding sentences. It is easy to see that a sentence can structurally correlate with another sentence, regardless of the language or the medium. With some imagination, we can also accept that it might correlate with the train of thought of the person who produced the sentence. What is subject to very much debate, however, is whether there is any way in which a sentence could correlate with events in the physical world. [8] This is a challenge, because sentences can often be traced back to such events through the transitivity of Translation.

An opening to this issue may come from an unexpected direction: vector images. Vector imaging is a modern approach to digitally storing images in a parametric way, where composite shapes with attributes (such as colour and line width)—rather than pixels—encode the visual scene. Vector images form a bridge between traditional pixel images on the one hand and sentences on the other hand, because they construct an image from discrete parts, like pixel images, but have the internal structure of a recursive language.

The rectangular arrangement of pixels in a traditional image is not any more faithful to the structure of the original projection than the custom ar-

[12] We think this is not a certain fact. It might actually be possible to design a accuracy metric for propositions, based on the Hausdorff distance between the set of possible worlds corresponding to a sentence and the set of possible worlds corresponding to the target proposition.

rangement of complex shapes in a vector image. However, it begs the question whether the non-positional relations between shapes that an artist may include in the encoding, such as layering and masking, are present in a visual scene. These aspects of vectorial image encoding are convenience tools for the artist rather than means of representation. Something similar may be true of sentences: the smallest constituents may correlate with aspects of a physical event, where some words or parts of words may function as indicators, and to some degree there might be a correlation between the order of the events and the linear order of the sentence. However, the further hierarchical organisation of the sentence is an addition that functions not to represent the world, but to help the producer and the consumer of the sentence to handle its contents. Nothing in the definition of Translation forbids a Translator to "enrich" the output Signal.

2.3.5 Summary

We have seen that the content of a representation lies in its structure. This is no surprise given our definition of Information; what Information is present in a Signal, such as a Representation, depends on the contents of the Signal. The aboutness of a representation however is ambiguous; it may be causal antecedent, intent or interpretation. Our framework mostly regards the causal antecedent because it is mostly concerned with Translation, a property inherited from Shannon's approach; it can however also capture the other modes of aboutness and explain misrepresentation.

Our definition of Representation has proven to stand up against all requirements for something to be a representation. It isn't an indication, but composed of indications; it has content and aboutness; and while Cummins and Poirier maintain that representation is intransitive, this is not a problem for our definition because, upon closer inspection, they would probably agree that causal antecedent is transitive.

2.4 Turing completeness

To prove that Communication is Turing-complete, we'll show one way to build a classical Turing machine with abstract Transducers. "Classical Turing machine" here refers to the type of Turing machine with one unbounded tape and one read/write device which can move to adjacent cells on the tape in both directions by single steps, as originally described by Alan Turing [9].

Traditionally, the "state" of a Turing machine is considered to be a scalar which enumerates the possible internal states of the read/write device, not taking into account its position relative to the tape. We'll refer to this state henceforth as the Tu-state. Turing proved that a finite number of Tu-states is sufficient for any computation. Correspondingly, we refer to the internal state of a Transducer as the Tr-state.

Note that in addition to Tu-states and Tr-states, we also need to distinguish Tu-symbols from Tr-Symbols. The former are the symbols that may be written on the tape of the Turing machine that is being designed, while the latter are the Symbols that may be transmitted between the Transducers in the realisation of the design. Formal underpinnings are provided in the Appendix.

In Figure 2.5 our design is shown for a classical Turing machine. The Transducers are represented by boxes, with a triangle pointing to the output end. The lines that connect the Transducers represent the Channels.

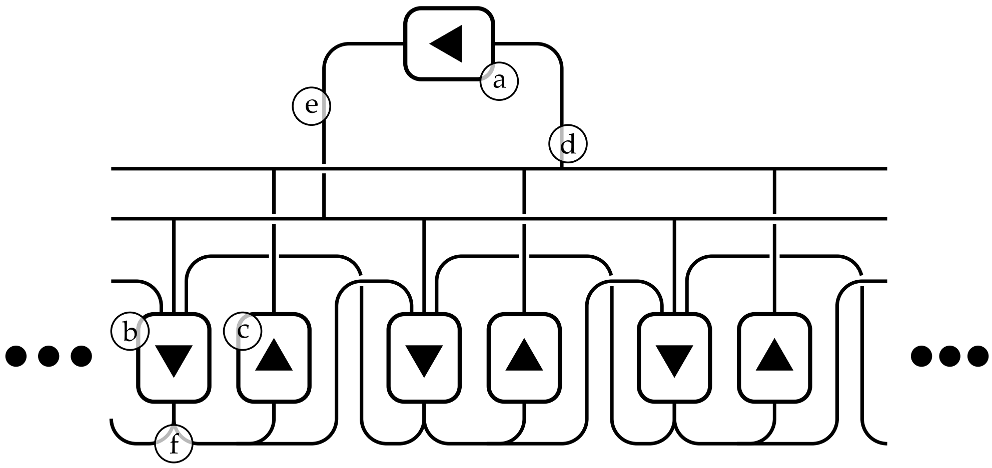

Figure 2.5: Design of a classical Turing machine.

Three (classes of) Transducers and three (classes of) Channels are labeled, respectively:

a. Device Transducer, corresponding to the read/write device of the designed Turing machine.

b. Switch Transducers, recognised by the downwards pointing triangle and the triple input. These keep track of the position of the read/write device relative to the tape.

c. Memory Transducers, the upwards pointing Transducers directly at the right of each switch Transducer. These store the content of a tape cell. One such pair of a switch Transducer on the left and a memory Transducer on the right fully implements one section of the tape. We may call the switch and memory Transducers that together correspond to the same tape cell each other's "partners". The triple dots at both sides of the drawing indicate that the number of tape cells may be unbounded.

d. Reading Channel. All memory Transducers connect to it on their output side, while the device Transducer receives input from it.

e. Writing Channel, connecting the output side of the device Transducer with the input side of the switch Transducers.

f. Private output Channel of a switch Transducer. It connects to the input side of the partner memory Transducer, as well as to the input sides of both adjacent switch Transducers.

The Tr-state of the device Transducer corresponds to the Tu-state of the designed Turing machine. The switch Transducers may be AtLeft, Active, AtRight or Inactive, where the former three options should occur exactly once directly next to each other in that respective order while all other switch Transducers are Inactive. The memory Transducers have one optional Tr-state Symbol$_i$ for each Tu-symbol that may be stored on the tape.

The following (classes of) Tr-Symbols may occur in some or all of the Channels:

- read, the tape moving Symbols left and right and the auxiliary control Symbols wake-left and wake-right.
- Tr-Symbols that match a particular Tu-symbol. These are indicated by the variable symbol$_i$, where i indexes over the set of possible Tu-Symbols.

The operational details are spelled out in the appendix. For now, we'll turn to Figure 2.6 and walk through the steps that occur during one execution cycle of the Turing machine:

1. The currently Active switch Transducer sends read. It is ignored by the neighbouring switch Transducers, but the partner memory Transducer is triggered.
2. The memory Transducer sends the symbol$_i$ matching its Tr-state Symbol$_i$. The device Transducer switches to a new Tr-state matching the Tu-state

that should be assumed according to the Turing machine's transition function and outputs two Tr-Symbols, which start the next two steps: the new Tu-symbol to be written to the current tape cell and the direction to move to.

3. The SYMBOL$_j$ triggers the ACTIVE switch Transducer.

4. a. The ACTIVE switch Transducer forwards the SYMBOL$_j$ from the previous step. This Tr-symbol is ignored by its neighbours but triggers its partner to switch to the SYMBOL$_j$ Tr-state.

   b. The moving Tr-Symbol triggers the three switch Transducers which are not INACTIVE. The new ACTIVE switch Transducer will output READ, completing the cycle.

5. As an example, suppose that the moving Tr-Symbol from the previous step is LEFT. In that case the ATRIGHT switch Transducer changes to INACTIVE, the ACTIVE switch Transducer changes to ATRIGHT, and the ATLEFT switch Transducer changes to ACTIVE as well as sending the Tr-Symbol WAKE-LEFT to its colleagues. This signal triggers only the INACTIVE neighbour at the left, causing it to change into ATLEFT. This coincides with the start of the new cycle.

Concluding, if the machine starts out with a valid state it is guaranteed to be in a valid state again each time it enters the cycle.

