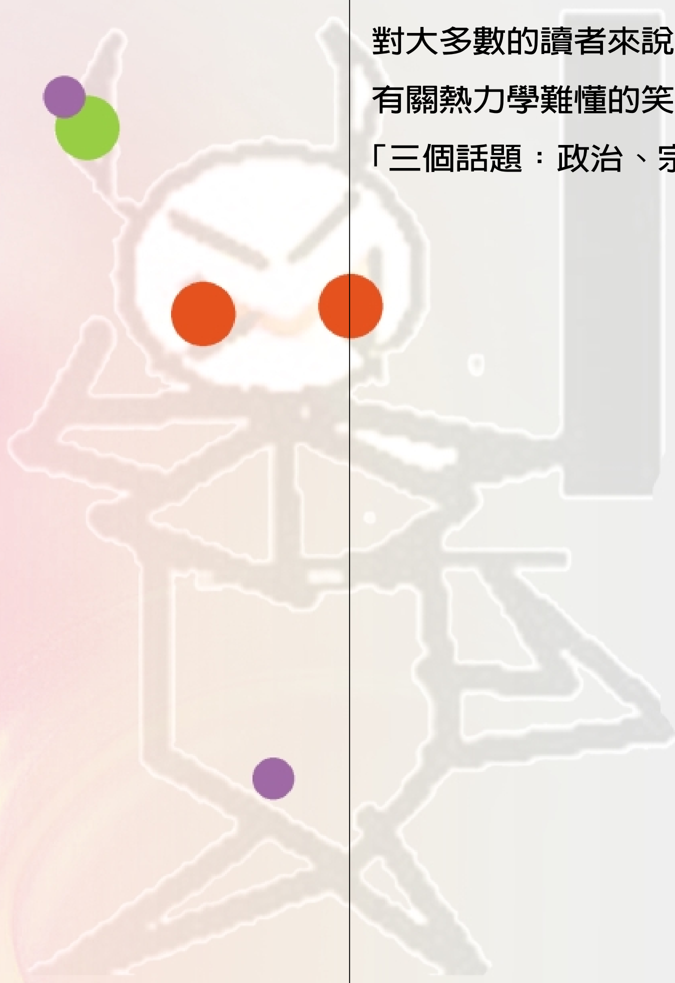
資訊與熵



對大多數的讀者來說，熱力學是一門抽象難懂的學問，
有關熱力學難懂的笑話很多，最常被引用的一段是：
「三個話題：政治、宗教及熱力學是不宜與朋友談論的。」

■汪上曉

豈止一個「亂」字了得

其實熱力學大部分的基本觀念，如功、能、熱、溫度等都非常直觀，引起混亂的只有「熵（entropy）」這一概念，光唸這個字就夠麻煩，如果照有邊讀邊的原則，「熵」應唸成「ㄕㄤ」，但你用注音輸入，卻要用「ㄉㄧ」，方能找到它。隨手查了一些線上字典及工具書，此字的漢語拼音多半是 shang，但也有 shang, di 並列的，教育部的國語辭典則沒有收錄此字，有夠混亂吧！？

根據大英簡明百科全書的說法，「熵」是：「物質系統不能用於作功的能量的度量。熵是一種廣延量，即它的量值由處於一定熱力學狀態的物質的量決定。熵的概念是德國物理學家克勞修斯（Rudolf Clausius）在一八五〇年代提出的，孤立系統的熵只會增加不會減少，此一現象有時也被說成是熱力學第二定律。根據這一定律，在熱氣體與冷氣體的自發混合、氣體往眞空方向自由膨脹以及燃料的燃燒之類不可逆過程中，熵都是增加的。在多數非科技使用上，熵被認爲是一個混亂和漫無目的系統的測量方式。」

