[This is adapted from a section of my dissertation. I have tried to patch it up so that it is self-contained, but I doubt I have succeeded completely. Please let me know if there is anything that requires further elaboration. - Tarun]

Agents are information processing systems. Landauer's principle [Landauer1961] describes certain thermodynamic constraints on information processing. Some physicists claim that these constraints engender an observation selection effect that can form the basis for an anthropic explanation of the asymmetry of our environment.[1] The precise content of the principle is a matter of some subtlety, as we shall see, but here is the usual gloss: the physical instantiation of a logically irreversible computational process will be thermodynamically irreversible. Any physically plausible agent of sufficient cognitive complexity should be expected to implement some logically irreversible processes, so agents will exhibit a thermodynamic arrow. This asymmetry in agents might then produce a selection bias that at least partially explains the asymmetry of our environment. While there are instructive aspects to this approach, it is ultimately problematic. The constraints set by Landauer's principle are not strong enough to perform the explanatory task at hand. To see why this is so, let us begin by examining the principle in more detail.

A logically irreversible operation is one that maps multiple distinct input states to the same output state, so the input cannot be fully recovered just by knowing the output. An example is the AND operation, which I represent as acting on two binary digits and producing one digit as output:

Input	**Output**
00	0
01	0
10	0
11	1

Clearly, three distinct input states are mapped to the same output state. If the output is 0 we cannot determine the input without further information. Now consider the operation where we perform AND but also copy the input bits and concatenate them to the output:

Input	**Output**
00	000
01	001
10	010
11	111

This operation is logically reversible, and the AND operation can be recovered from it by just looking at the first bit of the output. In principle, any computational

[1] See [Hawking1994], [Schulman2005] and [Hartle2005] for arguments of this sort.

procedure involving logically irreversible steps can also be implemented using only logically reversible steps. We just rewrite each logically irreversible step so that the inputs are retained in memory. Unfortunately, for a complex computational process this can make quite significant demands on the available memory. Any physically instantiated computation will have to work with finite memory, so after a certain point a logically irreversible operation will be required. One cannot keep writing new data into memory without overwriting existing data. Landauer considers the simplest irreducible process, resetting the state of a memory register to a given value no matter what its initial value:

Input	**Output**
0	0
1	0

This is a logical many-to-one transition. Landauer argues that any instantiation of it would have to involve a *physical* many-to-one transition. The phase space of the system instantiating this computation would have to have two disjoint regions corresponding to the register being in either the 0 or the 1 state. Prior to the reset, the system could be anywhere in these two regions, but after the operation it can only be in the 0 region. There will be a compression of phase space along this degree of freedom, which Landauer calls an information bearing degree of freedom (IBDF). Liouville's theorem tells us that this compression must be compensated by an expansion of accessible phase space along the system's other degrees of freedom. In a reliable computer, this expansion should not affect other IBDF. We don't want a reset operation performed on one register opening up new possible states for other registers. The compensating expansion of phase space will affect non-information bearing degrees of freedom (NIBDF). In other words, we have a transfer of fine-grained entropy from the IBDF to the NIBDF, since fine-grained entropy is conserved.

So far we have not assumed that the second law of thermodynamics governs the evolution of this system. If it does, then the decrease in entropy[2] along the IBDF must be either matched or surpassed by an increase in entropy of the NIBDF, so that the total entropy of the system either stays the same or increases. Since the NIBDF are usually environmental degrees of freedom, logically irreversible processes will be accompanied by dissipation of heat into the environment. A simple one-molecule memory device is often used to illustrate this fact. Consider a box with a removable partition separating it into two regions of equal volume, left (L) and right (R), corresponding to the two distinct states of the memory device. The walls of the box are movable pistons that can be used to compress the volume of the box. There is a single molecule in the box, and the box is in contact with a heat bath. If the molecule

[2] When I use the term "entropy" without qualification in a Gibbsian context, I am referring to the coarse-grained entropy.

is in region L the device is recording a 0, and if it is in region R the device is recording a 1.

In order to perform the reset operation, we need to ensure that the molecule ends up in the L region, no matter where it starts. Here is how this might be done. First remove the central partition. Then push the right wall in isothermally until the volume of the box has been halved. This ensures that the molecule is in region L. Now reinsert the partition, trapping the molecule in this region, and move the right wall back to its original position. The reset operation is complete.