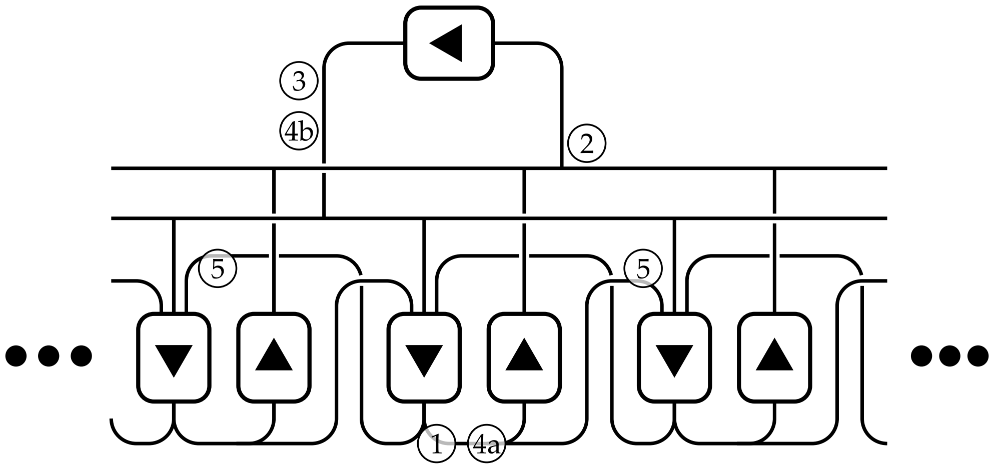

Figure 2.6: The same design as in Figure 2.5, now with numbers for the consecutive steps during one execution cycle of the Turing machine transition function.

Chapter 3

Applications

3.1 Thermodynamics and statistical mechanics

Thermodynamical entropy is often interpreted as a form of information or knowledge, or a lack thereof. In this section we will explore the origins of these information-oriented interpretations, of which there turn out to be several. We will also consider the compatibility of information-oriented notions with interpretations that equate entropy with disorder, and suggest that there is no conflict.

Finally we show that our framework can account for the different ways in which physical entropy is linked to "information" in a completely objectified way. If by "information" we mean the Mutual Information between macroscopic and microscopic state, then thermodynamical entropy is negative information. If we consider the microscopic state in isolation, thermodynamical entropy is positive information. We show that only the latter interpretation is in agreement with the definition of non-mutual Information as well as with the notion that entropy is disorder. For this reason it seems more natural to equate thermodynamical entropy to positive information.

3.1.1 Origins of information-oriented interpretations

Maxwell's demon

Not convinced that the irreversibility of the Second Law of thermodynamics could be derived from statistical mechanics alone, J.C. Maxwell invented a famous thought experiment in which the Second Law is violated. In this experiment, a very fine-fingered being with very acute senses—the demon, as it was later called by Lord Kelvin—operates a frictionless valve between two gas containers of equal volume. Initially the valve is open and the containers are equal in pressure and temperature. The demon then starts operating the valve, allowing only relatively fast molecules to pass to one container while allowing only relatively slow molecules to pass to the other. As the experiment progresses there is a net heat transport from the latter container to the former, without any energy being added to the system. Entropy decreases, thus violating the Second Law. [10]

According to Maxwell, the experiment shows that it is a lack of acute senses or a lack of *knowledge* that prevents us from violating the Second Law. This idea suggests a further line of reasoning where *more* knowledge would allow for a *greater* violation of the Second Law.

Maxwell's thought experiment was later taken as a threat to the Second Law itself, prominently by Leo Szilard. Szilard found that the irreversibility of processes had to be protected from cases like Maxwell's demon. He proposed that the demon itself must be introducing additional entropy to compensate for the entropy-reducing heat transport, either through the act of observing or *measuring* the state of nearby gas molecules, or by returning to a state in which it does not know yet whether it should open or close the valve. The latter is commonly referred to as "erasing the memory" of the demon. [11]

Léon Brillouin (1951, 1962) is notable for attributing an entropy increase to the measurement, and for making a particularly messy argument. On the one hand, he chooses a subjective interpretation of Boltzmann's W (see also page 32). W originally meant the number of possible microscopic states given the macroscopic state (the objective interpretation), but it can also be interpreted as the number of possible microscopic states given *what is known about the microscopic state*. Because Brillouin assumes the latter, entropy immediately decreases when an observer learns more about the microscopic state. For this reason he equates information to negative entropy. He then postulates a new version of the Second Law where the sum of information and negative entropy cannot increase. This, however, leads to a situation where the validity of the Second Law itself depends on the knowledge of the observer.

Others, in particular Bennett and Landauer, have chosen the memory erasure of the demon as the source of additional entropy. They conclude from the logical irreversibility of memory erasure that it must also be physically irreversible and therefore increase entropy—already assuming that the Second Law applies. [11] In this way they suggest that information is equal to entropy, rather than to negative entropy as claimed by Brillouin! Thus all parties involved in the discussion seem to agree that knowledge or information has something to do with entropy, but in radically opposed ways.

The mixing paradox of Gibbs

J.W. Gibbs (1875) observed that mixing two ideal gases of equal pressure and temperature wil raise the entropy when the gases are different, but not when they are of equal kind. First consider the case where the gases are equal: if a valve is opened between the containers, the gases are allowed to mix and the valve is then closed again, the original situation is restored without any work having been done. Hence the entropy has not changed.

In order to determine the entropy change for different gases Gibbs again considered what it would take to first mix the gases and then restore the initial situation. In order to separate the gases he proposed a telescopic arrangement of containers where the inner container had a semipermeable membrane on one end for the first gas and the outer container had a semipermeable membrane on the other end for the second gas—which however was somehow able to move on the *inside* of the inner container[1]. The procedure is illustrated in Figure 3.1. Gibbs calculated that no work would be required to separate the gases by extending this telescopic arrangement and that the entropy would

[1]This would not require magic, even though that would be allowed for the sake of the thought experiment. For example consider a loose membrane with a frame that is kept in place by a strong permanent magnet.

not change. At this point however the volume is twice that of the initial situation, so the two gases still have to be compressed. Doing so requires work and decreases the entropy; since this restores the initial situation, it follows that the mixing of the gases in the first place must have increased the entropy. [10]

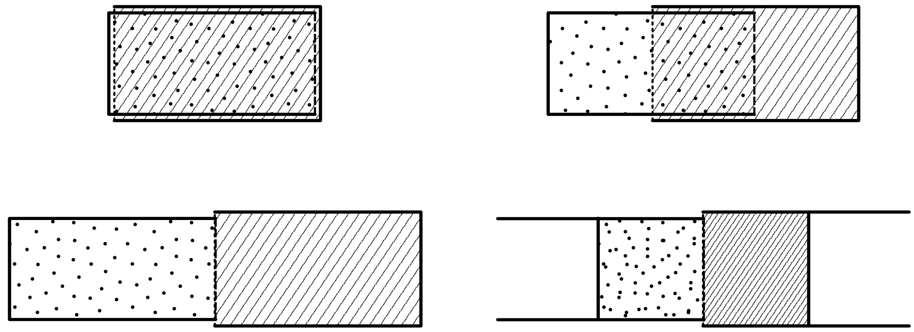

Figure 3.1: Gibbs' arrangement for different gases.
Top left: an equal mixture of different gases (polka dots and slanted lines) is held in a telescopic arrangement of containers. Top right: the containers are allowed to separate quasistatically, sorting the gases in the process. Bottom left: the gases are fully separated and no work has been done, but the total volume is now twice the original. Bottom right: the gases are compressed back to the original volume. Removing both semipermeable membranes would restore the original situation.

The paradox is that the entropy change does not depend on which gases are mixed. Very similar gases will have the same entropy increase when mixed as very different gases. At some point however gases stop being similar and they become the same; it is at this point that mixing the gases suddenly doesn't affect entropy anymore.

While Gibbs himself did not relate this observation to knowledge or information, the paradox does invite for such explanations. This can be done either in a subjective or an objective setting. The subjective setting, which is found in Jaynes and Grad among others [10], states that it is up to the observer to decide whether to consider the gases to be different. In this case the observer also decides whether the entropy has changed or not, so entropy depends on (subjective) knowledge. In the objective setting, the entropy will always increase if the gases are objectively different, but the observer will not be aware of that if they don't know that the gases are different. In this case entropy depends on information that may or may not be accessible to an observer, and the entropy itself may therefore also not be completely accessible.

