The Use of the Information Theoretic Entropy in Thermodynamics

James Ladyman[1,*], Stuart Presnell[1], Anthony J. Short[2]

[1] Department of Philosophy, University of Bristol, 9 Woodland Road, Bristol, BS8 1TB

[2] H.H.Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol, BS8 1TL

[*] email: james.ladyman@bristol.ac.uk

October 11, 2006

Abstract

When considering controversial thermodynamic scenarios such as Maxwell's demon, it is often necessary to consider probabilistic mixtures of states. This raises the question of how, if at all, to assign entropy to them. The information-theoretic entropy is often used in such cases; however, no general proof of the soundness of doing so has been given, and indeed some arguments against doing so have been presented. We offer a general proof of the applicability of the information-theoretic entropy to probabilistic mixtures of macrostates, making clear the assumptions on which it depends, in particular a probabilistic version of the Kelvin statement of the Second Law. We briefly discuss the interpretation of our result.

1

1 Introduction

Suppose that we have two copies, A and B, of a thermodynamic system, but in distinct thermodynamic states Ω_A and Ω_B, with entropies S_A and S_B. Suppose also that there is a third copy, C, of the system, and that it is prepared in Ω_A or Ω_B with equal probability. The natural way to describe the state Ω_C is to say that it is a probabilistic mixture of Ω_A and Ω_B. Suppose we need to carry out a thermodynamic calculation involving C. What should we say about its entropy S_C? Consider the following three possible answers to this question:

1. We cannot proceed since S_C is either S_A or S_B but we do not know which.

2. S_C is the weighted average of S_A and S_B, namely $\frac{1}{2}(S_A + S_B)$.

3. S_C is the weighted average of S_A and S_B plus an additional term that represents a contribution to the thermodynamic entropy due to the probability distribution itself. In this particular case, $S_C = \frac{1}{2}(S_A + S_B) + k \ln 2$ (where k is Boltzmann's constant).

Which of these three answers is correct is the subject of controversy. For example, Norton (2005) criticizes those (see, for example, Leff and Rex (2003)) who defend the last answer. While we do not fully accept his critique we agree with him that so far no general argument for its correctness has been given. In what follows we offer a proof of it based on the Kelvin statement of the Second Law of Thermodynamics. This formulation has the advantage that it directly relates to practical and operational matters. Our result is in fact more general than answer 3 above since we consider a mixture of any (finite) number of states with an arbitrary probability distribution.

In particular, consider a system M in the state Ω_{mix}^M which is a probabilistic mixture of n states Ω_i^M with probability p_i, and where the different Ω_i^Ms may have different temperatures. We show that, if each Ω_i^M has a well-defined thermodynamic entropy S_i^M, the correct expression for the thermodynamic entropy S_{mix}^M of Ω_{mix}^M is:

$$S_{mix}^M = \sum_i p_i S_i - k \sum_i p_i \ln p_i \tag{1}$$

We refer to $-k \sum_i p_i \ln p_i$ as the information-theoretic entropy because it is a form of the Shannon entropy of Information Theory (although rescaled by $k \ln 2$ from its usual form).

We make the following assumptions:

1. The standard Kelvin formulation of the Second Law of Thermodynamics (see for example Uffink 2001) can be can be generalised to cover probabilistic processes as follows:

   > It is impossible to perform a cyclic process with no other result than that *on average* heat is absorbed from a reservoir, and work is performed.

   It is clear that if you could violate this generalisation you could, with arbitrarily high probability, violate the original Kelvin statement by performing the relevant cycle many times to generate a total quantity of work that is approximately proportional to the average work produced in a single cycle.

2. The n states Ω_i^M are all distinct thermodynamic states in the sense that they can be perfectly distinguished by a single measurement.

3. In keeping with standard practice in discussions of thermodynamics and the Second Law, we assume the legitimacy of the following idealisations about physical systems and operations:

   (a) There is a heat bath at temperature T_0.

   (b) There is a box with volume V containing a single molecule gas and a number of removable partitions, and the molecule can be treated thermodynamically as an ideal gas such that when it is confined to a region of volume V its entropy is given by $k \ln V$ plus a constant that depends on the temperature. We refer to this combined system as G (note that any system that had equivalent properties would do just as well).

   (c) The controlled operations we use to couple G and M can be performed in a thermodynamically reversible way.

   (d) The moving of pistons and partitions can be done frictionlessly.

2 Derivation of the Main Result

Consider a process involving the systems M and G. Heat flow from the reservoir will be written as positive ΔQ.

1. Initially M is in a standard (non-probabilistic) thermodynamic state Ω_0^M with entropy S_0^M, and G is in the thermal state with all partitions removed.