The standard calculation of entropy change during this process proceeds as follows[3]: Assume that we start with a canonically distributed ensemble of memory devices, half of which are in state 0 and the other half in state 1. This represents initially random or unknown data. After the device has been reset, we have a canonically distributed ensemble of systems that are all in state 0 (region L). The canonical distribution over phase space is given by

$$\rho(q,p) = \exp(-E(q,p)/kT)/Z$$

where E is an energy function, T is the temperature of the system, k is Boltzmann's constant and Z is the partition function. The partition function is given by

$$Z = \int_{\Gamma} \exp(-E(q,p)/kT)\,dq\,dp$$

where Γ is the region of phase space accessible to the system. The Gibbs entropy of the system is

$$S = -k\int_{\Gamma} \rho(q,p)\ln\rho(q,p)\,dq\,dp\,.$$

Let Γ_L be the phase space region corresponding to the particle being in region L of the box and Γ_R be the phase space region corresponding to the particle being in region R. Let Γ_B be the union of these two phase space regions. We have assumed that initially all of Γ_B is accessible to the system, but after the reset operation the system is restricted to Γ_L. The energy function changes after the operation, since part of the phase space that was accessible is no longer accessible. But we assume that the energy function does not change in the region that remains accessible, Γ_L. Furthermore, we assume that the initial energy function (prior to the reset operation) is the same in regions Γ_L and Γ_R. The two regions are initially symmetric. With this apparatus in place, the entropy of the final state is

[3] A version of this calculation is presented (among many other places) in the introduction to [LeffRex2003].

$$S_f = -k \int_{\Gamma_L} \rho_f \ln \rho_f \, dq \, dp$$

where ρ_f is the canonical distribution over the final ensemble. The entropy of the initial state is

$$S_i = -k \int_{\Gamma_B} \rho_i \ln \rho_i \, dq \, dp$$

where ρ_i is the canonical distribution over the final ensemble. Because of the symmetries between Γ_L and Γ_R, the partition function for the initial ensemble is simply twice the partition function of the final ensemble, so that $\rho_i = \rho_f/2$ in the region Γ_L where both distributions are non-vanishing. This gives us

$$S_i = -k2 \int_{\Gamma_L} \frac{\rho_f}{2} \ln \frac{\rho_f}{2} \, dq \, dp = -k \int_{\Gamma_L} \rho_f \ln \rho_f + k \ln 2 \int_{\Gamma_L} \rho_f \, dq \, dp$$
$$= S_f + k \ln 2 \, .$$

The initial entropy of the memory device is greater than the final entropy by k ln2. This means the entropy of the environment must increase by at least this much. We have Landauer's principle in its most precise form: Resetting a single bit increases the entropy of the NIBDF by at least k ln2.

There are a number of points worth noting about this argument. The first is that the argument assumes that the second law of thermodynamics governs the system under consideration.[4] This is what allows us to infer an entropy increase in the NIBDF from an entropy decrease in the IBDF. This places significant restrictions on the form of any anthropic argument that relies solely on Landauer's principle. It rules out region selection arguments. One might have hoped that the asymmetry involved in agency that the principle highlights could somehow be used to explain why agents must live in asymmetric environments. But we have seen that Landauer's principle only tells us that agents are asymmetric *if* they live in an asymmetric environment. Any region selection argument based on the principle will be circular.

Things look more promising when we turn to direction selection arguments. The principle might be able tell us why agents in a second law governed environment must all "point the same way". An agent in such an environment will only be able to perform logically irreversible operations in the direction of increasing entropy. This

[4] This point is made forcefully in [EarmanNorton1999]. The authors criticize attempts to rule out Maxwell's demons (and thus vindicate the second law) on the basis of Landauer's principle. They argue that any such attempt will be question-begging, since the principle presupposes the truth of the second law.

in itself doesn't tell us much, but perhaps it might be leveraged into an argument that an agent's subjective future must be in the direction of increasing entropy. Indeed, those who look to Landauer's principle for an anthropic explanation tend to focus on direction selection. For instance, here's Stephen Hawking:

> "[W]hen a computer records an item in memory, the total amount of disorder in the Universe increases. The direction of time in which a computer remembers the past is the same as that in which disorder increases." [5] [Hawking1987, p. 47]

And here's James Hartle:

> "If [a] robot processes information irreversibly, then its psychological arrow of time must generally be congruent with the thermodynamic arrow of time… Entropy increase is necessary… for the erasure of a record." [Hartle2005, p. 106]