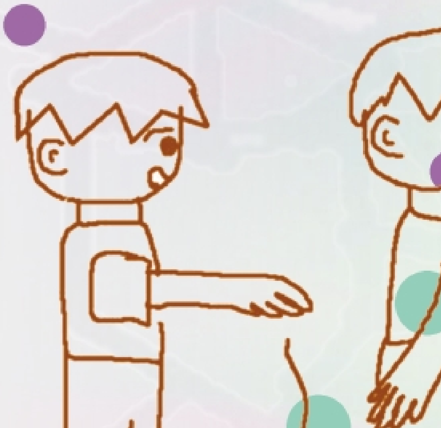
看了上述敘述後，很難理解爲什麼「在多數非科技使用上，熵被認爲是一個混亂和漫無目的系統的測量方式」？

一般大學物理或工程教科書大概沿用克勞修斯的方式定義「熵」：一個與外界沒有物質交換的封閉系統會有一熱力學狀態函數「熵」，當系統經由一可逆路徑從狀態 1 變化到狀態 2 時，若熵的變化爲 A，則當系統經由一不可逆路徑從狀態 1 變化到狀態 2 時，熵的變化必然大於 A。

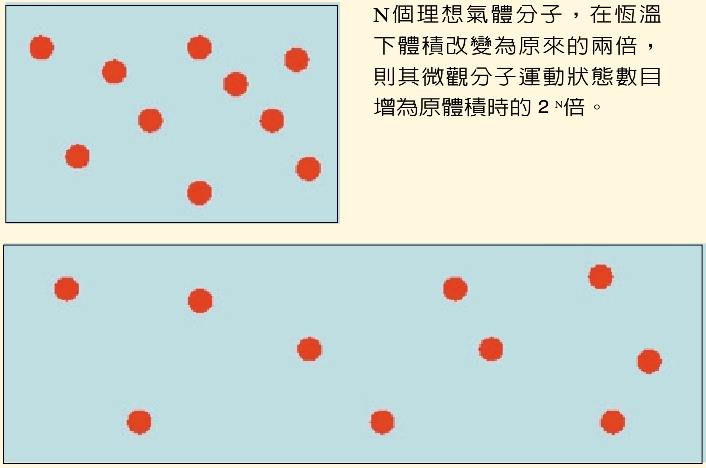
此一定義顯然陳述得不乾不淨，因爲熱力學狀態函數的變化只跟系統的起始及最終狀態有關，與系統經歷的途徑無關，所以熱力學狀態函數的定義應與途徑無關！但上述的定義使用了「可逆路徑」的概念，而當問及什麼是「可逆路徑」，什麼是「不可逆路徑」，人們往往會再拿出熵的定義倒過來解釋，也就是把「熵」與「可逆路徑」的定義綁在一起了。

微觀的概念

要單獨賦予「熵」的定義，通常採用波茲曼（Ludwig Boltzmann）的微觀解釋，假設W代表某一宏觀熱力學狀態所對應的──包括每個分子的運動速度、位置、鍵接旋轉角度及振盪頻率等──微觀分子運動狀態數目，那麼該宏觀狀態的「熵」與W的對數成正比。



此一「熵」的微觀定義又與克勞修斯的概念有什麼關係呢？讓我們舉一簡單的例子來說明：有N個理想氣體分子在恆溫下占據一體積爲V的容器，假設容器體積增加爲2V，系統的「熵」改變多少，對應的微觀分子運動狀態的數目改變多少？

