傳播的數學理論之最近研究成果

作者：華倫‧韋弗　　譯注：謝清俊

壹、導論：分析性傳播研究的通則[1]

一、傳播

本文所談的「傳播」(communication)具有相當廣泛的意涵；舉凡一個心智可能影響另一個心智的所有過程、步驟，都包括在傳播的範疇。在此定義下，傳播則不限於書寫和口語的文辭，還包括音樂、繪圖藝術、戲劇、舞蹈，乃至所有的人類行為。有時候，在某些情況下，傳播的定義似乎有必要界定得更廣些。就像傳播可以界定為：一個機制(mechanism)影響另一機制的程序；就如會追蹤且能預知飛機航線的自動設備，在導向飛彈中追逐飛機。

本文所用的文辭常指語藝傳播(communication of speech)這個特殊的學門，其範疇廣泛而重要。雖然如此，本文的論點卻不限於語藝傳播，一樣適用於所有的音樂、漫畫、卡通及電視的傳播。

二、傳播問題的三個層次

與傳播這個廣泛主題相關的問題，大概可分為三個層次，以下依序說明。

- 第一層（屬技術問題）：如何將傳播所用的符號正確傳送(transmitted)出去？

[1] 本論文主要分三部分。其中第一和第三部分的內容之觀點和表達形式都是華倫‧韋弗一人之作。第二部分，名為「貳、第一層次的傳播問題」，則是針對在貝爾電話實驗室任職的克勞德‧夏農博士的數學論文所做的詮釋。誠如紐曼(von Neumann)指出，夏農博士的研究可追溯自波茲曼(Boltzmann)的觀察發現。波茲曼於 1894 年在統計物理學的研究中，談到熵(entropy)和「失落的信息」(missing information)有關。也就是說，在一個物理系統中，把所有大到可觀察到的信息都記錄下來之後，仍然可能遺留下幾種狀態，而此與熵有關。濟拉得(L. Szilard)將此觀念引申到物理學中一般信息的討論(Zsch. F. phys. Vol.53, 1925)，而紐曼則從量子力學(quantum mechanics)和粒子物理(particle physics)的領域討論信息的問題(Math. Foundation of Quantum Mechanics, Berlin, 1932, Chap. V)。夏農博士的這個研究與大約二十年前尼奎斯特(H. Nyquist)和哈特利(R. V. L. Hartley) 的某些想法有直接的關連。尼奎斯特和哈特利也和夏農一樣，都在貝爾實驗室工作。夏農博士自己一再強調其傳播理論的基本哲學主要源自威那教授（Norbert Wiener）的影響。威那教授卻指出，夏農早期對交換理論(switching)和數理邏輯(mathematical logic)的研究，激發了他自己對此領域的興趣。威那還大方地補充說，夏農個人獨自發展了這個介紹熵的基本理論，因此，發展理論的功勞自當歸於夏農。夏農的關注自然較偏向於通信工程的應用，而威那則較關心生物方面的應用（譬如中樞神經系統的現象等）。

- 第二層（屬語意問題）：傳送的符號如何精確地傳達欲表達的意義(meaning)？
- 第三層（屬效果問題）：如何有效地使接收到的意義產生預期的反應、行為？

所謂技術問題是指傳送者(sender)與接收者(receiver)之間信息轉移之正確性的相關問題；此處所說的信息是指一群符號（書寫的文辭）、一種連續變化的訊號（電話或廣播時所傳送的語音或音樂訊號），或者是一種連續變化的二度空間圖案（電視訊號）。以數學性質而言，第一個例子所涉及的是傳送一群屬於一有限集合的獨立附號的問題；第二個例子是指傳送一個對時間連續變化函數的問題；而第三個例子則是指傳送許多對時間連續變化函數的問題，或者亦可視為是傳送一個對時間及二度空間座標連續變化函數的問題。

所謂語意問題涉及認同性(identity)，或是接收者所詮釋之意義是否能滿意地接近傳送者欲傳達的意義。這問題牽涉到情境的複雜程度既深且廣；即使是以語文來溝通相當單純的問題，在此層次也一樣複雜。

導致此複雜情況的主要原因之一，可用以下例子說明。若是王先生懷疑他不了解丁先生說些什麼，那麼，丁先生無論用多長的時間，只要不是無窮無盡的時間，繼續和王先生對話，而不做其他方式的溝通努力，如此在理論上，丁先生是不可能完全澄清王先生的疑慮的。如果丁先生說：「你現在了解了嗎？」而王先生答道：「當然，我了解了。」，這樣的對話並不能保證雙方已獲致了解；也許王先生並不了解問話的意義。如果你覺得上面的例子有點兒呆，那麼再試試這個：若是問：「czy pan mnie rozumie?」而回答：「Hai wakkate imasu.」。我認為溝通的根本困難[2]，至少在語藝傳播的領域裡，是可以經由「解釋」而將它降低到可接受的地步。事實上，溝通的困難是永遠無法完全消弭的。在此所說的「解釋」，是指(一)如前述可能永遠不會比已解釋過最近似的意義好多少，但是，(二)因為這些語言的指涉已經經過某些運作方式成為相當清晰的語言，因此，是可以了解的。舉例來說，任何一種語言中，表示「是」(yes)的符號，並不需要經過太久的比劃，就能讓人懂得它的意思。

就一般的溝通而言，語意問題會變得枝節叢生。譬如說，一個俄國人對一捲美國新聞影片所理解的意義。

所謂效果問題是指接收者收到信息後，會不會成功地有預期的行為反應。初看之下，這種說法似乎狹隘而不合情理：難道所有的傳播都只是在影響接收信息者的行為？然而，對行為(conduct)做相當廣義的界定後，很清楚的情景是：傳播要麼會影響行為，要不然就根本不會有任何效果的可能。

---

[2] 「1911 年龐斯特(pfungst)展示艾伯菲爾得(Elberfeld)的馬兒時，這些馬兒表現出驚人的語言和數學能力。其實，牠們只是對馴馬師頭部動作的反應應而已。這情形使得馬兒的主人柯若爾(Krall, 1911)，備受非議。他問馬兒，是否能看到如此細小動作。馬兒竟拼出『不』(No)字作為答覆。不幸的是，我們不能完全確定：馬兒是不是了解我們的問題？我們是否獲得清晰的回答？」詳見賴西利(K. S. Lashley)的”Persistent Problems in the Evolution of Mind” in Quarterly Review of Biology, V.24, March, 1949, P.28。

