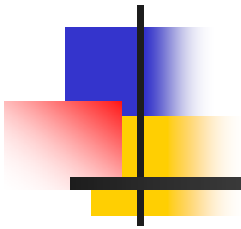
Trường Đại học Bách khoa
Khoa Công Nghệ Thông Tin



BÀI GIẢNG MÔN HỌC
LÝ THUYẾT THÔNG TIN

Giảng Viên: Hồ Văn Quân
E-mail: hcquan@dit.hcmut.edu.vn
Web site: http://www.dit.hcmut.edu.vn/~hcquan/student.htm

NỘI DUNG MÔN HỌC

- Bài 1 Giới thiệu
- Bài 2 Một số khái niệm cơ bản
- Bài 3 Chuẩn bị toán học
- Bài 4 Lượng tin
- Bài 5 Entropy
- Bài 6 Mã hiệu
- Bài 7 Mã hóa tối ưu nguồn rời rạc không nhớ
- Bài 8 Mã hóa nguồn phổ quát
- Bài 9 Kênh rời rạc không nhớ, lượng tin tương hỗ

NỘI DUNG MÔN HỌC (tt)

- Bài 10 Mã hóa chống nhiễu, định lý kênh
- Bài 11 Mã khối tuyến tính
- Bài 12 Cơ sở toán học của mã hóa chống nhiễu
- Bài 13 Mã vòng
- Bài 14 Giới thiệu về mật mã hóa
- Bài 15 Một số vấn đề nâng cao

TÀI LIỆU THAM KHẢO

1. Information Theory - Robert B.Ash, Nhà xuất bản Dover, Inc, 1990.
2. Introduction to Information Theory - Masud Mansuripur, Nhà xuất bản Prentice–Hall, Inc, 1987.
3. A Mathematical Theory of Communication - C. E. Shannon, Tạp chí Bell System Technical, số 27, trang 379–423 và 623–656, tháng 7 và tháng 10, 1948.
4. Cơ sở Lý thuyết truyền tin (tập một và hai) - Đặng Văn Chuyết, Nguyễn Tuấn Anh, Nhà xuất bản Giáo dục, 1998.

HÌNH THỨC ĐÁNH GIÁ

- Sẽ có thông báo cụ thể cho từng khóa học. Tuy nhiên, thường là có hình thức như bên dưới.
- Thi trắc nghiệm
  - Giữa kỳ: **30** câu / **45** phút
  - Cuối kỳ: **45** câu / **90** phút
  - Được phép xem tài liệu trong 2 tờ giấy A4
- Làm bài tập lớn cộng điểm (không bắt buộc)
  - Nộp bài tập lớn và báo cáo vào cuối học kỳ
  - Cộng tối đa 2 điểm

CÁC MÔN LIÊN QUAN

- Lý thuyết xác suất
- Kỹ thuật truyền số liệu
- Xử lý tín hiệu số

Bài 1 Giới thiệu

1.1 Thông tin là gì?

1.2 Vai trò của thông tin

1.3 Lý thuyết thông tin nghiên cứu những gì?

1.4 Những ứng dụng của lý thuyết thông tin

1.5 Lý thuyết thông tin – Lịch sử hình thành và quan điểm khoa học hiện đại

Thông tin là gì?

- Một vài ví dụ
  - Hai người nói chuyện với nhau. Cái mà trao đổi giữa họ gọi là thông tin.
  - Một người đang xem tivi/nghe đài/đọc báo, người đó đang nhận thông tin từ đài phát/báo.
  - Quá trình giảng dạy trong lớp.
  - Các máy tính nối mạng và trao đổi dữ liệu với nhau.

  - Máy tính nạp chương trình, dữ liệu từ đĩa cứng vào RAM để thực thi.

Thông tin là gì? (tt)

- Nhận xét
  - Thông tin là cái được truyền từ đối tượng này đến đối tượng khác để báo một "*điều*" gì đó. Thông tin chỉ có ý nghĩa khi "điều" đó bên nhận chưa biết.
  - Thông tin xuất hiện dưới nhiều dạng *âm thanh*, *hình ảnh*, ... Những dạng này chỉ là "*vỏ bọc*" vật chất chứa thông tin. "Vỏ bọc" là phần "xác", thông tin là phần "hồn".
  - *Ngữ nghĩa* của thông tin chỉ có thể hiểu được khi bên nhận hiểu được cách biểu diễn ngữ nghĩa của bên phát.
  - Một trong những phương tiện để diễn đạt thông tin là *ngôn ngữ*.
  - Có hai trạng thái của thông tin: *truyền* và *lưu trữ*. Môi trường truyền/lưu trữ được gọi chung là môi trường chứa tin hay *kênh tin*.