Mathematical formulations of entropy in statistical mechanics

In the works of L.E. Boltzmann we find

- $H = \sum_i f_i \ln f_i$ with f_i the number of particles in a cell of the state space;
- $H = \int f(\vec{p}, \vec{q}) \ln f(\vec{p}, \vec{q}) d\vec{p} d\vec{q}$ with $f(\vec{p}, \vec{q}) d\vec{p} d\vec{q}$ the number of particles with momentum in the range $[\vec{p}, \vec{p} + d\vec{p}\rangle$ and position in the range $[\vec{q}, \vec{q} + d\vec{q}\rangle$;
- $H = \int f(\vec{v}) \ln f(\vec{v}) d\vec{v}$, like the above but with $f(\vec{p}, \vec{q})$ uniformly distributed over $\vec{q}$;
- $S = k \ln W$ with W the volume of the part of the state space that corresponds to the macroscopic state of the gas, and in equilibrium $\ln W \propto \sum_i n_i \ln n_i$ with n_i the number of particles in a cell of the state space and $S = -H$. [10]

In the works of Gibbs we find

- $\sigma_\rho = -\int \rho(x) \ln \rho(x) dx$ with $\rho(x)$ the probability density distribution over states x in an *ensemble* of states that are related by their macroscopic properties;
- $\Sigma(t) = -\sum_i V_i \rho_t(i) \ln \rho_t(i)$, the "coarse-grained" or *pixellated* version of the above that develops over time. [10]

Apart from being related and looking similar to each other, all of these formulas also bear striking resemblance to Shannon's information entropy:

$$H = -K \sum_i p_i \log p_i$$

with p_i the relative frequency of a symbol in a signal and K an arbitrary constant that determines the unit of information. Note that Shannon derived this formula completely independently based on reasonable criteria for a measure of the information density in a signal; he only borrowed the name "entropy" and the symbol H because he was aware of the resemblance with the above formulas from statistical mechanics. [1]

While it would be perfectly well possible for two different quantities in two different fields of study to have a similar formula by coincidence, especially such a relatively simple one, this particular resemblance goes deeper because both quantities are fundamentally based on probability density distributions. In addition it is not hard to accept, perhaps even uncontroversial, that there might *somehow* be more to know about gases with greater entropy—even though more precise formulations of what the relation between entropy and information may look like vary wildly. The apparent relation between the quantities is so strong that E.T. Jaynes (1983) decided to go the other direction and attempt to derive the Second Law using Shannon's information entropy. [10]

Summary

We have seen that there are various reasons to relate physical entropy to "information" and various formulations of such a relation have been put forward. Many of those formulations seem compatible with a notion that there is more to "know" about a system with greater entropy. For example, Brillouin suggested that learning about the microscopic state of a system also increases its entropy. On the Gibbs paradox, one could suggest that it is impossible to know that entropy has increased if you don't know that two different gases were mixed.

In addition, Shannon independently arrived at a measure for the Information density in a Signal which is based on probability density distributions, in the same way as the formulas for physical entropy that Boltzmann and Gibbs derived in statistical mechanics. This is suggestive of the relation between physical entropy and Information being fundamental.

Differences between the ways in which authors relate physical entropy to "information" seem to arise partly from the lack of a rigid definition of information, and partly from subjective versus objective interpretations of both "information" and physical entropy. We will see that our framework allows us to abstract away these differences and give a definitive answer on how the concepts should be related.

3.1.2 Entropy as disorder

Thermodynamical entropy has traditionally been described as disorder at microscopic scale. A modern variation is to speak of *energy dispersal* at microscopic scale, meaning that the energy carried by the particles manifests in more diverse ways at greater entropy. The basic idea remains the same, i.e. that entropy "complicates" categorization of particles in classes of states. Entropy is zero when all particles share the same state. The analogy with information entropy is easily made. In signals with greater information density categorization of the symbols is more complicated, either because there are more different types of symbols, because the symbols are more evenly distributed over the types, or both. A signal with only one type of symbol has zero entropy. Below we will discuss three concrete examples.

Expansion of a gas

It may not seem immediately obvious that an expanded gas is less ordered at microscopic scale than the same gas before expansion. This can be made more intuitive by comparing the gas with a crate of Lego bricks. Lego bricks in a crate do not appear particularly well ordered, for example they are not stacked in a regular pattern of rows and columns, nor will they remain sorted by colour for very long. Still, when the collection of bricks is allowed to expand, for example by overturning the crate above the floor, the bricks do become even more disordered. They are not "neatly in one place" anymore. The exact same thing is true of particles in an expanded gas. The analogy is very strong: when you topple over the Lego crate it will require work to get all the bricks back in their orderly place.

A more formal way to put it is as follows. For any compartimentalisation of space, particles in the gas can be categorised by the compartiment that they reside in. Given any regular spatial compartimentalisation at any resolution finer than the greatest diameter of the expanded gas, the particles of the expanded gas will be distributed over more distinct categories than the particles of the unexpanded gas. Hence the particles of the expanded gas are less ordered.

The direct analogy with information theory is that the symbols in an "expanded" signal belong to more different types than the symbols in an "unexpanded" signal. The "surprise value" per symbol is greater when the number of symbol types is greater. The same thing is true of particles in a gas: it is a greater "surprise" to find the position of a particle in the expanded gas than it is to find the position of a particle in the unexpanded gas. Hence, the expanded gas contains more information.

Mixing

It is intuitive that a mixed system is less ordered than two "sorted" systems each half the size of the mixed system. In addition, in the particular case of mixing two different ideal gases by removing a wall between the containers, as we saw in the Gibbs paradox, we may observe that each gas is allowed to expand to twice its original volume. It follows that twice as many classes of states are required to categorize the particles of each gas. However, we can also approach this from a different angle without depending on intuition or expansion.

In a gas where all particles are of the same kind, the state of each particle can be fully categorized by its position and momentum. If the gas on the other hand is a mix of different kinds of particles, each particle is not only categorized by its position and momentum but also by its kind. Since kind, position and momentum may vary independently between particles, the number of classes of states that was required for the pure gas is now multiplied by the number of kinds of particles. It follows that the information density or "surprise value" per particle also increases, by an amount logarithmically proportional to the number of state classes. In the special case where there are equal amounts of two kinds of particles, the added information per particle compared to a pure gas is exactly 1 bit.

Phase change

Apart from the expansion that tends to occur when matter changes to a more energetic phase, particles also behave less regularly in the more energetic phase than in the less energetic state. For ease of discussion, we will consider the position of particles in the transition from solid to liquid phase only.

In the liquid phase, particles move more or less freely around. Categorization by position in a liquid is therefore similar in principle to categorization by position in a gas.[2] In the solid phase, particles vibrate around fixed positions in a regular grid. Therefore position needs only be categorized relative to the expected position of the particle. Obviously, this results in a much smaller number of state classes for the same spatial resolution. It follows that particles in a solid are more ordered and that a solid contains less information than a liquid.

3.1.3 Reconciliation

It is time to solve the confusion about the relation between thermodynamical entropy and information that arose from Maxwell's demon and the Gibbs paradox. According to our framework, Information is reduction of a Signal space. To apply this notion to the discussions in this section, we need to identify the candidates for the Signal first. There are two candidates: on the one hand, any complete encoding of the macroscopic state of a gas such as the combination of pressure and temperature; on the other hand, the microscopic state of a gas.

Let us first examine the Mutual Information between both kinds of Signal. Macroscopic states that correspond to high entropy provide less Mutual Information about the microscopic state than macroscopic states that correspond to low entropy, because a larger fraction of the microscopic state space corresponds to high entropy macroscopic states. So, if by "information" we mean Mutual Information between macroscopic and microscopic state, Brillouin appears to be right that "information" corresponds to negative thermodynamical entropy.

If, however, we take "information" to be reduction of the microscopic Signal space alone, we find that higher entropy systems contain more Information per particle. This is the pure, non-mutual type of Information. Hence, a high entropy gas as a whole contains more Information than a low entropy gas with the same number of particles. On this account, Bennett and Landauer were right that "information" corresponds to positive thermodynamical entropy.

Paradoxically, this means that for a higher entropy gas, a Transducer will learn less from reading the macroscopic state, but more from reading the mi-

[2]The entropy difference between a liquid and a gas lies mostly in the expansion as well as in the velocities of the particles.

croscopic state. In other words, whether a Transducer knows much or little about a microscopic state depends both on the physical entropy and on whether it learned from the microscopic or the macroscopic Signal. This is not a peculiarity of our framework, but of the laws of physics; emergent properties cause Information on one organisation level, such as the position and momentum of particles, to be unequally "compressed" into Information at a higher organisation level, such as temperature and pressure of a gas.

The interpretation of entropy as disorder or energy dispersal is solely based on the microscopic state. It is no surprise that entropy corresponds to positive information in this case, too. We have seen that there is a very clear mapping between disorder among particles in a gas and diversity among Symbols in a Signal, and hence between entropy in statistical mechanics and entropy in information theory. For these reasons, it is most straightforward to equate thermodynamical entropy to positive Information, even though our framework can fully account for the "negative information" interpretation as well.

3.2 Quantum information

Quantum systems behave differently from the systems that have been studied in classical physics. It is a natural question to ask what difference quantum systems can make for communication and information processing; this is the subject of the new field of quantum information theory. Quantum systems turn out to offer interesting new opportunities for information processing, such as quantum computation and teleportation, and the study of quantum information theory has deepened understanding of quantum theory as a whole. These outcomes of quantum information theory have motivated some to suggest that information (however defined) may be ontologically more fundamental than matter.

C.G. Timpson (2006) provides a comprehensive overview of these matters [12] and we will use his discussion as a starting point for our own discussion. In the next few subsections, we will first compare quantum information with classical information as analyzed by Shannon. We will then briefly touch upon the ontological claims made by some researchers in the field and the associated problems, using teleportation and superdense coding as case studies. In both cases, the problem is that information appears to be transported either in larger amounts than is theoretically possible, or before it is created. After that we discuss Timpson's answer to these claims, which solves the problem by observing that information is abstract but which has the disadvantage of conflating Signal and Information. Finally, we show how both problems can be avoided if we recognise that teleportation is a transfer of state, rather than a transfer of a Signal.