2. $n - 1$ partitions are inserted into the box so that the volume of the ith region of the box is $p_i V$. Let Ω_i^G (with entropy S_i^G) be the thermodynamic state in which the gas is confined to the ith region, where

4

$i = 1, ..., n$. The gas is now in a probabilistic mixture of n states Ω_i^G with probability p_i. There is no heat flow during this process, hence

$$\frac{< \Delta Q >}{T_0} = 0. \tag{2}$$

3. A controlled operation is performed from G onto M such that if G is in the state Ω_i^G, M is evolved reversibly from the state Ω_0^M into the state Ω_i^M which has entropy S_i^M. M is now in a probabilistic mixture of n states Ω_i^M with probability p_i. Note that if the process of transforming Ω_0^M to Ω_i^M involves a heat flow to or from the heat reservoir, this is done via a Carnot engine to ensure reversibility.

$$\frac{< \Delta Q >}{T_0} = \sum_i p_i (S_i^M - S_0^M). \tag{3}$$

4. A controlled operation is performed from M onto G such if M is in the state Ω_i^M, the partitions of G are all removed except those around region i, and then pistons are inserted from either end up to the remaining partitions. The last two partitions are then removed and the gas is allowed to expand isothermally, doing work against the pistons and inducing a heat flow into the box from the reservoir. When the gas is originally in the ith partition we have:

$$\frac{\Delta Q_i}{T_0} = \Delta S_i^G = k \ln V - k \ln p_i V = -k \ln p_i. \tag{4}$$

Hence,

$$\frac{< \Delta Q >}{T_0} = -k \sum_i p_i \ln p_i. \tag{5}$$

5. G is now in its initial state but M is now in the state Ω_{mix}^M, which is a probabilistic mixture of the states Ω_i^M with probability p_i. Every step of this process is reversible so the effect has been to transform the state of M from Ω_0^M to the state Ω_{mix}^M reversibly.

The average total heat flow in the process is given by

$$\frac{<\Delta Q_{rev}>}{T_0} = \sum_i p_i(S_i^M - S_0^M) - k\sum_i p_i \ln p_i \tag{6}$$

If there were another reversible process from the same inital state to the same final state, it would have the same value for $<\Delta Q_{rev}>$ (because otherwise by the correct choice of the order in which to combine them we could violate the probabilistic Kelvin statement of the Second Law of thermodynamics of assumption 1). Hence we can define a unique entropy difference between the two states as follows:

$$\Delta S = \frac{<\Delta Q_{rev}>}{T_0} \tag{7}$$

Similarly for an irreversible process from the same initial state to the same final state, the total heat flow into the system $<\Delta Q_{irr}>$ must be less than $<\Delta Q_{rev}>$ (because otherwise, again, by the correct choice of the order in which to combine them we could violate the probabilistic Kelvin statement of the Second Law of thermodynamics of assumption 1). Hence we recover a probabilistic analogue of the entropic form of the Second Law: For all processes, irreversible or otherwise, connecting the same initial and final states,

$$\Delta S \geq \frac{<\Delta Q>}{T_0}. \tag{8}$$

Finally from equations (??) and (??), and the fact that $\Delta S = S_{mix}^M - S_0^M$ we have, as claimed,

$$S_{mix}^M = \sum_i p_i S_i - k\sum_i p_i \ln p_i. \tag{9}$$

6

3 Discussion

We now offer a brief discussion of the interpretation of our result.

We imagine someone objecting that if what we have said is correct then the thermodynamic entropy cannot be an objective property of a system. This might be so if the probabilities p_i were merely subjective, but we take it that even in a deterministic universe it makes sense to talk about the objective chance of a fair coin landing heads. In other words, even purely epistemic probabilities may be objective.

What we mean by the objective probabilities is illustrated by the following case. Consider two parties, a mixer and a resetter, who go through the following cycle many times. The mixer begins with state Ω_0^M and transforms it into the state Ω_{mix}^M, extracting work W_{mix} on average. He or she then passes the state to the resetter, who transforms it back to Ω_0^M with certainty, doing some average amount of work W_{reset}, and then passes it back to the mixer. In the best case, in which both implement their transformations of the system in the most efficient way (i.e. reversibly), $W_{mix} = W_{reset}$. However, suppose that the mixer misinforms the resetter about the probability distribution for macrostates, that is he or she says that the state is Ω_{false}^M. Then the resetter will calculate that the maximum amount of work the mixer has extracted in preparing Ω_{false}^M is W_{false}. Then the resetter may find that $W_{reset} < W_{false}$, and hence falsely believe that together they have implemented a cycle that violates the probabilistic Kelvin statement of the Second Law. This is the sense in which the probability distribution used in the information theoretic calculation of the thermodynamic entropy must be the correct one.

Note also that our main result can be used in a direct proof of Landauer's Principle (Ladyman et al (forthcoming)), and is also useful in explaining

7

how Maxwell's Demon fails to break the Second Law because its memory is a probabilistic mixture of macrostates (Bennett (1987)). On reflection it is not so strange that uncertainty about the macrostate should carry a thermodynamic cost since, as we also argue in Ladyman et al, it's harder to build machines that operate correctly on a family of macrostates rather than a specific one.

Acknowledgements

References

Bennett, C. H. (1987). Demons, Engines and the Second Law. *Scientific American, 257*, 108-116.

Ladyman, J., Presnell, S., Short, A.J. and Groisman, B. (forthcoming), The Connection between Logical and Thermodynamical Irreversibility. *Studies in History and Philosophy of Modern Physics*

Leff, H. S., and Rex, A. F. (Eds.) (2003). *Maxwell's demon 2: Entropy, classical and quantum information, computing.* Bristol: Institute of Physics.

Norton, J. D. (2005). Eaters of the lotus: Landauer's principle and the return of Maxwell's demon. *Studies in the History and Philosophy of Modern Physics, 36*, 375-411.

Uffink, J. (2001). Bluff Your Way in the Second Law of Thermodynamics. *Studies In History and Philosophy of Modern Physics, 32*, 305-394.