效果問題涉及像美學對精緻藝術的考量。以語文的情形為例，無論是說的或是寫的，都涉及全面廣泛的考量：從僅僅是僵化的寫作格式，到宣傳理論的心理學和情緒等的所有層面，以至於價值的判斷等均如此；就像在談效果問題的第一段中的第一句話時，必須先對「成功」和「預期」等詞彙賦予有用的意義後，才能判斷其價值。

效果問題和語意問題有密切的相互關係，而且彼此還有含糊重疊之處；事實上，以上所談的三個層次之間，都有互相重疊的地方。

三、評述

談到這裡，讀者可能認為第一層次是比較膚淺的；它只涉及設計一個良好傳播系統的工程細節而已，而第二層次和第三層次則似乎涵蓋了所有傳播問題的大部分哲思。

這主要由夏農在貝爾電話實驗室發展出來的傳播（通信）在工程上的數學理論，最初是針對第一層次而發展的；也就是討論將各式各樣的訊號，由傳送者正確傳送到接收者之間所涉及的技術問題。然而，我認為這個理論卻有深遠的意義，顯然上一段的看法是不正確的。

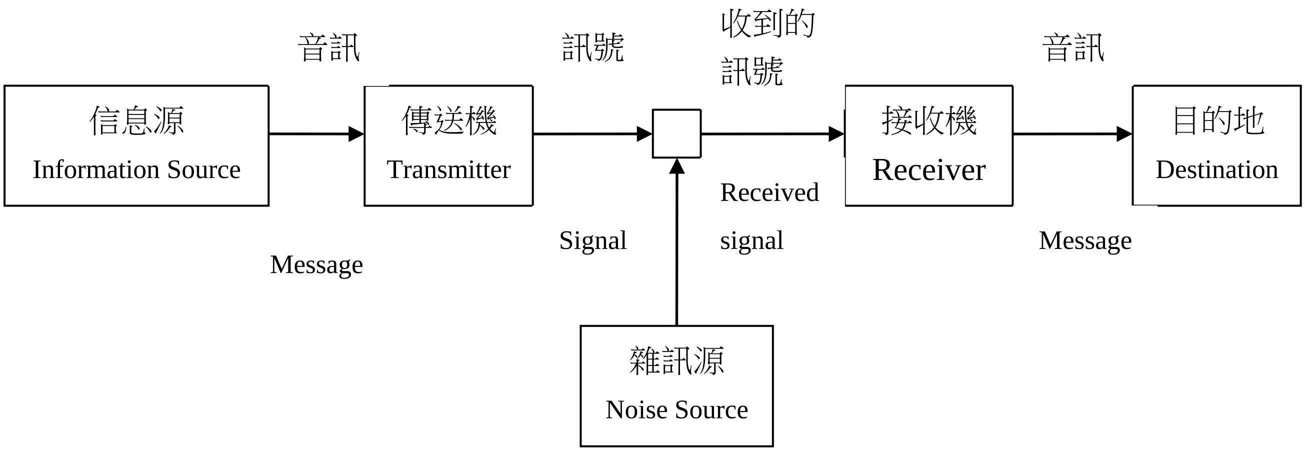
這個新理論的深遠意義是其適用的範圍，不只在第一層次的通信工程方面，也包括第二層次及第三層次方面。詳細的說，在第二層或第三層中，如果要分析訊號的正確性時，只有應用第一層次的分析才能做到。所以，在第一層中數學理論發現任何通信上的限制(limitations)，勢必適用於第二層和第三層。此外，傳播的數學理論的意義之所以深遠在於它的分析顯示：第一層與其他兩個層次之間有相當程度的重疊關係，而其重疊程度遠超過一般人的想像。如此說來，這個源自第一層次的數學理論，也是第二層次和第三層次的理論，或多或少能適用於第二、三層。

我希望下文能剖析、證明上述的觀點。

貳、第一層的傳播問題

一、傳播系統及其問題

上文討論的傳播系統，可用下圖表示。「信息源」(information source)從一組可用的音訊(message)中選取欲傳送的「音訊」（此點是特別重要的說法，稍後將作詳盡解釋）。音訊可以是書寫的文字，或是口語的言辭，也可以是圖片繪畫，或是音樂等。



「傳送機」(transmitter)把音訊轉變成訊號(signal)。事實上，傳送機經由傳播通道(communication channel)傳送至接收機的是訊號。以電話為例，導線就是傳播通道，而訊號則是流經導線的大小不等的電流，至於傳送機則是將語音的聲壓(sound pressure)轉變為大小相當的電流之一組裝置；亦即電話傳送（話）機。以電報為例，發報機（亦即傳送機）將書寫文字轉換成長度不等的電流脈波序列（即「的」、「答」、和空白的序列）。口語傳播時，信息源是大腦，傳送機是指喉部的發聲機制，由發聲器官產生的聲壓就是訊號，它可經空氣（此即通道）傳送。以收音機為例，其通道是空間(space)，而其發射的電磁波(electromagnetic wave)則是訊號。

「接收機」(receiver)是反其道而行的一種傳送機，它把收到的訊號還原為音訊，並將之送達目的地。譬如說，我對你說話時，我的大腦是信息源，而你的大腦便是目的地；我的發聲器官是傳送機，而你的耳朵及其關連的第八神經就是接收機。

在傳送過程中，不幸的是訊號中會徒增一些非信息源想要的東西。這些不速之客像是：電話中干擾的聲音，收音機裡的靜電干擾，電視圖像的變形或是陰影，電報或傳真中所產生的錯誤音訊等等。所有諸如此類使傳送的訊號變形的因素叫做「雜訊」(noise)。

對於傳播系統，人們常常會問的是：

(1) 如何量測「信息的量」(amount of information)？

(2) 如何量測一個傳播通道的「容量」(capacity)？

(3) 傳送機把音訊轉換為訊號時，通常會牽涉到「編碼」(coding)的工作。此時，有效率的編碼應具有那些性質？當編碼的效率已達極致時，傳播通道能以何種速率傳遞信息？

(4)「雜訊」的一般性質為何？在接收的目的地收到音訊前，雜訊如何影響音訊的正確性？如何才能把雜訊的干擾降至最低？而且降至什麼程度時，可望把雜訊消除掉？

(5) 如果傳送的訊號是屬於「連續性質」(continuous)，譬如口語演說或音樂，而非單個不連續的「離散」(discrete)符號，譬如書寫文辭、電報之類，那麼此種訊號性質的差異對上述的問題有何影響呢？