3.2.1 Quantum vs. classical information

Quantum information theory is completely compatible with Shannon's mathematical framework, as well as our generalisation, but certain simplifying guarantees are not applicable anymore. The starting point is again the Signal, a sequence of Symbols. A Symbol corresponds to the state of a quantum system, which may be any continuous superposition of the basic states it can assume, which in turn may be discrete—for example, the polarization of a photon. Since the superposition is continuous it may take an infinite amount of classical Information to exactly describe a quantum state, even if the basic states are discrete. Transducers are still characterised by a transition function where the next state and the next sequence of output Symbols depend on the current state and the current input Symbol.

The critical complication is in the uncertainty of measurement. If the consuming Transducer does not measure (read) the Symbol in the same orthogonal basis that the producing Transducer prepared (wrote) it, the outcome is random. For example, consider the case where the producer sends the *qubit* $|1\rangle$. If the consumer reads the Symbol in the same basis, it will also receive

$|1\rangle$. If the Symbol is read in a different basis however, the consumer may receive either $|1\rangle$ or $|0\rangle$ with equal probability. What basis a Symbol is prepared and measured in depends on the state of the Transducers that send and receive it, respectively. The general consequence of this complication is that the transition function of a Transducer is nondeterministic[3] and that an output–input connection between Transducers is not a Trivial Translator, unless the Transducers involved are predetermined to always use the same basis or the Transducers pass non-quantum Symbols to identify the basis of each quantum Symbol.

Apart from the uncertainty of measurement, quantum information theory adds entanglement. Entangled systems have a joint state which can be influenced by affecting either system, after which the state change is reflected in both systems. This gives quantum Computation an efficiency advantage over classical Computation, in particular through the possibility to evaluate disjunctions without needing to evaluate the terms, and enables new communication techniques such as *superdense coding* and *teleportation*.

In teleportation, an arbitrary quantum state can be transferred between Transducers that share a maximally entangled pair of quantum systems using only minimal classical communication. This is illustrated in Figure 3.2. After the systems have been prepared with their shared entangled pair (steps 1 and 2), they only need to Communicate two classical bits in order to be able to transfer any arbitrarily superposed quantum state. The sending Transducer performs a joint operation on its half of the entangled pair and the quantum state to be teleported (step 3) and then uses two classical bits to convey the result of the operation (slanted dashed line). The receiving Transducer performs the same operation on its own half of the entangled pair, reconstructing the thus teleported quantum state (step 4). The entanglement is destroyed in the process.

Superdense coding, as illustrated in Figure 3.3, is similar to teleportation in that the Transducers share an entangled pair on which measurements are performed. The purpose and the procedure, however, are mostly opposite. Rather than sending two classical bits in order to restore a quantum state at the receiver, a qubit—which belongs to the entangled pair—is sent which enables the receiver to restore two bits of classical information. The key is that the entangled pair of qubits is known to be prepared in one of the four Bell states (steps 1, 2). The sender in the protocol performs one of the Pauli operators to its half of the pair, which flips the joint state of the entangled qubits to another Bell state (step 3). When the receiver in the protocol comes in possession of both qubits in the pair (step 4), it can perform a measurement in the

[3]At first sight this is directly in conflict with our definition of Transducers, which are supposed to be deterministic. However, determinism is not strictly required for the formalism to work. The only thing that is lost with determinism is the possibility to have non-singular Transducers, which most applications of the framework do not depend on. We will see that quantum communication techniques specifically reintroduce determinism when needed.

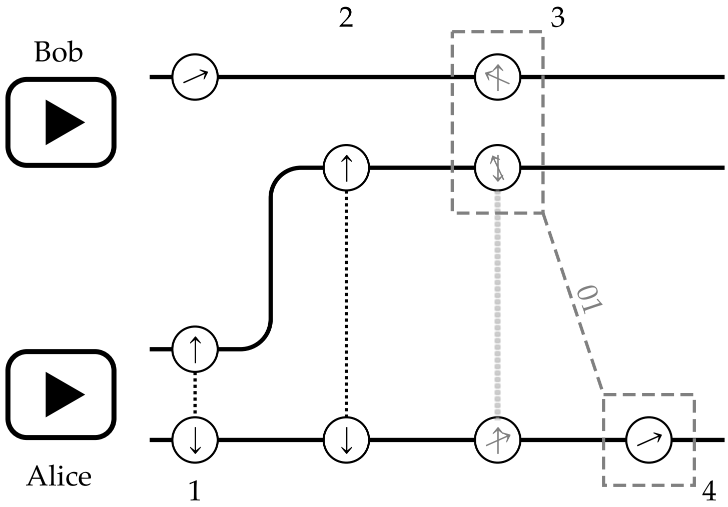

Figure 3.2: High-level overview of the quantum teleportation protocol. The vertical axis represents space, the horizontal axis represents time with the future to the right. Circles with arrows inside represent qubits, dotted lines indicate entanglement.

Bell basis on both systems to determine which of the four Pauli operators has been applied; this corresponds to two classical bits of information.

3.2.2 Ontological problems

Under the motto "information is physical", some (prominently Wheeler and Landauer) have gone so far to suggest that information is the most fundamental physical stuff or, equivalently, that matter can be reduced to information [12]. This view suggests that any quantum state is properly considered to be a piece of information—an infinite amount of Information in fact, as it would require an infinite amount of classical bits to perfectly specify the superposed state of a quantum system. This in turn complicates understanding of teleportation.

The apparent problem is that the Transducers involved in the teleportation transmit only two bits of Information, yet the consumer of those two bits reproduces a large amount of Information that originated at the producer. Timpson names three solutions that could be or have been proposed:

1. According to Jozsa and Penrose, the quantum state Information travels backward in time from the producing Transducer to the time and place where the entangled pair was created, and then forward in time to the consuming Transducer. In effect, this places Information in a peculiar

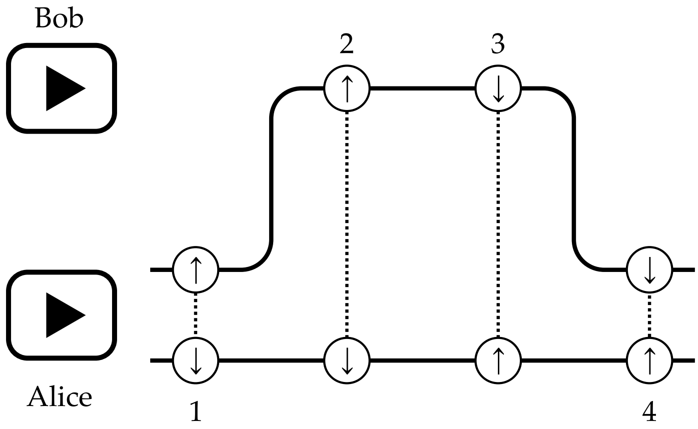

Figure 3.3: High-level overview of the superdense coding protocol. The visual language is the same as in Figure 3.2.

position as it is not just the most fundamental stuff, but also the only stuff that can travel backward and then forward in time.

2. Deutsch and Hayden provide an alternative interpretation of quantum mechanics, which allows all the Information of the quantum state to be transferred in those two classical bits after all. This amounts to introducing a third species of Information: Information of which the Symbols appear in all ways like classical Symbols, even being measurable with certainty, but which nonetheless contain a hidden quantum state.

3. Timpson suggests that we may just as well consider the idea that the quantum state Information makes a *non-local* jump from the producer to the consumer. Like the solution proposed by Jozsa and Penrose, this puts Information in a peculiar position of being both the most fundamental stuff and the only stuff with a special power, but this time the special power is to jump non-locally.

In each case, the solution seems to introduce as many problems as it solves.

Similar ontological concerns arise with superdense coding. At first sight, only one qubit is transferred but the receiver is somehow able to extract two classical bits of information. The fact that two qubits are involved does not immediately resolve the paradox, because the sender applies a Pauli operator only after the qubits have been entangled and the sender needs to see only one half of the entangled pair. This again invites for explanations where information is travelling back in time, being magically compressed or jumping non-locally.

The proper solution, as Timpson rightly points out, is to abandon the idea that Information is physical and accept that it is abstract instead. In this case the paradox of teleportation does not arise in the first place, both because the teleported quantum state by itself is not Information and because abstracta are not subject to the same restrictions of causality and locality that concreta are subject to. Similarly, in superdense coding we should not ask how those two bits of classical information could "travel" through only one qubit. To maintain that information is abstract, however, one needs to identify what is information in the first place.

3.2.3 Timpson's argument

Timpson suggests, like we have done in the present work, that Shannon's framework provides an answer not only to the question of how much information is transferred in a Signal, but also of what is information proper. Timpson however attempts to extract this answer in a rather more straightforward manner. His starting point is the following quote from Shannon:

> The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another.

Timpson takes the liberty to interpret "the fundamental problem of communication" as more or less directly referring to information. He then arrives at the following definition which basically equates "information" to the "message" (p. 23):

> Information$_t$ is what is produced by an information$_t$ source that is required to be reproducible at the destination if the transmission is to be counted a success.