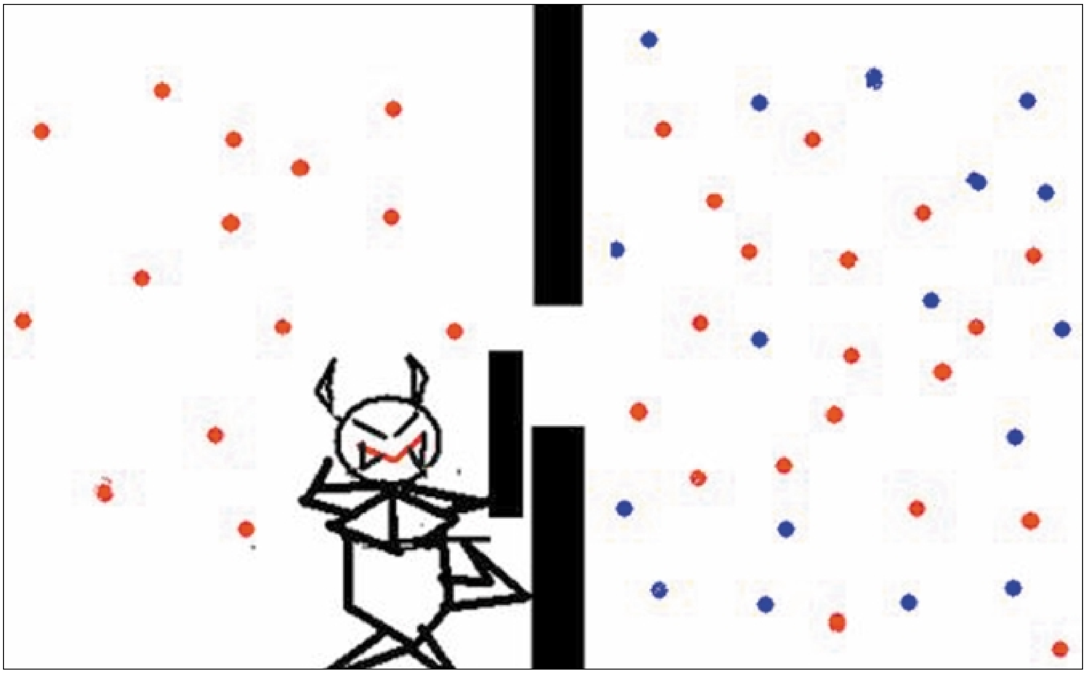


N個理想氣體分子，在恆溫下體積改變為原來的兩倍，則其微觀分子運動狀態數目增為原體積時的 2^N倍。

在恆溫條件下，理想氣體的內能沒有變化，膨脹過程中吸收的熱等於膨脹所作的功，根據克勞修斯的定義，過程中「熵」的變化是 $R\,ln2$，其中R是一常數。如果從波茲曼的微觀定義出發，在原有體積爲V時，每一瞬間N個分子中，任一理想氣體分子只可能在原有的空間V中；當體積膨脹爲2V後，每一瞬間任一理想氣體分子可能在原有的空間，也有可能出現在新增的空間，所以對應的分子運動狀態數目增加爲 2^N倍，由此所計算出來的「熵」變化也正好是 $R\,ln2$。

從上述討論可知，系統可能存在的對應微觀分子運動狀態數目越多，「熵」就越大。系統可能存在的對應微觀分子運動狀態數目越多，表示我們對系統的詳細狀態越不確定，所以「熵」往往被視爲亂度或不確定度的測量。

最早提出此一想法的是馬克斯威爾（James Clerk Maxwell），他提出一個有趣的難題來挑戰熱力學第二定律：根據克勞修斯對熵的定義，一個與外界完全隔絕的系統，它的「熵」只會增加，但假設我們在此一與外界完全隔絕的氣體容器內，加裝一堵設有一道活門的隔間，由一「人」把守，此「人」有辦法測量氣體分子移動的速度，讓高速分子從某一方向通過活



馬克斯威爾小妖在把關，它有辦法偵測分子速率的快慢，藉由控制活門，讓高速分子從某一方向通過活門，而低速分子僅允許自另一方向通過活門，最後使系統出現二個不同溫度區域，而系統的熵也減少了。

閂，低速分子只允許從另一方向通過活閂，最後會得到兩個溫度不一樣的氣體，而系統的「熵」也會減少。

要破解此一難題，在現實世界中沒有「人」辦得到，能夠擔任把守活閂這一角色的只能說是一個妖精，我們姑且將它稱爲馬克斯威爾小妖（Maxwell demon）。在馬克斯威爾的挑戰中，我們已隱約看到了資訊與熵的關聯性，亦即掌握資訊是降低系統熵的重要關鍵。