以下將針對上述問題，以最少的數學術語且不用任何公式證明的方式，陳述夏農博士的研究結果。

二、信息

在傳播的數學理論中，信息(information)一詞，有其特殊的涵義；它與日常用語的涵義不同，不應混淆。特別要指出：切勿將此信息與其含意(meaning)混為一談。

事實上，如果有兩個音訊，一個意義非常豐碩，另一個卻全無意義，從數學理論的信息觀點來看，這兩個音訊不分軒至是完全一樣的。正因如此，夏農在他的論文中曾說「傳播的語意部分，與傳播的工程部分無關。」可是，這個命題並不意味著傳播的工程部分必然和語意的部分無關。〔譯按：這段話所說的意思至為緊要，是本書內容的精華之一，也是了解書中理論的關鍵所在。這段話的意思是說：在工程設計上，由於只考慮處理「形式（form）」，因此與形式所承載的意義，即「內容（content）」，無關。換言之，從工程角度而言，不會對「形式」之於「內容」的彼此關係作任何限制。這是為了讓所設計的系統適用於任何「內容」的關係。這是設計「通用系統」的做法。然而，從應用角度考慮，勢必要涉及「內容」。一般而言，對於「內容」的表示常須借重於某種「形式」。因此，從使用者的角度考慮，就不是如上述的情境了。使用者欲傳達某意義時，必需將欲傳之「內容」用工程上所能提供的「形式」忠實地傳送出去。因此，對現成的傳播機器適合傳輸什麼樣的「內容」？和如何傳輸此內容？就會有所權衡和取捨。所以說，在處理第二層和第三層的傳播問題時（如語意問題），則不能不和工程扯上關係。換言之，欲傳之內容必須套在機器能提供的形式之內的這個條件，就造成第二、三層問題對第一層問題之依賴。〕

為了確立此概念，此傳播理論中「信息」一詞，和「你說什麼」的關係，遠不如和「你能說什麼」的關係來得密切。換句話說，「信息」是一個人在挑選音訊時，關於其選擇自由的量測。舉個很根本的例子：假如一個人必須在兩個音訊中挑選一個，那麼，這種情形絕對可說是：此信息與此處境是一體的(unity)。該注意的是，如果說成這兩個音訊的其中之一傳遞著某信息單位，就誤解了原來的意思。這是常

易發生的誤解。此處所談的信息概念，並不能像意義的概念那樣適用於單一的音訊，而僅適用於視處境和音訊為一整體的情形；亦即，一個單位的信息表示在此情境下一個選擇音訊的自由的量，如此設定此單位量的理由是為了方便表達此選擇的處境。

上述二必選一的情況，可施用於任何音訊內容的選擇。譬如其一可以是欽定的聖經譯本(the King James Version of the Bible, 譯按：指 1611 年經英國國王詹姆斯一世所核定的聖經譯本)；其二可能只是"Yes"這個字。在傳播系統中，傳送機可能把這兩個音訊編碼；前者以「0」編碼，後者編為「1」。或者是在執行時，以通電流 (current flowing) 為訊號表示前者，而以斷電流表示後者。如此，可用一只單純的繼電器(relay)之開或關來代表此二音訊。

說得更明確些，信息量在最單純的情況下是指：對其可以選擇的數量所取的對數值。此對數[3]的基數（或稱為底）以 2 最為方便（比 10 為基數的對數方便）。如前述二必選一的情況，信息量為取 2 的對數（以 2 為基數），其值為 1。信息量的單位，最初是由特基(John W. Tukey)命名的，叫做「二進位數元」，或簡稱「位」(bit)，是由二進位數元(**bi**nary digi**t**)兩個字取首取尾組合而成。在二進位的數字系統中，每個位數(digit)的值只可能有兩個，即 0 與 1；正如十進位系統中，每個位數只能有 0 到 9 十個位數的情形一樣。0 與 1 可用來表示任何兩種被選取的音訊，因此「二進位數元」或「位」自然與一個單位信息的二選一情況有關聯。

假設我們可以在 16 種音訊中任選一種，那麼，因 $16=2^4$，故 $\log_2 16=4$；我們稱此情形是由 4 位信息所決定。

無疑的，當人們首次聽到將信息界定為可選擇次數的對數時，會覺得怪異。然而，當你逐漸深入了解此理論後，就越來越會明白這種對數的量測是很自然的。現在，我們用一個現象來說明。如前述，一個只能開或關的繼電器，可以用 0 或 1 來標示開或關的情況，它只能處理一個單位的信息，亦即在兩個音信中作一選擇。若如是，一個繼電器可處理一單位的信息，那麼三個繼電器可以處理多少？很合理的，我們可以說三個繼電器應可處理三倍於一個繼電器能處理的信息量。若是我們用對數來定義信息量，就正是這種情況。三個繼電器有 2^3 即 8 種選擇，若以符號表示，則為：000,001,010,011,100,101,110,及 111，其中第一個表示三個繼電器均為斷路，而最後一個則代表三個均為通路。而對 2^3 取對數時，就是 3。是故對數的量測指示了三個單位的信息量給這種情況，正如吾等所預先期望的。同理，若將選擇的次數加倍，則使音信呈平方數目的增加，此正是使對數加倍；若是信息用對數來測量的話，正好得到倍增的信息。

至此，我們的說明僅限於人為設定的簡單情況：即信息源可以自由地從某些固定的音信中作選擇。這就像是到電信局從標準的祝賀生日電文中選一個拍電報是一

[3] 當 $m^x=y$ 時，我們說 x 是 y 以 m 為底的對數。當 $m^x=y$，則 x 稱為（以 m 為基數的）y 的對數值。

樣的。一個更自然且更重要的情況是從一群基本符號中選擇，而使選取的符號串形成我們的音信。據此，一個人可以一個字一個字的挑，而組合的字串就是我們的音信。

說到這兒，有一個一直居於幕後的要角要隆重現身了，那就是扮演選擇音信角色時的機率。至少，從通信（或傳播）系統的觀點來看，這種連續選擇符號的行為是被機率主宰著；而且，每次選擇都與當時之前的選擇有關係，彼此並不是不相干的獨立事件。以英語為例，如果前一個詞選的是 the，那麼下一個能選擇的詞是冠詞，或是某種形式的動詞的機率是非常小的。事實上，這種機率的影響延申到兩個以上的詞。例如：in the event 之後，下一個詞是 that 的機率非常高，而是 elephant 的機會當很低。