Unfortunately, to equate "information" to "message" seems to deny the possibility that information is something distinct from a Signal that may have both a quantity and a quality. Apart from being intuitively undesirable, this would amount to trivialising the concept of information as a means to dissolve the apparent paradox of infinite information transfer. We will return to this issue below.

The next step is to identify the ontological status of "what is produced by an information source that is required to be reproducible at the destination if the transmission is to be counted a success". Timpson correctly observes that a message is characterized by Shannon to be a sequence of elements (Symbols) drawn from an alphabet (Symbol set). At this point he introduces the distinction between *token* and *type*. The token is the physical realisation of a Symbol while its type is the abstract set element it corresponds to. For example, in a digital wire the token may be high voltage while the corresponding type is **1**.

Timpson convincingly argues that what is required to be reproducible for the transmission to be counted a success is the type rather than the token.

In the case of the digital wire it is the sequence of **1**s and **0**s that must be reproduced, not the exact voltage. Likewise, in a handwritten text it is not our concern to transmit the precise shape of the glyphs on the paper; we only wish to preserve the sequence of elements from the alphabet. Given that the type is what is required to be reproduced and that the type is abstract, Timpson finally concludes that information must be abstract.

While the conclusion that information is abstract is arguably desirable, the way it is arrived at unfortunately breaks it all. Shannon's mathematical framework and the entire field of information theory that it gave rise to is only concerned with the type, not the token. From the composability of Transducers we know that the message is just another Signal. It follows that Timpson has equated information to Shannon's Signal. This is redundant and ineffective to explain what information really is. For example, it provides no justification for the idea that the amount of Information (Signal) should be proportional to the logarithm of the length of the Signal, as a framework more faithful to Shannon's model would.

3.2.4 A new solution

Staying firmly rooted in our framework, we can offer a rigid and surprisingly simple solution. We have already defined Information, and it is abstract: to be precise, it is reduction of a space of Signals. As discussed before, this notion captures both quantity and quality of Information.

In order to tackle the teleportation issue, it is instructive to identify the Transducers and Signals. Two Transducers are involved, each holding one half of the entangled pair. The producing Transducer does a joint operation on its half of the entangled pair and the quantum state to be teleported, which is part of its transition function, and then sends a Signal of two classical bits. The consuming Transducer receives that Signal and uses the contained Information to perform the operation on *its* half of the entangled pair that reconstructs the quantum state thereby teleported. Again, this is part of its transition function.

In summary, the teleported quantum state appears in the transition functions of both Transducers, but not in any Signal. Hence, teleportation is primarily about transferring state, not about transferring information (which is also intuitively justified). There is a Signal involved in the protocol, but all it does is transfer two classical bits of Information to assist the state transfer. The state transfer itself is just that, so asking how sufficient amounts of Information could be transferred is to make a category mistake.

We can proceed and shed some light on how the state transfer is made possible as well. Prior to teleportation the Transducers have received their halves of the entangled pair, either because one of them created the pair and sent one half to the other Transducer, or because a third Transducer created the pair and distributed it. This sending and receiving of an entangled system is a regular Signal transmission. The receiving Transducer must have had a transition function which caused it to incorporate the entangled system. Effectively, the teleportation Transducers offer an interface for altering their transition functions, and the creator of the entangled pair uses this to *program* the receivers so they are able to teleport together.[4]

The resolution to the superdense coding paradox is analogous. In the initial situation, the sending and the receiving Transducer each hold one half of an entangled pair, which was prepared in one of the four Bell states. Since the pair is entangled, any state transition in the sender that involves its half of the pair is also a state transition in the receiver that involves *its* half of the pair—the Transducers themselves share part of their state. Next, the sender applies a Pauli operator, which is a state transition for both Transducers, and sends its half of the pair to the receiver. This is a Signal of one qubit. The qubit is incorporated in the state of the receiver, which means that the receiver is now in possession of both halves of the pair.

This is where the story ends! Prior to transmission, the receiver holds one qubit from the entangled pair that can affect its future transitions; after transmission, it holds two qubits and both qubits can affect its future transmissions. In conclusion, there is really only one qubit of Information transferred, hence only one classical bit. The other qubit had been available at the receiving end all the time; it just happens to be the case that it featured in the transition functions of both Transducers.

[4]This perspective is elegant in describing the principle of *entanglement swapping*, but this is left as an exercise to the reader.

3.3 Biology

There has been a vivid discussion in biology on the merits and faults of the use of information-related concepts, especially with regard to genetics. Most of the debate can be attributed to either fuzzy, somewhat confused notions of Shannon information, or to the absence of any definition of information (as pointed out by Sarkar 1996 [13]). In their entry on Biological Information in the Standard Encyclopedia of Philosophy, Godfrey-Smith and Sterelny [14] provide an excellent overview of the arguments. They explain that a distinction between "weak" and "strong" senses of information is of crucial importance to understand the discussion:

> One common way to start organizing the problem is to make a distinction between two senses of "information," or two kinds of application of informational concepts. One of these is a weak or minimal sense, and the other is stronger and more controversial. In the weaker sense, informational connections between events or variables involve no more than ordinary correlations (or perhaps correlations that are "non-accidental" in some physical sense involving causation or natural laws). A signal carries information about a source, in this sense, if we can predict the state of the source from the signal.

The authors suggest that the "weaker sense" of information was established by Claude Shannon. The "stronger sense" of information is later identified as *teleosemantic*. We will investigate both notions in the next two sections.

In the "weaker sense", applying informational concepts to natural phenomena fails to add any explanatory power because direct causal relationships are insufficiently acknowledged, while the "stronger sense" is fuzzy and controversial. It turns out that a rigid and consistent appeal to Signals and Transducers removes the need for maintaining two separate "levels" of information. We can address all motivations for the problematic notions of teleosemantic information by re-emphasizing the causal structure of Communication and by applying the three modes of "aboutness" we discussed in the section on Representation.

3.3.1 Weak information: correlation

Directly after the distinction quoted above, Godfrey-Smith and Sterelny continue with the following interpretation of Shannon, adopted from Sterelny (2000) [15], in which a causal relation between a piece of information and the antecedent is optional:

> …This sense of information is associated with Claude Shannon (1948), who showed how the concept of information could be used to *quantify* facts about contingency and correlation in a useful way, initially for use in communication technology.
>
> For Shannon, anything is a *source* of information if it has a number of alternative states that might be realized on a particular occasion. And any other variable *carries information* about the source if its state is correlated with the state of the source. This is a matter of degree; a signal carries more information about a source if its state is a better predictor of the source, less information if it is a worse predictor.

This interpretation appears to be too liberal. Shannon defined the amount of *mutual* information in such a way that it could conceivably be applied to causally unrelated Signals, but he designed the measure under the assumption that one Signal would be a Translation of the other. This is a natural assumption because Shannon was primarily interested in the task of reproducing a message from a source at a destination through a chain of Transducers. Indeed, in general, increasingly long Signals are increasingly unlikely to have much mutual information unless they are related through Transducers.

It is a digression, however, to focus on the mistakes in this interpretation. The heart of the matter is that interpretations of Shannon where causality is not required are common ground, and that such interpretations still have operational value because they facilitate quantitative analyses. This operational value is shown clearly in a field like Computational Biology. [14]

The "just correlation" notion of information falls short when the goal is to *explain*, for example, the relation between a gene and the protein it encodes. The fact that there is a correlation between the sequence of base pairs in a gene and the sequence of amino acids in the protein does not explain anything about how genes and proteins come about, or why there is a correlation anyway.

It is exactly the omission of causality from the notion of information that causes this problem. In the case of the gene and the protein, the gene and the protein are Signals and the Transducer—the causal link—is the molecular machinery that Translates the former into the latter[5]. It is the existence of this Translator that explains why there is a correlation and how the protein comes about. The transition function of this Translator is known as the genetic code.

All other motivations for a "stronger sense" of information are ultimately closely related to this same consideration. We will address these motivations in the next section.

[5]This is a composite Transducer, consisting at the very least of RNA polymerase which transcribes the DNA sequence to an RNA sequence, and a ribosome which translates the RNA sequence to the final protein.

3.3.2 Strong information: teleosemantic

There are several intuitive reasons why a stronger, more profound concept of information may be attributed to genes in particular. Among them are the idea that genes have a *prescriptive*, rather than a *descriptive*, function (telic) and the idea that a gene may "stand for" or (mis-)represent a protein (semantic) even if it fails to accurately encode it, for example due to a mutation. These ideas are strongly inspired by analogies with human communication systems. Another prominent idea is that a gene may "stand for" a macrophenotypic trait rather than just a protein. We will discuss these ideas and their associated problems in turn.

Description versus prescription

Genes (and associated regulatory mechanisms known as *epigenetic information*) are often regarded as the blueprint or program for the development of an organism. As such, they do not only describe or predict the phenotype, but carry *instructions* for its realization. The intuitive justification for such a viewpoint is that genes, like recipes, blueprints and musical scores, seem to encode whatever is going to be produced by the entity that reads the instructions. The reading entity might be the cellular machinery that translates DNA to proteins for genes, while for the human counterparts it may be a chef, a construction worker and a musician, respectively.

