Maxwell's demon flourishes

The perennial question of whether suitably instructed microscopic beings could falsify the second law of thermodynamics is still alive, but only just.

JAMES Clerk Maxwell's demon stands in relation to the second law of thermodynamics as does King Canute to the tides in the North Sea, between Britain and Scandinavia. Canute (995–1035), anxious to demonstrate to an over-trusting people that even he could not halt the diurnal tide, allowed his feet to be wetted uncomfortably by the sea. Maxwell, the founder of the kinetic theory of gases, was similarly concerned to show that the second law of thermodynamics could be avoided only by magical means: his demon was meant to be a sentient creature of atomic dimensions, able to discriminate between atoms and molecules.

If properly instructed, a small gang of Maxwell's demons could ensure that the entropy of an isolated system would spontaneously decrease. Take a volume of gas in equilibrium with itself, partition it into two by means of a solid screen and station a demon at an orifice in the screen. If the demon is instructed to prevent high-velocity molecules from moving in one direction through the orifice, and low-velocity molecules from moving in the other, the result will be that the temperature of one part of the accessible volume will increase and that of the other will decrease. Equilibrium will be upset.

Maxwell's argument resembles that of Canute in obvious ways. Canute demonstrated that he lacked supernatural powers. Maxwell showed that a microscopic sentient demon is the only way in which the consequences of the second law can be avoided and then appealed to the general knowledge that such creatures cannot exist. But neither Canute nor Maxwell would be entirely satisfied with the outcome. Belief in the supernatural flourishes, in various forms of parasitism. And there is a rich and flourishing "what if . . . ?" literature of Maxwell's demon which, by contrast, has the virtue of being interesting as well.

That is not surprising. There is, for example, the prior question of whether Maxwell's demon could function as described. Can a microscopic sentient being interfere with the motion of individual molecules so as to destroy thermodynamic equilibrium, and without the expenditure of effort? And then, because deviations from the second law would imply that strange phenomena — perpetual motion, for example — would be possible, could Maxwell's demon (if it exists) be harnessed to useful purposes?

Unsurprisingly, most people have concluded that there is not much scope for it. Carlton M. Caves, from the University of Southern California at Los Angeles, is the latest to have done so (*Phys. Rev. Lett.* **64**, 2111; 30 April 1990).

In principle, a demon should be able to do what is expected of it (him or her) without significant expenditure of mechanical effort, called work. Standing at the orifice in a partition, for example, the process of selecting between molecules of different characteristics would entail three consecutive steps — deciding whether or not an approaching molecule should be allowed to pass through the orifice, closing a frictionless door when the decision is to exclude it and then returning the system to its original condition in preparation for the arrival of the next particle. Making a strictly frictionless door would not, of course, be a simple matter, but there is no intrinsic lower limit to performance.

That is, of course, a recipe for perpetual motion; simply instruct the demon to maintain a constant difference of temperature between the two parts of the enclosed volume and use them as the upper and lower heat reservoirs of a heat engine.

Caves notes that the simplest engine of this kind is the *gedanken*-engine devised by Leo Szilard, in the century after Maxwell. Put a single molecule in an enclosure bounded by two frictionless pistons and equipped with a movable partition dividing the enclosure exactly into two. The partition is dropped, the demon decides in which half of the box the molecule is trapped, moves the piston closing the other half up to the partition (there is no resistance because that side is empty) and then removes the partition. The result is that the one-molecule gas does work on the piston of amount $kT \log_2 2$, where k is Boltzmann's constant and T the temperature. The quantity $k \log_2 2$ is simply the entropy of mixing between the two halves of the box. So is the second law false and, in particular, is perpetual motion possible?

Attempts to save the appearances by pleading that the demon must do some work have mostly been defeated by the argument that the friction of its equipment can be indefinitely reduced. But the demon's Achilles' heel turns out to be cognitive, not mechanical. A demon deciding that a molecule is in one half of a box and not the other must carry out a computation, store the result and then

clear its mind (or computer) for the next cycle, which entails the manipulation of entropy. Rolf Landauer (see *Nature* **335**, 779; 1988) is chiefly responsible for having made the case that the only unavoidable energy cost in the demon's work is that involved in clearing its mind for the next cycle of the operation. The intelligence that the molecule is on one side of the box rather than the other can evidently be represented by a single bit of information, but Landauer shows that the energy needed to erase it is precisely $kT \log_2 2$, so that the second law remains intact.

Caves goes a step further, by conceptually coupling together several (say N) Szilard *gedanken*-engines. He accepts Landauer's argument that the erasure of information is the superintending demon's most probable energy cost, but explores the question raised by W. H. Zurek (*Nature* **341**, 119; 1989) of whether it may be possible to avoid the stringency of the second law by collecting information economically, and thus having less of it to erase from the memory.

The trick lies in the briefing of the superintending demon, which must be told to extract work from the engine only when the storage of information can be handled economically. In the extreme case, for example, the demon might be told not to extract work except when the N molecules in the N containers are on the left-hand side of their respective partitions. Then, the work the demon can arrange to be produced is $NkT \log_2 2$, which state can be represented by only a single bit, the erasure of which will require only the expenditure of energy equivalent to $kT \log_2 2$. The argument, inevitably, proceeds by a combinatorial description of the states of the machine and a calculation of the shortest algorithm that will erase information of unpredicted complexity.

The outcome is teasing, but familiar. Caves concludes that Maxwell's demon could indeed extract work from a *gedanken*-engine, but only by waiting for thermodynamic fluctuations that are, by definition, rare. Designers of perpetual motion machines (an energetic company) will throw up their hands at this point, looking for something a little less chancy with which to impress their sponsors, but Caves seems to have hit on the foundation for yet another wave of speculation about what Maxwell's demon might accomplish.

John Maddox