據此，機率對於英語顯然有某種程度的控制，再看看下面的例子，就更明白：在任一本字典中，找不到起首為 J 的字後面跟著 b, c, d, f, g, j, k, l, q, r, t, v, w, x 或 z 的字母，是故這些字母跟著 J 的機率為 0。同理，任何人都會同意像這樣的一串字：Constantinople fishing nasty pink，的機率很低。無疑的，它的機率是低，但不是 0；例如，前一句的句尾可能是 Constantinople fishing，下一句的句首正好是 Nasty pink……。然而，由於閱讀這一段，讀者亦可發現這個不尋常的四字詞確實出現在我們的一個句子中。

一個系統若依某種機率而產生一連串符號（這些符號當然可以是字母，音樂裡用的音符，而不必是詞）者稱之為隨機程序(stochastic process)；隨機程序中若其機率是依前面的事件而定時，則稱為馬可夫程序(Markoff process)或馬可夫鏈(Markoff chain)。馬可夫程序可產生各種性質的音信字串，其中有一種對通信（傳播）理論特別重要的，即是遍歷性程序(ergodic processes)，相關的分析細節與推理步驟是很複雜且深入的，要最好的數學家傾其全力才能釐清它關聯之理論；然而，此遍歷性程序粗略性質却不難理解。遍歷性程序能產生具有「理想」性質的一串符號；此所指之「理想」是由於：只要在符號串中取得合理大的樣本，其即能顯示出此遍歷性程序所可能產生的全部符號串的性質。它所產生的符號串是公認最理想的，因為任何合理足夠大的樣本即能夠有代表此類全部符號串的性質。假設有兩個人從不同的方法來選樣本，而研究其樣本趨大時有那些統計上的性質，若是處於遍歷性的情形，則無論它們如何選擇樣本，對全部性質的估計將會一致。此遍歷性程序，從另一角度看，顯現出統計規率上一個特別安全而且舒適的系統類型。

現在讓我們再回到資訊的概念。當我們有一個信息源，它以連續的方式選擇離散的符號產生音信（離散符號如字母、詞彙、音符、標點、或大小不同的點(spots)等等），且每一次選擇其選擇符號的機率是依以往選擇而決定（如，一個馬可夫程序），那麼，這樣的程序究與信息有何關聯呢？

唯一符合這種自然要求來界定「資訊」的量，就是熱力學中的熵(entropy)。它由許多相關的機率所組成：包括在形成音信過程中到達某系統狀況的機率，以及在此狀況下選取下一個符號的機率。進一步說，它的公式，也涉及求取機率的對數值，於是不難明白，上文所談到很基本對數量測所做的自然推廣。

對於已學過物理的人而言，有一個像量測熵的公式出現在資訊量測的理論中，是一件很有意義的事。近一百年前，克拉氏(Clavsius)將之介紹給我們，波茨曼(Boltzmann)和它的研究工作有密切的關聯，吉伯士(Gibbs)在他的古典統計力學中給它深層的意義，自此熵已變成一個非常重要的基本概念，愛丁頓(Eddington)曾詮釋說道：「熵的定律是：它的值永遠在增加—此即熱力學的第二定律常真，我想，它在自然定律中佔著超級重要的地位。」

在物理學中，熵是對系統紊亂（隨機）程度(degree of randomness)的一種測量，若你願意的話也可將之稱為「洗」（洗牌的洗(shuffledness)）的程度；依此，所有物理系統結構的趨勢是將變得越來越沒有組織，變得越洗越紊亂。這種情況是所有物理系統的基本性質，以致愛丁頓(Eddington)辯稱，時間及其流向即主要依此性質而來。我們可以想像，若物理世界的演變像一個電影，那麼熵在決定此電影是向前放映或向後放映。

因此，當一個人在通信理論裡遇到了熵的概念，他是應該感到興奮，並且有權來懷疑他是否擁有一些能轉變為基本且又十分重要的東西。終究，用熵來測的信息是自然的；在通信理論中，信息是關聯到我們要建立音信時究竟有多少選擇的自由。因此，對一個信息源我們如果說：「這個系統是有高度組織的，它不是很紊亂，沒有許多選擇；換言之，這也就是說此系統的信息量（或熵）是很低的」。我們以後還會回到這論點，因為除非我錯得離譜，否則，此論點將是此傳播理論更廣的一個重要推論。

當計算好了某個信息源的熵（或說是信息、或說是選擇的自由度），吾人可以把它與能獲得的最大的熵值做個比較。在不改變其使用的符號集合的限制下，此比值，即實際的熵比之於最大的熵，便叫做此信息源的相對熵值。如果，有一信息源的相對熵值是 0.8，它即概略指出此信息源在選擇符號產生音信時，均可到達 80%最大的自由程度。用 100%減去相對熵值即稱為重覆程度(redundancy)，它是表示在音訊結構中不能由送者自由選擇的部份，是取決於選擇符號的統計規律的。重複程度是個感性的稱呼。從常識上來看，它的確是可表示重複的一部份音信；換言之，這部份的音信是不必要的（亦即重複），若它遺失了，原音信仍然完整，或至少能設法恢復完整。

英語的重複程度是大約 50%[4]，有趣吧！依此，在我們說或寫時，約有一半的字母或詞彙，是可以任由我們自由選擇的，而另一半（雖然我們平常並未意識到）是由語言的統計結構來控制的。暫時撇開嚴肅的推論，說些輕鬆的；我們會發現一個語言若是要能建構一個令人滿意的拼字遊戲，則此語言至少要有 50%能選擇字母的自由。若是它有 100%的自由，那麼任何可能的字母排列，都將是一個可拼的字。若只有 20%的自由，則不可能建構一個拼字遊戲以使它有足夠的變化與複雜程度來吸引大眾。夏農曾估計過，若英語只有 30%的重複，那麼就可以建構一個三度空間的拼字遊戲。

在此節結束之前，我們應該注意在第一層次的分析、論說以及信息的概念。此概念掌握了信息源全部的統計特性，而不是直接干係到個別的音信（而且也和個別音信的含義毫無關連）。因此，從工程觀點來看，其實際的理由是一個通信系統必須面對「要能處理任何一個信息源可能產生的音訊」的問題。若是不能設計一個系統將每件事都處理妥當，則應將此系統設計得能把經常要遇到的事做好，而不去做一些他不能做好的少數事情。這樣的想法，馬上引導到為各種信息源能產生那些音信，以及必須界定其所有音信的統計特性的問題。在通信理論中用到的「信息」，正是用解決此問題的。