There is a clear problem with this analogy: all established examples of instructions involve the participation of human agents, especially a designer who gives the Signal the *purpose* of providing instructions to the recipient. At least, this is how instructions are generally understood. Genes, on the other hand, are assumed not to have any conscious designer that could give them the purpose of providing instructions.

This apparent conflict leads some authors to conclude that genetic information cannot be telic, while others search for alternatives to the conscious designer. In most cases, natural selection takes this role. The reasoning works as follows. Assume a gene X that translates to a protein Y and suppose that individuals with Y in a previous generation got more offspring that individuals without it. In that case the relative frequency of X has been increased in the present generation thanks to its property of translating to Y. Thus, natural selection has given X the purpose or *function* of translating to Y. A notable proponent of this type of approach is John Maynard Smith (e.g. Maynard Smith 2000 [16]).

While authors have their own specific solutions, the conflict remains and the field is largely divided into two camps, each with its own problems. In the non-telic camp, the opportunity is lost to use information as an explanatory device. In the telic–by–natural–selection camp, a new natural category is postulated which is hard to justify: apart from conscious entities, which are capable of producing instructions, and "normal" nonconscious phenomena, which are not—a distinction which in itself is already problematic—, we now also have natural selection, which is not conscious but nonetheless capable of producing instructions. Instead of taking sides in this conflict, let us instead review its origin in terms of our framework.

The alleged examples of instructions so far, i.e. both genes and the human analogies, are all Signals which may be Translated into something else by the recipient (a Transducer). In both cases, the latter Translation is optional; the recipient may, for example, destroy rather than Translate the Signal[6]. On the producing side, however, there is a difference between genes and human instructions.

In human instructions, the producing Transducer is typically Translating a prototype Signal that the recipient will ultimately recreate. For example, a composer will typically write down a score while playing the music that it represents. In genes on the other hand, the DNA sequence is copied from a template DNA sequence, never from the protein that it may later be Translated into. Indeed, the molecular machines that copy DNA sequences do not have states that allow them to *intend* (here meaning one of the three modes of aboutness) proteins.

That said, we need to look at two organisation levels. Molecular machines that copy DNA sequences explain why a particular gene is present in a particular cell. Natural selection explains why the same gene has a particular frequency within a particular population. The latter does involve a causal link from protein to gene; however, the former is more directly analogous to our examples of human instructions, where a Signal is Translated to its target. While gene frequency at population level is a Signal on its own, it is not Translated into proteins; it is Translated into a new frequency for the next generation, where the population is the Transducer and the expression of the gene affects the state of the population.

Error

Building on the notion that genes are instructions, it seems natural that a mutated gene is not merely an imperfect copy of the template from the previous generation, but a *failure to correctly represent* whatever the original gene encoded. In other words, apart from being instructions, genes may also "stand for" a particular protein, even if following the instructions faithfully would lead to a different protein. This may remind us of the section on Representation: if I intend to show what Anna Moffo looked like, but inadvertently hand over a picture of Maria Callas, does the picture still "stand for" Anna Moffo?

[6]For sheet music, this could be a shredder, while for DNA, there exist specialized enzymes that depolymerize it.

An appeal to the three modes of aboutness is exactly what we need. We can trace the semantics of error both at the molecular level and at the population level. At the molecular level, a mutation occurs when the DNA copying machinery creates a DNA sequence that deviates in any way from the original. If the protein-producing Translator is able to distinguish the new sequence from the original, we have a misrepresentation. This is not the case if both sequences translate to the same protein according to the genetic code.[7]

At the population level, misrepresentation can occur but has nothing to do with mutations. The frequency of a gene is a misrepresentation of the frequency from the previous generation if the frequency did not increase while the protein encoded by the gene had a positive effect on the number of offspring per individual, or if the frequency did not decrease while the encoded protein had a negative effect. New sequences are just new sequences with associated proteins; they may be beneficial or detrimental.

Macrophenotypic traits

The concept of genes predates the discovery of DNA. Before the development of molecular biology, genes were commonly assumed to encode obvious macroscopic traits, such as the neck length of giraffes and okapis and the seed colour of garden peas. The notion that genes may provide a blueprint for macroscopic traits, rather than just proteins, has never really disappeared. While current state of the art reveals that most macroscopic traits are not actually explained by any particular gene, and that genes interact in very diverse and complicated ways, there are a few observations that still fit the idea of genes providing programs or blueprints for macroscopic traits.

The prime example is the *Eyeless* gene, of which the expression seems to be necessary and sufficient for the creation of an entire eye anywhere on the body surface of a fruit fly. The gene controls the expression of other genes, leading to a complex cascade of gene expressions and ultimately the creation of an eye. It may appear that *Eyeless* basically says “make an eye here”, but this is highly controversial.

A relatively minor objection against the notion of genes as the blueprint for macroscopic traits, is the threat of preformationalism. If the genome would be a blueprint of what the entire organism will look like, this would suggest that the organism carries a small copy of itself as the hereditary information. This leads to an infinite regression. Most authors, including proponents of the macroscopic blueprint, agree that the genome does not work like that.

More importantly, the notion of a macroscopic blueprint seems to neglect the accepted knowledge that most macroscopic traits come about through a complex interaction of many genes as well as many environmental factors.

[7]There is, however, a second way in which a misrepresentation may be found at the molecular level. This is left as an exercise for the reader.

For Godfrey-Smith (2000), this is reason to argue that genes can only "stand for" proteins [17]. Jablonka (2002) adopts an opposite stance, where environmental factors take a prescriptive role just like genes.

We can offer some relief based on our previous observations. In general, we can point out that in order for a Signal to be a Representation of another Signal, the latter must be the complete input of the Transducer that output the former. Therefore, at the molecular level we can say that *Eyeless* caused an eye, while at the population level we can say that the benefits of eyes cause the frequency of *Eyeless*. For macroscopic traits in general, where there is a very fuzzy web of causal links, we cannot say much at the molecular level. At the population level, we can say that the state of the population as a whole, including the diversity of genotypes and phenotypes, causes the state of the next generation. The environment provides the input Signal to the population.

3.3.3 Summary

When we explicitly identify all Transducers involved, thereby realising that Transducers are composite and taking multiple organisation levels into account, and we use the three modes of aboutness to scrutinize any representation-related concepts, we find that a distinction between "weak" and "strong" senses of information is not needed. Moreover, our Shannon-based framework provides us with a sharp and definitive way to determine where information–centric concepts are and are not helpful in genetics.

We can explain the "weak" and "strong" notions of information very directly in terms of the Signal-Transducer framework. The "weak" variant arises when the concept of Mutual Information is applied while the causal relation between the Signals is ignored. The "strong" variant is the result of confusion between the molecular level, where a DNA sequence is Translated into an amino acid sequence, and the population level, where expression of a gene is Translated into an increased frequency of the same gene. This confusion leads to a comparison with systems in which the output of a Translator tends to recreate the input of a previous Translator, which is common in human Communication. The divide is entirely avoided if we carefully keep track of the causal relations between Signals.

Genes have in common with human instructions that the Signal has a standard target to which it is commonly (but not always) Translated by recipients, but the difference is that genes are never Translated from a prototype of the target Signal. Genes can be misrepresented at the molecular level, if the copy from the previous generation is not perfect and the protein Translation machinery can distinguish between the variants. Misrepresentation is also possible at the population level, when the frequency change of a gene is not in agreement with the fitness effects of the target protein. Informational relations can be drawn between genes and macrophenotypic traits, but this is generally not very helpful except in corner cases such as the *Eyeless* gene.

Chapter 4

Conclusion

While Shannon's mathematical theory of communication is mostly known for its quantitative approach to signals, entropy and channel capacity, it has more to offer. In particular, the way in which information is quantified allows us to also derive the quality of information and thus offers a complete definition of information: it is reduction of a signal space. One could say that the content is the direction of the reduction while the amount is the size of the reduction.

Equally important is the definition of transducers. Transducers model the way in which we usually understand physical entities. In addition they are closed under composition, which enables them to form feedback loops. The ability to compose and to make feedback loops render transducers Turing-complete as we have proven by simulation. Together, these properties lead to a formal prediction of nature's ability to create computation engines such as brains.

We have also been able to give a fully naturalistic account of representation based on signals and transducers: it is crucial to distinguish three modes of "aboutness", i.e. the causal antecedent of the representation, the intention of the sender and the interpretation by the receiver. Concluding, we were able to expand Shannon's quantitatively rigorous framework to also answer profound questions of meaning and reference while maintaining a strictly naturalistic stance.