These direction selection arguments are based on the claim that information processing systems must be thermodynamically irreversible. But does Landauer's principle really license this claim? Landauer's argument only tells us that there must be an increase in the entropy of the environment greater than *or equal to* the decrease in the entropy of the memory device. If the entropies gained and lost balance each other so that there is no net increase in entropy the process may well be reversible. And in fact the molecule-in-a-box memory device we considered need not involve any change in total entropy. The initial removal of the central partition does not change the entropy of the initial ensemble, and the compression of the box by moving the right wall can be implemented completely reversibly, so that the entropy gained by the environment is equal to the entropy lost by the system. If Landauer's principle does not guarantee us irreversibility, what's all the fuss about?

Our toy system doesn't experience any change in entropy during the first step where the partition is removed. This is because we assumed that our initial ensemble, before the removal of the partition, had half the devices in the 0 state and the other half in the 1 state. This ensemble does not change when the partition is removed. However, if our initial ensemble is skewed in any way, with slightly more devices in one state than in the other, then the removal of the partition will lead to an increase in the total entropy of the system, and the system will be genuinely irreversible. This means that the reset process can only be implemented reversibly if the initial data being reset is random or completely unknown. If we have some idea about the initial state of the system – say we believe there is a 60% chance it is in state 0 – this will skew the initial probability distribution. When the partition is removed, the

[5] Hawking mistakenly attributes entropy increase to the process of recording an item to memory, rather than erasing an item from memory. Charles Bennett has demonstrated that measurement (or recording) can be done in a thermodynamically reversible manner [Bennett1982].

distribution will flatten out, leading to an increase in total entropy. In the limiting case, where we know the initial state of the device, removing the partition will lead to an entropy increase of k ln2. So when a computing device erases known data, the process will be thermodynamically irreversible. The entropy increase in the NIBDF will not be compensated by any entropy decrease in the IBDF.

John Norton argues that the preceding analysis of Landauer's principle is confused because it is based on an illicit choice of ensemble [Norton2005]. Recall that we started with an ensemble of systems in both 0 and 1 states. However, our actual system is in either the 0 state or the 1 state. This means that in the actual phase space of the system either region Γ_L is accessible or Γ_R is accessible but not both. We may not know which one of these regions is accessible *but we know that it is only one of them*. But by choosing an ensemble that contains both 0 and 1 states we are working in a phase space where both regions are accessible. This corresponds to an energy function that we know for sure doesn't accurately describe the system. In fact, this would be the canonical ensemble for the box *without* the partition in the center. Needless to say, this is a different system. A proper canonical ensemble for our system would only consist of devices in the same memory state as the system we are studying, irrespective of our epistemic limitations. This eliminates any dependence of the probability distribution on our state of knowledge about the initial data.

One way around this objection is to interpret statistical mechanical ensembles as probability distributions that codify an observer's ignorance about the system (or, equivalently, the observer's information about the system). On this view, the canonical ensemble is a special case of distribution applied over Γ_L (or over Γ_R) is the distribution that best represents our uncertainty about which of the many configurations within this region the system instantiates. But this does not capture all of our uncertainty about the system, so to reason using one of these distributions would be to reason as if we had information that we do not in fact have. Norton is right that the distribution we started with is not a canonical distribution. If interpreted as one, it seems clearly inappropriate; it posits an energy function we know does not obtain. But of course we call that parameter in the distribution an "energy function" only if we believe the distribution is canonical. The problem is with the interpretation, not the distribution.

So how should we interpret the distribution we use? Let ρ_L represent the canonical distribution over phase space region Γ_L and ρ_R represent the canonical distribution over Γ_R. These distributions represent our ignorance about the configuration *within* these regions.[6] But we also need to represent our uncertainty *between* these regions. If I have no information about which of these two regions the system is actually in, I should use the distribution

[6] The canonical distribution is the maximum entropy distribution constrained by a given mean energy for the system. If the mean energy of the system is all that is known, this is the distribution that best represents this epistemic state.

$$\rho = \frac{1}{2}\rho_L + \frac{1}{2}\rho_R .$$

Plugging this distribution into the Gibbs entropy formula, we get

$$S = -k\int_{\Gamma_B} \rho \ln \rho \, dq\, dp = -\frac{k}{2}\int_{\Gamma_B} (\rho_L + \rho_R)\left[\ln(\rho_L + \rho_R) - \ln 2\right] dq\, dp .$$