本文的目的絕不是來介紹數學細節的，然而，為了測是信息，還是能夠了解熵的算式越多越好。設若有 n 個獨立的符號，或是一群 n 個獨立的音信，它們選擇的機率是 $P_1, P_2, \ldots, P_n$，那麼實際上資訊的算式是：

$$H = - [P_1 \text{ long } + P_2 \log P_2 + \ldots\ldots + P_n \log P_n]，$$

或 $$H = -\Sigma P_i \log P_i$$

其中Σ，像一般數學裡的用法，是表示各項的和，而一個項就如同 Pi log Pi 表示的[5]。

這樣看來有些複雜，但是讓我們看看在一些簡單情形下，這個式子有什麼樣的表現吧。

首先讓我們在兩個可能的音信中作一選擇，它們的機率分別一個是 P_1，另一個是 $1-P_1$，在此情形下，H 之值是當兩者的機率相等時最大，等於 1；也就是說當 $P_1=P_2=1/2$ 的時候，你有完全的自由從這兩個音信中選取其一。當一個音信變得比另一個易於發生的時候（例如 P_1 大於 P_2），H 的值就會減少。而當一個音信變得有很大的機會發生時（如，P_1 幾乎等於 1 而 P_2 幾乎是零的情形）則 H 的值亦將變得很小（幾乎是 0）。

---

4 （原書第 13 頁）這 50%的估計值只是由八封信裡統計出來的，因此，其值究竟有多少，一般相信會比此值略高。

5 （原書第 15 頁）不要擔心此處之負號。任何一個機率的值都是小於或等於一的數值。而對於小於 1 的數字取其對數時，它的值是負的。因此，在此加負號是有必要的，它可使 H 為正的數值。

在極端的情形下，一個的機率是 1（完全肯定會發生），而所有其他的機率是 0（完全不可能發生）時，那麼，H 之值將是 0（完全沒有不確定之因素，沒有選擇的自由，即亦無信息）。

因此，當兩者機率相等時（此時可以毫無偏見地自由選擇）H 最大，而當你失去選擇自由時，H 的值也就不見了。

上述的狀態是一個典型。若此時不只兩個選擇，而是有許多選擇時，則當在環境的允許下，各個的機率越接近則 H 的值越大—當選擇的人有更多的自由作選擇時，而自然會儘量減少被驅入於做一些必然的選擇。設若，從另一角度來說，有一個選擇的機率幾乎是 1，而其他的則趨近於 0，這種情形是很明白易懂的，此時的選擇有強烈選擇某件特殊音信的傾向，因之沒有多大的選擇自由。而此時計算所得之 H 值亦很小—此信息（選擇的自由，或不確定的因素）是很低的。

當可選擇的數目是固定的時候，我們已知道當每個選擇都有相等的機率時，其資訊量最大。此外，還有一個重要的途徑來增大 H，那就是增大可選擇的數目。更精確地說，若是每一個選擇的機會都是一樣大，而能夠選擇的數目越多，則 H 的值就越大。若是你自五十種音信中隨意選出一種來，自然比自廿五種音信中隨意選取一種時有更多的「信息」。

三、傳播通道的容量

在看完上一節的討論之後，對於通道之容量不以它能傳輸多少個符號來敘述，而以它能傳輸多少信息來界定，就不會感到意外。更好的說法是，一個通道的容量，是以它能傳輸多少信息源的產出來描述。

如果這個信息源是一種簡單的類型，它的每一個音信都佔用相同長度的時間（例如，電傳打字 teletype 系統即是如此），若是每一個被選取的符號都以 s 筆資訊表示（可自由地從 2^s 個符號中自由地選擇），以及假設此通道每秒可以輸送 n 個符號，那麼，此通道的容量 C 則定義為每秒 ns 筆。

對於較普及的個案，則必須對每個符號它可能用不同長度的情形加以考慮。因此，通道容量的一般公式將考慮在某種的時間長短下，一些符號個數的對數值（當然，這就涉及資訊的觀念，對應於前一段中個案裡的 s）；而且，也涉及處置了多少個這樣的符號（在上一段中，即指 n）。因此，在通常的情況下，容量的測量不是以每秒傳送了多少符號，而是每秒傳送了多少資訊量，用每秒多少筆作為單位。

四、編碼

在一開始，我們就已指出：傳送機接收音信並將之轉換為叫做訊號的東西，而訊號才是真正通過通道送給接收機的。以電話來說，傳送機僅僅將可聽見的語音改換為某種東西（即是在電話線上變化的電流），它顯然地不同於語音但又顯然地與

之等值。可是，傳送機對音信可以做更複雜的運作來產生訊號。例如，它可以用某些碼將一個書寫的音信加密為一連串的數目字，而這串數目字就由通道當作訊號傳送。

因此，有人說，一般而言，傳送機的功能就是對音信編碼(encode)，而接收機的功能就是解碼(decode)。對於一些設計得很成熟的傳機及接收機而言，例如，能擁有一些「記憶」(memory)的，這個理論提供了如何對於某個音信的符號編碼，它不僅要依此符號而定，而且也依先前音信用的符號以及它們的編碼方法而定。

現在，我們來談一談由這個理論，為了傳送個別符號的一個無雜訊通道，所導出的基本定理(theorem)。這個定理與通信通道，其容量為每秒 C 筆，並從一個熵（或資訊）為每秒 H 筆的資訊源接收訊號，等有關。這個定理說：若是為傳送機設計適當的編碼步驟，是可能在通道上以接近平均 C／H 之速率傳送符號[6]；然而，無論編碼有多聰明，其速率永遠無法超過 C／H。

等一會兒，當面對更普及的情形，即有雜訊時，對這個定理的意義將有更有用的說明。而目前，重點在注意碼所扮演的決定性角色。

記得吧，一個產生音信或是訊號的程序的熵（或資訊）是由這個程序的統計性質來決定的—即選擇下一個符號的狀態時，由其到達某音信的情況和選擇時之各種機率來決定。資訊源的性質可以完全決定音信的統計性質。可是，實際上經由通道中傳送的訊號的統計性質，亦即此通道之熵，是由我們試圖餵進通道些什麼，以及處置各種情況的訊號時此通道的容量，此兩者共同來決定的。以電報為例，在點(dots)與點之間必須有空隔(spaces)，點與劃(dash)之間，劃與劃之間等也都是一樣要有空隔，否則點與劃就無從分辯。

現在，變成這種情況，即當一個通道有這類的限制(constraints)時，它將使訊號的自由程度無法完全呈現，有某些統計上訊號的性質會導致一個訊號的熵比它其他任何一種統計上的訊號結構的熵都要大，而這是一個重要的案例，此訊號之熵正好等於其通道容量。