Our Shannon-based framework proved well-equipped for current scientific discussions. In thermodynamics and statistical mechanics, we could account both for theories that equate entropy with negative information and theories that equate entropy with positive information. We were also able to make a clear case why the latter interpretation is more natural. In quantum information, we were able to solve the infinite information paradox of teleportation, by determining that it is a transfer of state, not information; we did so without being caught by the same pitfall as Timpson, who equated information to Shannon's "message". We could also explain superdense coding. As for biological information, we eliminated the need for a distinction between "weak" and "strong" senses of information, serving the purposes of both using our own causality-based information concepts, and we could outline in which cases information-related terminology was helpful. Concluding, our postulated definitions truly stood up to the task of providing a general approach to information in nature.

Consequences for Artificial Intelligence

Given that our framework predicts that computation is possible in nature independently of any particular brain architecture or human intervention, and given that we could account for profound aspects of meaning in examples of human communication without any appeal to hard-to-naturalize concepts such as "mind" or "consciousness", there appears to be no reason why general intelligence should be impossible outside a human (or any animal in general). In addition, given the very general properties of transducers, there appears to be no principal reason why such a general intelligence couldn't be constructed as a (human) artifact. We hope that this consideration could be a valuable contribution to the discussion on whether "hard AI" is possible.

Another potential contribution to AI lies in the composability of transducers. Systems composed of many small, relatively simple components are common ground in AI: this includes neural networks and Bayesian networks, but also integrated circuits. Our transducer-based approach might offer a tool to study such composed systems in a unifying way.

Natural levels of organization

Both in the application to thermodynamics and statistical mechanics and in the application to biology, the separation of distinct levels of organisation turned out to be of vital importance for clearing up existing confusion. In thermodynamics, the choice whether "information" is supposed to be *mutual information* between the macroscopic and the microscopic level or *information proper* about the microscopic level only, determines whether the relation between "information" and thermodynamical entropy was positive or negative. In biology, DNA sequences are signals at the molecular level which represent the sequences they were copied from while gene frequencies are signals at the population level which represent the reproductive consequence of the expression of that gene in the previous generation, and confusion of these two levels leads to highly problematic comparisons with human communication.

The fact that levels of organisation play a role in both cases is interesting, partly because it shows that levels of organisation are a fertile ground for confusion, but more importantly because transducers might prove an excellent tool for analysing the relationships between organisation levels. After all, transducers are closed under composition, so any natural transducer can be analysed either as a component of a larger transducer at a higher organisation level, or as a network of smaller transducers at a lower organisation level. The potential value of separating organisation levels has been illustrated many times before, for example in [18]; we believe it would be worth investigating whether the transducer framework could play a general role in multi-level analysis of complex systems.

Further applications

We believe our framework can capture many more concepts than the ones discussed in this paper, for example knowledge. A possible definition is that a Consumer *knows* a Signal if it has taken that Signal as its input and its state has changed accordingly, such that it would not have had the same state if it had not read that same Signal. The act of reading the knowable Signal by the to-be knowing Consumer may, for convenience, be called *learning*. Note how this definition of knowledge is strongly related to our definition of intention in Section 2.3.2; we consider this a desirable property. The *amount* of knowledge could be defined as the amount of Information present in the known Signal, or as the mutual information between the known Signal and the state Signal of the knowing Consumer. Rigid definitions of information-related concepts like this may be valuable in any philosophical or social field concerned with meaning.

In Section 2.3.4, we briefly discussed structural correlation of a sentence with its antecedent. This is a promising ground for cooperation between linguistics and information theory. Possible directions of further research include the question how the distinction between continuous and discrete channels interacts with our concept of an "enriched" signal, or how to account for the role of the context in the meaning of a sentence. In the latter case, the answer may very well be found in the state of the transducers that are involved.

The presented framework may be valuable in many other fields. In particular studies of decentralized systems come to mind, such as the structure of human organisations, self-organising behaviour and aynchronous multiprocessing using producer-consumer queues. A comparison with the *actor model* [19] and similar concurrency models from computer science may be worthwhile.

Appendix A

Formal characterizations

A *channel graph* is a tuple $G = \langle S, R \rangle$ where each node in S is connected to all nodes in R and all edges are directed from S to R.

A Communication Network is a tuple $W = \langle T, C \rangle$ with T the set of all Transducers in the Network and $C \subseteq T \times T$ the set of all output–input connections in the Network.

A Channel c in W is any subset of C such that the pair $\langle S_c, R_c \rangle$, with $S_c \subset T$ the starting vertices of c and $R_c \subset T$ the ending vertices of c, is a channel graph and there is no superset c' of c that is also a Channel.

A Transducer $t \in T$ in turn is a tuple $\langle X_t, Y_t, A_t, F_t \rangle$ with X_t the finite set of possible input Tr-Symbols, Y_t the finite set of possible output Tr-Symbols, A_t the finite set of possible Tr-states and $F_t : X_t \times A_t \to Y_t^* \times A_t$ the transition function where * is the Kleene operator.

A Turing machine is a tuple $M = \langle I, K, G \rangle$ with I the finite set of Tu-Symbols that may occur on the tape, K the finite set of Tu-states that the read/write device may assume and $G : I \times K \to I \times K \times \{right, left\}$ the transition function, where the last factor is the set of allowed directions for the read/write device relative to the tape.

The next section will describe the type of Communication network that implements the Turing machine design in Section 2.4, then show that it truly works. The following variables are used:

- $n \in \mathbb{N}$ indexes over time;
- $p \in \mathbb{Z}$ indexes over the positions on the tape, assuming that $p - 1$ refers to the cell left of p;
- i and j index over the Tu-symbols in I, with an individual element denoted SYMBOL_i (without capitalization) or Symbol_i (with capitalization) depending on whether a Tr-Symbol or a Tr-state is concerned;
- k and l index over the Tu-states in K, with individual elements denoted as State_k;

- $x_n \in X_t$ is the nth input Symbol to the Transducer t under discussion;
- $y_n \in Y_t^*$ is the nth sequence of output Symbols of the Transducer t under discussion;
- α_n is the nth state of a Transducer.

Appendix B

Proof of Turing completeness

In this section we start out by formally describing the general structure of a Communication network W_{Turing} that implements a classical Turing machine $M = \langle I, K, G \rangle$. We do so by specifying the components of W_{Turing}: T_{Turing} (the Transducers) and C_{Turing} (the Channels). After that, it will be straightforward to verify that W_{Turing} does indeed fully model M.

Transducers

Recall that we labeled three types of Transducers in Figure 2.5: a. device Transducer, b. switch Transducers and c. memory Transducers. Templated component sets and transition functions for each of these types of Transducers are listed below. Note that in this subsection, MOVE serves as a template stand-in for the Tr-Symbols LEFT and RIGHT. Also note that for the sake of brevity no-ops are not shown in the transition tables; i.e. those rows where $\alpha_n = \alpha_{n+1}$ and y_n is empty.

The device Transducer $t_{\text{dev}} \in T_{\text{Turing}}$ has the following components:

- $X_{\text{dev}} = I$;
- $Y_{\text{dev}} = I \cup \{\text{LEFT}, \text{RIGTH}\}$;
- $A_{\text{dev}} = K$;
- see Table B.1 for F_{dev}. The second row is a special case of the first, which corresponds to cases where M switches to the *Halt* state.

Table B.1: Transition function of device Transducer t_{dev}.

x_n	α_n	y_n	α_{n+1}
SYMBOL$_i$	STATE$_k$	SYMBOL$_j$ MOVE	STATE$_l$
SYMBOL$_i$	STATE$_k$	SYMBOL$_j$	HALT

A switch Transducer $t_{\text{switch},p} \in T_{\text{Turing}}$ corresponding to tape position p has the following components:

- $X_{\text{switch},p} = I \cup \{\textsc{left}, \textsc{ricth}, \textsc{wake-left}, \textsc{wake-right}, \textsc{read}\}$;
- $Y_{\text{switch},p} = I \cup \{\textsc{wake-left}, \textsc{wake-right}, \textsc{read}\}$;
- $A_{\text{switch},p} = \{\textsc{Active}, \textsc{AtRight}, \textsc{AtLeft}, \textsc{Inactive}\}$;
- see Table B.2 for $F_{\text{switch},p}$.

Table B.2: Transition function of a switch Transducer $t_{\text{switch},p}$.

x_n	α_n	y_n	α_{n+1}
SYMBOL$_i$	ACTIVE	SYMBOL$_i$	ACTIVE
LEFT	ACTIVE		ATRIGHT
LEFT	ATLEFT	READ WAKE-LEFT	ACTIVE
LEFT	ATRIGHT		INACTIVE
RIGHT	ACTIVE		ATLEFT
RIGHT	ATLEFT		INACTIVE
RIGHT	ATRIGHT	READ WAKE-RIGHT	ACTIVE
WAKE-LEFT	INACTIVE		ATLEFT
WAKE-RIGHT	INACTIVE		ATRIGHT

A memory Transducer $t_{\text{mem},p} \in T_{\text{Turing}}$, corresponding to tape position p and partner of $t_{\text{switch},p}$, has the following components:

- $X_{\text{mem},p} = I \cup \{\textsc{wake-left}, \textsc{wake-right}, \textsc{read}\}$;
- $Y_{\text{mem},p} = I$;
- $A_{\text{mem},p} = I$;
- see Table B.3 for $F_{\text{mem},p}$.