Using the fact that ρ_L and ρ_R have disjoint support, we can show that

$$S = -\frac{k}{2}\int_{\Gamma_L} \rho_L \ln \rho_L \, dq\, dp - \frac{k}{2}\int_{\Gamma_R} \rho_R \ln \rho_R \, dq\, dp + k \ln 2$$

Let S_C be the entropy of a canonically distributed molecule-in-a-box system restricted to either Γ_L or Γ_R.[7] Then the entropy of the system under consideration reduces to

$$S = S_C + k \ln 2$$

But this is precisely what we had earlier. Complete lack of knowledge about the data stored in the device adds k ln 2 to the entropy. The earlier analysis was foundationally sloppy, relying on an unjustified choice of ensemble, but we have seen that a more careful analysis leads to the same result as long as we recognize that statistical mechanical probability distributions are relative to the available information about the system.

Norton responds that we must stick to the canonical ensemble because we are dealing with a thermodynamic phenomenon, and the canonical ensemble is what gives us the thermodynamic entropy. Perhaps there is some other quantity, call it the *augmented entropy*, that reflects a particular observer's uncertainty about the state of the system, but why would this quantity be the one we use to determine how much heat is dissipated by the system? Thermodynamic entropy is a function of the state of the system. If we use the notion of entropy I have suggested above, then two observers will legitimately assign different entropies to the same memory device depending on whether or not they know the data stored in the device. This entropy, then, is not a function of the state of the memory device, so it cannot be the thermodynamic entropy.

Norton is correct that we cannot use a notion of entropy that is relative to the probability distribution over states if this distribution can be fixed willy-nilly. And

---

[7] The entropy would be the same for canonical distributions over both these regions, since they are completely symmetric.

indeed, if we take a completely subjective approach to the epistemic probabilities involved in the argument, they will be fixed more or less willy-nilly. However, one can interpret the probabilities in statistical mechanics as epistemic, even agent-relative in a certain sense, without being committed to this problematic subjectivity. An example of this approach is the *Jaynesian* or *objective Bayesian* tradition. Like the subjective approach, the probability distribution describing the thermodynamic properties of the system isn't a feature of the system alone; it captures something about the relationship between an observer and the system. However, unlike the subjective approach, the relevant relationship does not depend on the idiosyncratic epistemic state of the individual observer. It is based on the broad epistemic and causal capacities that the observer shares with all other similarly structured information gathering and utilizing systems.

Return to the notion that the thermodynamic entropy is a function of state. The "state" here is individuated by a set of independent thermodynamic variables, and these variables are picked based on the ways in which work can be done on the system. A method of doing work is given by a generalized force (an intensive variable) and the associated generalized displacement (an extensive variable). For instance, one might be able to perform work on a system by changing its volume, a generalized displacement, and the generalized force against which this work is done is the pressure of the system.[8] Specifying one member of each force-displacement pair is, in conjunction with the internal energy of the system, sufficient to fix the system's thermodynamic state.

What is important here is that the thermodynamic state space depends on the ways in which work can be done on the system, and these are just the ways in which energy transfer to and from the system can be *controlled*. The thermodynamic coordinates are macrovariables that can be directly manipulated. Any energy transfer through degrees of freedom that cannot be controlled is heat. But this emphasis on control should suggest a certain agent-relativity. In particular, theoretical agents with more refined epistemic and causal capacities might be able to extract more work from a system. As a simple example, consider a system consisting of a gas in thermal equilibrium with a paramagnet. An agent only able to manipulate the volume of the gas will in general be able to extract less useful work from the system than an agent who, in addition to the volume, can also manipulate and exploit the magnet's magnetization. The former's thermodynamic state space will be a projection of the latter's three-dimensional space onto the two-dimensional energy-volume subspace.

Built into the thermodynamic state description of a system, then, is a particular conception of the agent interacting with the system. This means the thermodynamic entropy can be a function of thermodynamic state while still being agent-relative to a degree. The statistical mechanical entropy inherits this agent-relativity because

---

[8] Other examples of force-displacement pairs are chemical potential and particle number, and magnetic field and magnetization. See [Kardar2007] for more.

the partitioning of phase space into macrostates is determined by the relevant thermodynamic variables. These macrostates delineate sets of microstates that are indistinguishable to the agent. There is an important ambiguity here, however. So far I have been saying that the thermodynamic variables depend on the agent's epistemic and causal capacities, without distinguishing between these two sets of capacities. However, there is no *a priori* reason that these capacities should pick out the same macrostates.