用這種想法，現在就可以精確地描述最有效的碼的性質。事實上，最好的傳送機是這樣將音信編碼的：它將選擇剛好有這種最佳統計特性的訊號，以適應（配合）所使用的通道—而事實上它是使訊號之熵變成極大（或許也可說是通道的），並使它大到等於通道的容量 C。

這種編碼，由於上述基本定理，將導致傳送訊號時 C／H 的最大值。可是為了獲得這種傳送速率的增益，得付出代價。比較糟的情形是，當把碼調整則越來越接近理想時，處理碼的過程亦將被迫越來越長。對抗這種禍害的方法之一是利用電子設

---

[6] （原書第 17 頁）我們記得，容量 C 涉及每秒傳送多少資訊的概念，而它是以每秒多少筆數來測量。此處之熵 H 是對每個符號資訊量之測量，是故 C 與 H 之比值是以每秒多少符號為測量單位。

備的速度，所謂「長」在電子設備中也許指的只是一秒鐘的極小一部份時間；另一種方法是做個妥協，在傳送速率的增益和編碼時間的損失兩者之間，求一平衡。

五、雜訊

雜訊是如何影響資訊的？我們必須記牢，資訊是對選取音信時，選擇自由度的一種量測。選擇的自由度越大，資訊量就越大，對於選到某特定音信的不可預定程度也就越高。因此，選擇的自由度越大，不確定的程度越高，手牽手似地傳送的資訊就越多。

如果雜訊來了，那麼接收到的音信中將會含有某些變形，一些誤差，一些多餘的東西，而這種情形無疑的將使我們說：因為受雜訊的影響，接收到的音信將表現出增加不確定的程度。可是，若不確定的程度增高，則資訊亦增加，這樣看來雜訊還蠻有好處的嘛！

通常，若有雜訊時，接受到的訊號是會表現出增大資訊的情形—或者說，接收到的訊號是從一個比可傳送的訊號集合更大的集合中取得的。這情形，漂亮地做了個語意陷阱的示範，如果不記得資訊在這兒有其特別的意義，是用以測量選擇的自由程度的，對於已做了那個選擇而言是測量其不確定性的，那麼就會陷入此陷阱。因此，對資訊這個詞就可能有好或是壞的內涵。不確定性，對於傳送者而言，是選擇自由度的德性，是我們所要的不確定性。由於錯誤或是由雜訊的影響而引起的不確定性質，是我們不想要的不確定性質。

因此，當明瞭了這些觀點，就知道在說到接收到的訊號有更多的「資訊」時，它在那裡開頑笑。有一些資訊是假的，是不想要的，却被雜訊引入。為了得到能用的資訊，我們必須把假的部份自接收到的訊號中抽除。

在我們能夠澄清這個觀點以前，我們須要停下來先繞個小彎子。假設我們有二組符號，像是由資訊源產生的音信符號，以及實際接收到的訊號符號。這兩組符號的機率關係是相互有關的，因為明顯的在收到某符號的機率將依其被送出的是什麼訊號而定。若從雜訊或其他原因中未生錯誤，收到的訊號將準確地檔當於送出的音信符號；而在可能有誤差出現的情況時，接收到的符號的機率將明顯地沈重地擔起相當於或接近於送出音信的機率。

現在，我們可以算算看稱之為一群等號相對於另一群之間的熵。讓我們用例子來觀察，想想看音信對於訊號的熵。不幸的是，若我們不稍涉及細節，將無法了解此處涉及的問題。現在，假設我們知道某一個訊號的符號已實際上收到了。那麼，每一個音信的符號便產生某種機率—若是此音信的符號和收到的相同或是相似，它的機率值相對的會較大，而其他的機率值相對的較小。用這一群機率，可以計算出一個臨時的熵值。這是在確定已知收到的符號或是訊號的符號之後前題下音信的熵。在任何良好的情況下，這個值是很低的，因為，此機率並不是依不同的個案平均地

各處分配，而是集中負載於一個或幾個個案上。若完全無雜訊則它的值應為 0，這是因為除了一個符號（即所收到者）其機率為 1 以外，其餘的均應為 0。

對於每一個收到的訊號符號，都可據此假設求其隔時的音信熵。算完它們的全部之後，可依計算時各個訊號符號之機率加權，以得其平均值。像這樣考量兩組符號而算出之熵值稱為相對熵(relative entropies)。剛剛談到的特別例子是音信相對於訊號的熵。夏農為他命名為等效熵(equivocation)。

從等效熵的計算過程中，我們可看出它的意義。它是在測量：當訊號是已知時所求得的「音信的平均不確定性質」。若無雜訊，則在訊號已知的情形下，根本音訊不會有不確定的情形。若是在訊號已知後，資訊源仍有任何殘餘的不確定性留存著，則它必然是由於雜訊所致的不確定性。

前幾段的討論都是圍繞著「當接收的訊號是已知的情形下，音信源的平均不確定性」這個量為中心。若把上句話重組為：「當傳送的音信是已知的情形下，接收訊號的平均不確定性」也是同樣能成立的，而且當沒有雜訊時，以後者之值亦當為 0。

至於這些量之間的彼此關係呢，不難證明是：

$H(X)-H_y(X) = H(Y)-H_x(Y)$

其中 H(X)是音信源的資訊或熵；

H(Y)是接收到訊號的資訊或熵；

$H_y(X)$是等效熵，即當訊號已知時音信源之不確定性；

$H_x(Y)$是送出的音信已知時，接收到信號之不確定性，或者說它是由於雜訊所引發的假資訊，是接收到訊號資訊中的一部份。

此等式的右邊是很有用的資訊，它與傳送中雜訊有多麼壞的影響無關。

現在可以解說一個有雜訊通道的容量 C 是什麼意義。事實上，它的定義是等於在有用的資訊（如總共的不確定性減去雜訊的不確定性）能夠在通道上傳送的情況下，最大的速率（以每秒多少筆表示）。

為什麼在此要說「最大的」(maximum)速率呢？我們能做些什麼，亦即，使此速率大些或小些？答案是我們能以選擇資訊源的方式來影響這個速率，若是此資訊源的統計性質能適合此通道的一些天然限制情況的話。因此，我們可以選擇適當的編碼來得到傳送有用的訊息的最大速率。