熵與資訊的關係

然而如何將資訊定量呢？學者薛農（Claude E. Shannon）利用類似「熵」的觀念來定義訊號傳輸中的資訊量，並稱之爲「資訊熵」(information entropy)：

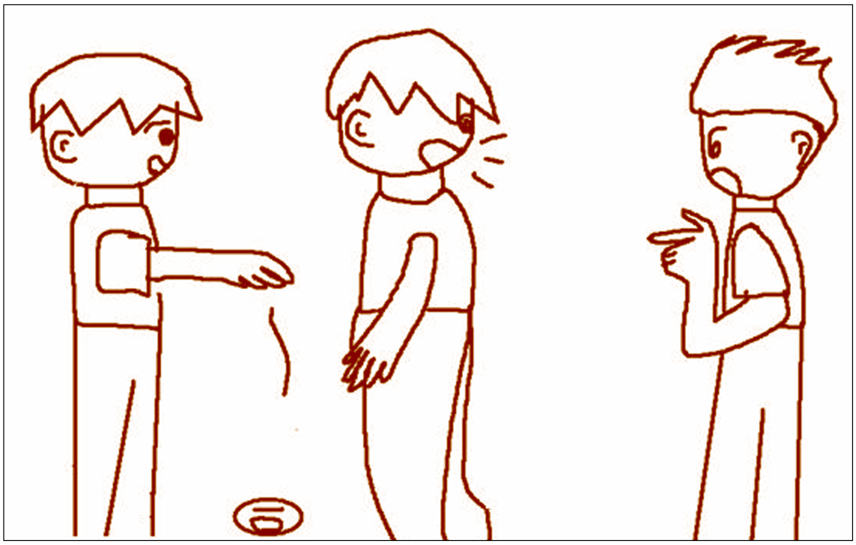
$$\triangle I = -A\log\frac{p_{correct}}{p_i}$$

其中$p_{correct}$是資訊傳播時維持正確的機率、p_i則是被傳播的事件發生的可能機率，A則是與「資訊熵」相關的比例常數。

舉例來說，假設甲拋了一個錢幣，乙對丙大喊一聲說：「是人頭！」乙對丙的通訊中到底包含了多少資訊呢？假設乙的眼力是百分百的可靠，因此傳播時維持正確的機率$p_{correct}=1$，而錢幣落地擲出人頭的機率$p_i=1/2$，因此$\Delta I=A\log 2$，如果乙老眼昏花，看錯的機率是20%，那丙獲得的資訊是ΔI=Alog1.6=0.47A，習慣上我們用bit（$\log_2 2$=1 bit）做爲資訊熵的單位，因此比例常數A在使用不同對數時有不同的數值。

「資訊熵」與「物理熵」是息息相關的。舉例而言，當我們將1莫耳（mol）理想氣體的體積在恆溫下壓縮一半時，「物理熵（ΔＳ）」減少了，但由於分子分布的空間縮小，其位置的確定性增高，因此「資訊熵（ΔI）」增加了，如果我們認定「資訊熵」的增加與「物理熵」的減少是相同的（$\Delta I=-\Delta S$），則取得一位元的資訊相當於每莫耳理想氣體溫度上升1K所需能量（J／mol・K）。

因此有名的熱力學學者路易士（Gilbert Newton Lewis）曾說：「熵的增加意味著資訊的流失，這是一主觀的概念，不過我們可以用不太主觀的方式表達。」（Gain in entropy always means the loss of information and nothing more. It is a subjective concept, but we can express in its least subjective form.）看了本文之後，你對熵的概念是更清楚還是更混亂呢？不要煩惱，願馬克斯威爾小妖祝福您！ □



資訊的傳輸 假設甲拋了一個錢幣，乙對丙大喊一聲說：「是人頭！」乙對丙的通訊中到底包含了多少資訊呢？

汪上曉
清華大學化工系教授