Table B.3: Transition function of a memory Transducer $t_{\text{mem},p}$.

x_n	α_n	y_n	α_{n+1}
READ	SYMBOL$_i$	SYMBOL$_i$	SYMBOL$_i$
SYMBOL$_j$	SYMBOL$_i$		SYMBOL$_j$

Channels

C_{Turing} contains c_{read}, c_{write} and an unbounded number of $c_{\text{switch},p}$, respectively corresponding to the labels d, e, and f in Figure 2.5. We will now define these Channels in turn.
The reading Channel c_{read} has

- $S_{\text{read}} = \{t_{\text{mem},p} \mid p \in \mathbb{Z}\}$;
- $R_{\text{read}} = \{t_{\text{dev}}\}$.

The writing Channel c_{write} has

- $S_{\text{write}} = \{t_{\text{dev}}\}$;
- $R_{\text{write}} = \{t_{\text{switch},p} \mid p \in \mathbb{Z}\}$.

For any given p the private output Channel of $t_{\text{switch},p}$, $c_{\text{switch},p}$, has

- $S_{\text{switch},p} = \{t_{\text{switch},p}\}$;
- $R_{\text{switch},p} = \{t_{\text{switch},p-1}, t_{\text{switch},p+1}, t_{\text{mem},p}\}$.

Verification

We can now demonstrate how our design for W_{Turing} implements M. We do this by following two transitions of M, a normal transition and one that leads to the *Halt* state, as they are modeled by the transitions of the Transducers in T_{Turing}. The transitions of M are shown in Table B.4 and the complex transitions and interactions of W_{Turing} that model them are shown in Table B.5. In both tables, time is represented on the vertical axis and flows downwards. Colours and dashed horizontal lines are used to highlight matching parts between the tables. For the normal transition a Tu-state change and a move to the right is assumed and in both transitions the Tu-Symbol that was read from the tape is overwritten; the other possible variations are easily derived in the same manner. This is left as an exercise to the reader.

Because of space considerations, only Transducers that are directly involved in the transitions of M are shown in Table B.5. Two Transducers that are not directly involved and are therefore not shown do change state during the process, however; these are $t_{\text{switch},p-1}$ and $t_{\text{switch},p+2}$, switching respectively from AtLeft to Inactive at $n = 5$ and from Inactive to AtRight at $n = 6$.

Table B.4: The transitions of M that are modeled by the process in Table B.5.

read	state	write	new state	move
i	*k*	*j*	*l*	*right*
i	*l*	*j*	*Halt*	

Equivalence

A Turing machine with two tapes A, B can elegantly model a Transducer in the following way. Initially, A contains the input Signal, the read/write device is positioned at the cell with the first Symbol and B is blank. A single transition of the modeled Transducer that produces a series of m output Symbols may then be simulated by the Turing machine in one or multiple steps:

Table B.5: The process in W_{Turing} that models the transitions in Table B.4.

	$t_{\text{mem},p}$			$t_{\text{mem},p+1}$			t_{dev}		
n	x_n	α_n	y_n	x_n	α_n	y_n	x_n	α_n	y_n
1	READ	Symbol$_i$	SYMBOL$_i$		Symbol$_i$			State$_k$	
2		Symbol$_i$			Symbol$_i$		SYMBOL$_i$	State$_k$	SYMBOL$_j$ RIGHT
3		Symbol$_i$			Symbol$_i$			State$_l$	
4	SYMBOL$_j$	Symbol$_i$			Symbol$_i$			State$_l$	
5		Symbol$_j$		READ	Symbol$_i$	SYMBOL$_i$		State$_l$	
6		Symbol$_j$		WAKE-RIGHT	Symbol$_i$		SYMBOL$_i$	State$_l$	SYMBOL$_j$
7		Symbol$_j$			Symbol$_i$			Halt	
8		Symbol$_j$		SYMBOL$_j$	Symbol$_i$			Halt	
9		Symbol$_j$			Symbol$_j$			Halt	

	$t_{\text{switch},p}$			$t_{\text{switch},p+1}$		
n	x_n	α_n	y_n	x_n	α_n	y_n
1		Active			AtRight	
2		Active			AtRight	
3	SYMBOL$_j$	Active	SYMBOL$_j$	SYMBOL$_j$	AtRight	
4	RIGHT	Active		RIGHT	AtRight	READ WAKE-RIGHT
5	READ	AtLeft			Active	
6	WAKE-RIGHT	AtLeft			Active	
7	SYMBOL$_j$	AtLeft		SYMBOL$_j$	Active	SYMBOL$_j$
8	SYMBOL$_j$	AtLeft			Active	
9		AtLeft			Active	

1. With initial state α_n, the device reads input Symbol x_n from its current position on A, moves one cell up on A and switches to state $\alpha_{n+1,1}$. If $m > 0$ the device also writes $y_{n,1}$ to its current position on B and moves one cell up on B.

2…m until all output Symbols up to $y_{n,m}$ have been written to B and state $\alpha_{n+1,m}$ is reached, the device keeps transitioning upwards on B, without moving further on A.

The final state $\alpha_{n+1,m}$ becomes the initial state for the next simulated Transducer transition. The Turing machine repeats this procedure until all input Symbols on A have been processed. The resulting output sequence on B will be identical to the output Signal of the simulated Transducer.

Turing machines with two tapes and Turing machines with a single tape have been proven to be equivalent [20, pp. 161–163]. It follows that a classical Turing machine with a single tape can also simulate a Transducer. Since we just proved that (cyclic) Transducers can simulate a Turing machine, we now have a complete proof that Transducers are Turing-equivalent.

Bibliography

[1] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948.

[2] M. Unser, "Sampling—50 Years after Shannon," *Proceedings of the IEEE*, vol. 88, pp. 569–587, April 2000.

[3] R. Hermsen, "On the physics of computability and the computability of physics," Master's thesis, Utrecht University, June 2003.

[4] R. C. Cummins and P. Poirier, "Representation and indication," in *Representation in Mind. New Approaches to Mental Representation* (P. S. H. Clapin and P. Slezak, eds.), pp. 21–40, Elsevier, 2004.

[5] T. Crane, *The Mechanical Mind: A Philosophical Introduction to Minds, Machines and Mental Representation*. Routledge, second ed., 2003.

[6] C. Houtekamer, "What about it? An essay on intentionality, perception and nonconceptual content," Master's thesis, Utrecht University, August 2005.

[7] R. G. Millikan, "Teleosemantic theories," in *Varieties of Meaning*, Jean-Nicod Lectures, ch. 5, Paris: Institut Jean-Nicod, 2002.

[8] M. David, "The correspondence theory of truth," in *The Stanford Encyclopedia of Philosophy* (E. N. Zalta, ed.), fall 2009 ed., 2009.

[9] A. Turing, "On computable numbers, with an application to the entscheidungs problem," *Proceedings of the London Mathematical Society*, vol. 42, no. 2, pp. 230–265, 1937.

[10] J. van Dis and S. Muijs, *Syllabus bij het college Grondslagen van de Thermische en Statistische Fysica*. Grondslagen van de Thermische en Statistische Fysica, Utrecht University, third ed., 1999.

[11] M. Mevius, "Entropie en informatie. Het duiveltje van Maxwell en andere verhalen," Master's thesis, Utrecht University, 1997.

[12] C. G. Timpson, “Philosophical aspects of quantum information theory,” in *The Ashgate Companion to the New Philosophy of Physics* (D. Rickles, ed.), Ashgate, 2007.

[13] P. E. Griffiths, “Genetic information: A metaphor in search of a theory,” *Philosophy of Science*, 2001.

[14] P. Godfrey-Smith and K. Sterelny, “Biological information,” in *The Stanford Encyclopedia of Philosophy* (E. N. Zalta, ed.), fall 2008 ed., 2008.

[15] K. Sterelny, “The “genetic program” program: A commentary on Maynard Smith on information in biology,” *Philosophy of Science*, vol. 67, no. 2, pp. 195–201, 2000.

[16] J. Maynard Smith, “The concept of information in biology,” *Philosophy of Science*, vol. 67, no. 2, pp. 177–194, 2000.

[17] P. Godfrey-Smith, “On the theoretical role of “genetic coding”,” *Philosophy of Science*, vol. 67, no. 1, pp. 26–44, 2000.

[18] M. de Pinedo-Garcia and J. Noble, “Beyond persons: extending the personal/subpersonal distinction to non-rational animals and artificial agents,” *Biology & Philosophy*, vol. 23, pp. 87–100, January 2008.

[19] C. Hewitt, P. Bishop, and R. Steiger, “A universal modular actor formalism for artificial intelligence,” in *Proceedings of the 3rd International Joint Conference on Artificial Intelligence*, IJCAI’73, (San Francisco, CA, USA), pp. 235–245, Morgan Kaufmann Publishers Inc., 1973.

[20] J. Hopcroft and J. Ullman, *Introduction to automata theory, languages, and computation*. Addison-Wesley series in computer science, Addison-Wesley, 1979.