最後，讓我們談一個有雜訊通道的基本定理。設若此有雜訊的通道有一容量 C，並假設它由一個熵為 H(X)的資訊源接收訊息，而接收的訊號則有熵 H(Y)。如果通道容量 C 是等於或大於 H(X)，那麼引用適當的編碼系統後，資訊源的輸出可以經此通道傳送得要多正確便可以有多正確。無論你把允許的誤差機率訂得多麼小，總有一

個碼能滿足需求。但是，如果通道容量 C 比 H(X)小，亦即比資訊源的熵小，那麼，就不可能去找到一個碼能將誤差頻率降到我們所期望的低。

無論將編碼程序設計得多巧妙,當訊號收到後總會有一些不是我們想要的關於音信是什麼的不確定性，這是不可避免的情況；而這個不想要的不確定性——這個等效熵——將永遠等於或大於 H(X)-C。更進一步說，經常至少有一種碼能夠減少這個不想要的不確定性，至於對音信來說，可將大於 H(X)-C 的部份降低到任意的小。

當然，最重要的事是，無論編碼程序多麼複雜或適當，這不想要的或是假的不確定性的極小值是不可能再減少的了。這個甚有威力的定理給我們一個精確的，也是幾乎令人訝異的簡單描述；當一個有雜訊的通信通道運作時，我們能從它獲得最大的依賴程度是什麼樣子。

夏農曾指出一個實用上的效果，值得一提。因為英語有約 50%的重複，因此若是經過一個無雜訊的通道，就可能找出一個適當的編碼程序，以節省一般電報一半的時間。然而若有雜訊，用一個沒有完全消除所有重複的編碼程序反而有一些實際上的好處。因為剩下的重複將協助我們與雜訊博鬥。這很容易明白的，例如，在我們收到經傳送來的電文時，我們毫不遲疑地會更正其中一些錯誤的字母拼寫。

六、連續型音訊

直到目前為止，我們一直在討論由個別（離散）符號所組成的音信，像是用字母併成詞，用詞組成句，用樂符組成譜，和用點組成的黑白圖案等。若是考慮到連續形態的音信，像是說話的語音有著連續的基音(pitch)和能量變化時，對所談的理論會有什麼事發生呢？

很粗略地我們可說，這理論的延申是涉及較複雜和較艱深的數學，然而，在重點上沒有主要的差別。有許多在前面為個別訊號所發表的論點根本不須要修改，有一些只需要做些小的修改。

有一個很有幫助的情況如下，當考慮現實時，我們通常會對連續的訊號產生興趣，這些連續訊號由一些單純的諧頻組成，然而它並不包含所有不同的頻率，而是在一段頻帶(band)之內，例如從 0 至每秒 W 週。因此，雖然人類的語音是含有較高頻的成分，非常令人滿意的電話通信却可由四千週以內的頻帶獲得。當頻率昇至十或二十千週，高傳真的交響樂無線電傳送也是可能的。

有一個很方便的數學定理說道：一個連續訊號，有 T 秒之久，而頻帶限制在 0 至 W 之間，則可用 2TW 個抽樣完整地界定它。這確實是一個值得一記的定理。通常，一條連續的曲線只能在其上取若干點作為代表此曲線之近似性質，若是要以此方式完全能表示其性質，則須無數個點。然而，若此曲線是由一些有限數目的單純的諧頻(simple harmonic)成份所組成，像是複雜的聲音由一群有限數目的純聲(pure tones)組成，那麼，它所需要的僅是有限數目的參數，即可完全表達原來的曲線。此點，對於要降低連續訊號複雜的通信問題而言，有極大的好處，因為它可以把處理一個無窮數目的變數的問題，化簡為處理有限變數的問題。

在連續的情況下，也有些公式已被導出來描述最大的通道容量 C，在頻寬為 W，當用以傳送的平均能量為 P，而通道在忍受雜訊能量為 N 的情形。此雜訊稱為「白的熱雜訊」(white thermal noise)是夏農界定的一種特殊雜訊。此種雜訊自己本身的頻寬已有限制，而其每個頻率分量的峯值又受到正常（高氏 Gaussian）機率分配的約束。在此情形下，夏農得到了相關的定理，它又是一個在簡單程度上和應用範圍上皆令人印象深刻的；它是：用最佳的碼，是可能傳送二進位數元於下列之速率：

W ・log2【（P+N）/N 】　　　　單位每秒多少筆

而能有隨你想要的誤差頻率。然而，無論如何編碼，在不計高誤差頻率之下，這個限制是不可能超越的。若是討論任意的雜訊，而不是特別的上述的「白的熱雜訊」，夏農並未能成功地導出一個計算通道容量的公式，然而，他找到了此時通道容量的上限和下限公式。而且，他也導出來當我們不界定平均傳送能量，而以傳送能量之峯值(peak)取代時，通道容量的極限。

最後應該要指出，夏農所得到的結果是必然會有些不是那麼具體而微的，然而它却是明顯的深奧，而且有橫掃全局的影響力，對一般的連續音信或是訊號而言，它說明了收到的訊音信的忠實性(fidelity)以及其相關的性質，並且對於和忠實度有關的量，諸如：一個資訊源產生音訊的速率，傳輸的速率，以及通道容量等的觀念和性質都有詳盡的描述。

參、傳播問題　三個層次間的相互關係

一、簡介

在本文的第一節裡，曾建議在考量一般傳播問題時可將之可分為三個層次，現以問話的方式表達如下：

第一層次：傳播的符號能如何正確地傳送？

第二層次：傳送的符號如何能準確地攜帶欲傳達之意義？

第三層次：收到的意義能如何有效地影響到預期之行為？

同時，也建議這個通信的數學理論，由夏農、韋弗和其他人發展出來的，特別是夏農導出了許多肯定的工程定理，雖然初看之下只應用於第一層次的問題，然而，實際上它對於第二層次及第三層次的問題均有幫助，而且值得參考。

然後，我們在第二節裡，看看這個數學理論究竟是什麼，發展了些什麼觀念，得到了些什麼結果。在這結束的一節中，目的在於回顧一下情況，看看到什麼程度和用什麼字眼，能夠證明第一節裡表示的論點：在第一層次中的進展成就能夠給第二層次與第三層帶來貢獻。這可由這三個層次間的相互關係的密切可以證明；也許由於三者之間如此密切，最後的結論可能使人覺得這三個層次的劃分是這麼不自然而覺得，沒有必要這麼做。

二、第一層次中此理論之一般性(generality)