In order to distinguish between these notions, consider an agent confronted with a decision problem about how to intervene on a system based on information gathered about the system. The epistemic capacities of the agent determine the amount of information it can gather about the state of the system. We can define a partition of phase space into macroregions such that microstates *a* and *b* are in the same macroregion iff the agent cannot distinguish whether the system is in state *a* or *b*. Call this the *epistemic partition*, and call the regions *epistemic macrostates*. Jaynes' maximum entropy principle, a version of the principle of indifference, applies to the epistemic partition. The probability distribution that best represents the agent's epistemic state will be the maximum entropy distribution over the epistemic macrostate containing the system. Intuitively, this is the distribution that assumes the least about the microstate of the system beyond what is already known by fixing the macrostate.

However, epistemic accessibility does not fully capture what is special about thermodynamic coordinates. The point is that we can often exploit our knowledge to extract work from the system. This depends on our ability to intervene on the system, not just record its state. Consider the one-molecule system discussed above, and assume we know which side of the box the molecule is in. If the molecule is in region L, I bring the right wall all the way up to the partition before removing it. I now allow the "gas" to expand isothermally, doing work as it pushes the wall back to its original position. If the molecule were in region R, I would perform a different intervention on the system, this time pushing in the left wall. If I could not know which state the molecule was in, I would not be able to extract work out of the system in the long run, because I would not be able to use the state of the system as a guide to the interventions I perform.

It is this feature of the macrostates, their connection with differential intervention allowing for the extraction of work, that connects them to the thermodynamic entropy. States of the system must be distinguishable for us to exploit them in this way, but the converse is not necessarily true. It may be the case that we are able to distinguish between two states of a system but are unable to exploit this distinction because our causal capacities are not sufficiently fine-grained. A simple example is when we know the histories of the system. Suppose I drop ink into two glasses of water in two different places and watch as the ink spreads. Soon the observable states of the glasses will be indistinguishable. But I have access to more than just the current observable state; I know the history of the two glasses. This allows me to distinguish the two glasses. I know that glass A is in a state that comes from a drop in the center, say, and glass B is in a state that comes from a drop on the periphery. However, I cannot exploit this knowledge. There aren't meaningfully distinct interventions I can perform on the systems based on their difference. This is the justification for coarse-graining. My fine-grained knowledge of the history of the system is pragmatically irrelevant, so I don't consider it when reasoning about the system.

To summarize, thermodynamics is a pragmatic science, a science concerned with our ability to control systems in order to channel energy flow. Statistical mechanical epistemic macrostates track this motivation pretty well, since knowledge about the system can often be exploited. However, it would be a mistake to think that the entropy calculated using epistemic macrostates is always the appropriate analog to the thermodynamic entropy. This is only the case when the knowledge is causally productive, when it can form the basis for meaningful intervention. If the knowledge is irrelevant to how one might efficiently interact with the system, then it should be ignored when reasoning about the thermodynamics of the system. One might distinguish between epistemic macrostates, defined above, and *causal macrostates*, which often coincide with their epistemic counterparts but may be more coarse-grained, since they ignore pragmatically irrelevant epistemic distinctions.

Returning to our discussion of Landauer's principle, it might seem that my digression has served no purpose. In our example, the Γ_L and Γ_R regions correspond to distinct epistemic macrostates, at least when the device is storing data. Furthermore, the distinction between the states clearly isn't pragmatically otiose. As I argued above, knowledge of the memory state allows for the extraction of work. More generally, the computational system might follow different computational paths depending on the state of the memory device, so knowledge of the state allows prediction of the macroscopic trajectory of the computational system. It seems that Γ_L and Γ_R must be distinct causal macrostates as well.

If this is right, then my earlier response to Norton cannot work. We can't draw thermodynamic conclusions from any old probability distribution. It must be the maximum entropy distribution over the causal macrostate of the system. And in this case, that would just be the canonical distribution over either Γ_L or Γ_R, but *not* over their union. I applied a probability distribution over both regions, claiming that this distribution represented the epistemic situation of an observer who lacks information about the initial data stored on the device. Norton's response was that the entropy of this distribution does not correspond to the thermodynamic entropy. My digression into the agent-relativity of thermodynamic entropy was supposed to dent Norton's criticism, but now it may seem that it has done no such thing.