第一個明顯的評語，也是內容最有份量的評語，是此數學理論的確是有非常普遍性的涵蓋範圍，對問題的處理是基層的，並且其結果達到典型的單純性和典型的威力。

這個理論是如此普遍以致不須要說用那種符號——是書寫的字母、詞彙也好，是樂符也好、是口語也好、是交響樂、圖畫都可以。這理論是如此深奧以致於它談到的科際關係可應用於各種不同形式的通信、溝通或傳播。當然，這表示這個理論充份由想像的動機觸發的，以致他能夠觸及傳播問題的真正核心——即以一些一般認可關係為基礎，無論實際的案情是什麼均可用得上。

由於這種通用性質的佐證，這個理論對於編碼形式有重要的貢獻，而對基本的加密過程亦是如此，因為加密本身就是一編碼過程。同理，此理論對於語言的翻譯，從一語言譯為另一語言，亦有貢獻；雖然翻譯之事明顯地須要語意的考量與資訊的考量。相似的，這兒所發展出來的理論和想法和偉大計算機的邏輯設計亦有密切的關聯。這也難怪夏農最近剛剛寫了一篇論文來討論如何讓計算機有能力下好一盤棋。而它更進一步中肯地針對目前的爭論在論文中作結語時指出，要不是我們承認這樣的電腦「能想」，就應該大幅修改所謂「去思考」的動詞一般的含義。

第二點，在這個理論所依據的基礎上，制式地發展一般性的傳播理論似乎是顯然可行的。起先，就像這個理論剛開始發展時一樣，可以劃個傳播系統的草圖，然而當我們相信了它而它順暢地引導我們到核心的一些重要事件後，就會深深體會並且賞識此理論分析問題的深入。幾乎可肯定的是在考量第二層次和第三層次的傳播問題時須要在本書第□頁的結構圖中加點什麼，可是也幾乎可以肯定所要加的都是一些細節，無關宏旨，不會有實質上的更新。

於是，當我們推移到第二層次或第三層次時，可能可以證明要顧及接收者端統計的性質是有其必要。我們可以想像，在結構圖中的接收機（即將訊號轉換為音信的設備）和目的地之間增加一個「語意接收機」。這個語意接收機主導接收音信的二次解碼，如此的話此語意接收機將會要求音信的語意統計特性要能匹配接收機的語意統計容量，或者是匹配到我們希望要影響的那一群聽眾。

同樣的，我們也可想像將另一個方塊「語意雜訊」加在結構圖的資訊源與傳送機之間，把原圖中只標了「雜訊」的方塊改成「工程雜訊」。從這個雜訊源，加之於訊號的是語意上的干擾或變形；而這些干擾或變形雖不是資訊源想要的，但是卻無可避免地會影響到接收者。此時語意接收機解碼時，須要能夠將這些語意雜訊列入考慮，似乎也可以想像來調整原始音信以使得它和語意雜訊之和正好是我們要送往目的地的語意。

第三點，對各層次的傳播問題來說，似乎無論編碼做得多麼好，若是試圖傳太多音信從一通道中傳過去（如 H＞C），則所傳音信的誤差和混淆將升高而傳真度(fidelity)將下降。同樣的理由再來一次，這個一般性的定理對各層次的問題都確切地表示：不僅要考慮通道的容量，而且還要顧及聽眾的容量（即使每個字都是正確的，若音信量超過聽眾的容量則無法完全接收）。如果給聽眾的超過聽眾的容量，那麼直接用比方來說，或許也是對的，就是只有讓多餘的溢了出去而他無法全收。更可能的是，也用比方來說吧，若你估錯了聽眾的容量而傳送的太多，則勢必造成一些無法規避的誤差和混淆。

第四點，如果說在第二層次和第三層次中不須要從這個理論所發展出來的「熵的觀念」以及它和「資訊的觀念」之間的關係等學到什麼，或是從這個理論解決這些問題的方法上學到些什麼的話，是令人難以置信的。

這個理論所導出的資訊觀念，初看之下似乎是令人失望且古怪的。對它失望是因為它與資訊的意義無關，覺得它古怪是因為它不是對付一個音信而是對付一個集團的音信的統計性質，古怪的另一個感覺來自它把「資訊」和「不確定性」的統計性質視為伙伴，是一類的。

然而，我想以上的反應是暫時的；最後，你會說這個理論的分析是如此地清淅透澈，說不定使你覺得第一次遇到這麼有意義的理論。一個工程的通信理論就像是一個端莊賢慧的小姐正在接到你的電報。她對其中的意義，無論是憂傷、或歡愉、或尷尬都不注意，只是她必須準備好來應付所有送到她桌子上的東西。這想法，一個通信系統應該能應付所有可能的音信，是聰明的，而依據資訊源統計的性質來達到此目的更是高明，一般而言不能認為它對傳播沒有意義。語言必須設計得能使人想說什麼就完全能做到；但是並不是要能做成每一件事，它亦應該盡可能常常做到。這也就是說它要從統計上來考慮如何應付它的任務。

將資訊的觀念施之於資訊源直接導致研究語言的統計結構；例如，用之於英語的研究顯示其資訊對於各級學習語言和傳播的學生都有確切的意義。利用馬可夫程序理論的威力來對付語意研究似乎特別有潛力，因為這個理論是特別用以處理有關「意思」(meaning)或稱之為內文的影響(influence of context)這種極重要且艱難的問題。我們有一粗略的感覺，資訊和意思兩者可能能夠證明它們像是一對在量子力學

中典型的互補變數，它們受到一些相依的限制，即當一個人要想對其中之一得到多些的話，他就被迫對另一個犧牲一些。

也許意思可以被證明像似和在熱力學裡一個集團的熵相依賴的量。在定理中熵所呈現的面貌無疑地是最有趣也最具意義的了。愛丁頓也已為這關係作了註解，然而在此所錄的是在「物理世界的本性」(The nature of the physical world)中摘出者，似乎是特別易懂且有建議性：

假設，我們要把下列之物安排為兩類：距離、質量、電力、熵、美麗、旋律。我想有充份的理由把熵和美麗、旋律歸為一類，而不和前三者相題並論。熵只有在審視各部份之間的關係時才出現，而且只有在洞悉各部份之間的美和旋律時才能被看見、被聽到。它們三個都是安排的美。這使人懷著要使其中之一能成為科學中常現的量的想法。為什麼這個陌生人能脫離他的原居住地到物理世界裡來，是它能夠說他們的話，即算數的語言。

我想愛丁頓應該會願意把意思這個詞加到美和旋律一起去；而且我猜想他會興奮得發抖，設若他看到在這個定理中熵不僅是能說算數的語言，而且會說語言的語言。