Fortunately for me, the argument sketched above is incorrect. When we are considering data that will be erased, the epistemic distinction between Γ_L and Γ_R is pragmatically irrelevant. In order to see this, we need a fuller conception of what genuine information erasure requires. Landauer describes it as a simple logical operation, the erasure of data from a single memory register. But this is clearly insufficent. As Landauer himself recognized, the erasure of data from a particular memory register is reversible if the data is simultaneously copied to another register. This just amounts to the reversible logical operation

Input	**Output**
0	00
1	01

There must be a requirement that there is no trace of the erased information left in any of the IBDF. But the requirement must be stronger than this if we have true erasure. For instance, consider the situation where the computational path of the cognitive system branches based on the pre-erasure state of the register. If this branching persists after the erasure of the register, we still do not have genuine erasure. The system obviously has access to its computational path, and so it has access to a record of the memory state. The computational path should be regarded as an IBDF. The finitude of memory argument that disallows the indefinite operation of a computer without erasure can be supplemented by a finitude of computational states argument ruling out indefinite branching of the computational path. For genuine erasure, branched computational paths must merge. We can extend this to any environmental traces of the memory state that are accessible to the cognitive system (such as, say, raised weights). Any such trace is only cosmetically different from an internal IBDF. Once the proponents of Landauer's principle acknowledge that if a trace is left in the system's memory, the process of erasure need not be thermodynamically irreversible, they commit themselves to very thorough conception of erasure.

Whatever physical process a cognitive system uses to implement erasure cannot be a function of the initial data on the memory device. Here it is crucial that it is the cognitive system itself, not some external force, that is implementing the operation. We can make the intuitive assumption that in order for the system to implement different physical processes it must go through different sequences of computational states. This means that if the physical erasure process is a (non-trivial) function of the initial data, there will be branching of the computational path of the system, leaving a trace of the initial data accessible to the system. True erasure would only be accomplished when the paths merge, and this merging would not, of course, depend on the initial data. Considerations of this sort suggest that the epistemic distinguishability of the memory states is pragmatically irrelevant to their erasure.

Even though we may know which state the system is in prior to erasure, this information is irrelevant to the implementation of erasure. This licenses us to ignore the distinction between these states in our thermodynamic description of the procedure. The relevant causal macrostate for this procedure encompasses both Γ_L and Γ_R, and the appropriate probability distribution is the maximum entropy distribution over this macrostate. After erasure, Γ_R is no longer accessible, so erasure does require a compression of the phase space of the IBDF, and (by Liouville's theorem) a corresponding expansion of the phase space along other degrees of freedom.

While the approach described here deals with Norton's objections, it does not establish the truth of Landauer's principle. In particular, it does not establish that erasure will be thermodynamically irreversible. In fact, it suggests that one argument for this irreversibility is fallacious. Recall that according to the standard argument the irreversibility is supposed to follow only if we have some information about the initial memory device, so that the ensemble of devices is "skewed". However, as Norton correctly points out, this is based either on an illicit choice of ensemble or an irrelevant notion of entropy. I have argued that the phase space reduction in the IBDF comes from *ignoring* our knowledge of the memory state. Perhaps it can be shown that the compensating increase in phase space volume must occur in the thermal degrees of freedom of the environment, thus saving Landauer's principle, but I do not see a reason why this would be the case. I suspect the details of whether or not erasure is dissipative will depend on the particular implementation of the process.

References

[Bennett1982] Bennett, C. H. "The thermodynamics of computation--a review."

[EarmanNorton1999] Earman, J.; Norton, J. D. "EXORCIST XIV: The Wrath of Maxwell's Demon. Part II. From Szilard to Landauer and Beyond."

[Hartle2005] Hartle, J. B. "The Physics of Now."

[Hawking1987] Hawking, S. W. *A Brief History of Time*

[Hawking1994] Hawking, S. W. "The Nature of Space and Time."

[Kardar2007] Kardar, M. *The Statistical Physics of Particles*

[Landauer1961] Landauer, R. "Irreversibility and heat generation in the computing process."

[LeffRex2003] Leff, H. S.; Rex, A. F. *Maxwell's Demon 2*

[Norton2005] Norton, J. D. "Eaters of the lotus: Landauer's principle and the return of Maxwell's demon."

[Schulman2005] Schulman, L. S. "A Computer's Arrow of Time."