Vai trò của thông tin

- Các đối tượng sống luôn luôn có nhu cầu hiểu về thế giới xung quanh, để thích nghi và tồn tại. Đây là một quá trình quan sát, tiếp nhận, trao đổi và xử lý thông tin từ môi trường xung quanh.
- Thông tin trở thành một nhu cầu cơ bản, một điều kiện cần cho sự tồn tại và phát triển.
- Khi KHKT, XH ngày càng phát triển, thông tin càng thể hiện được vai trò quan trọng của nó đối với chúng ta.
- Ví dụ, hành động xuất phát từ suy nghĩ, nếu suy nghĩ đúng, thì hành động mới đúng. Suy nghĩ lại chịu ảnh hưởng từ các nguồn thông tin được tiếp nhận. Vì vậy thông tin có thể chi phối đến suy nghĩ và kết quả là hành động của con người.

LTTT nghiên cứu những vấn đề gì?

- Ở góc độ khoa học kỹ thuật, LTTT nghiên cứu nhằm tạo ra một "cơ sở hạ tầng" tốt cho việc truyền thông tin chính xác, nhanh chóng và an toàn; lưu trữ thông tin một cách hiệu quả.
- Ở các góc độ nghiên cứu khác LTTT nghiên cứu các vấn đề về cách tổ chức, biểu diễn và truyền đạt thông tin, và tổng quát là các vấn đề về xử lý thông tin.

- Ba lĩnh vực nghiên cứu cơ bản của môn học
  - **Mã hoá chống nhiễu**
  - **Mã hoá tối ưu (hay nén dữ liệu)**
  - **Mật mã hoá**

Những ứng dụng của LT thông tin

- Cuộc cách mạng thông tin đang xảy ra, sự phát triển mạnh mẽ của các phương tiện mới về truyền thông, lưu trữ thông tin làm thay đổi ngày càng sâu sắc xã hội chúng ta.
- LTTT đóng một vai trò quyết định trong sự phát triển này bằng cách cung cấp cơ sở lý thuyết và một cái nhìn triết học sâu sắc đối với những bài toán mới và thách thức mà chúng ta chạm trán – hôm nay và mai sau.
- Những ứng dụng phổ biến của LTTT là **truyền thông và xử lý thông tin** bao gồm: *truyền thông*, *nén*, *bảo mật*, *lưu trữ*, ...
- Các ý tưởng của LTTT đã được áp dụng trong nhiều lĩnh vực như *vật lý*, *ngôn ngữ học*, *sinh vật học*, *khoa học máy tính*, *tâm lý học*, *hóa học*

Những ứng dụng của LT thông tin (tt)

- Mối quan hệ giữa LTTT và *thống kê* đã được tìm thấy, các phương pháp mới về phân tích thống kê dựa trên LTTT đã được đề nghị.
- Ứng dụng vào *quản lý kinh tế*. Ví dụ, lý thuyết đầu tư tối ưu xuất hiện đồng thời với lý thuyết mã hóa nguồn tối ưu.
- Ứng dụng vào *ngôn ngữ học*.
- Ứng dụng đến *tâm lý thực nghiệm* và đặc biệt là lĩnh vực dạy và học.

Lịch sử hình thành

- Cuộc cách mạng lớn nhất về cách nhìn thế giới khoa học là chuyển hướng từ *thuyết quyết định Laplacian* đến *bức tranh xác suất của tự nhiên*.
- Thế giới chúng ta đang sống trong đó chủ yếu là *xác suất*. Kiến thức của chúng ta cũng là một dạng xác suất.
- LTTT nổi lên sau khi cơ học thống kê và lượng tử đã phát triển, và nó chia xẻ với vật lý thống kê các khái niệm cơ bản về entropy.
- Theo lịch sử, các khái niệm cơ bản của LTTT như entropy, thông tin tương hỗ được hình thành từ việc nghiên cứu các *hệ thống mật mã* hơn là từ việc nghiên cứu các *kênh truyền thông*.
- Về mặt toán học, LTTT là một nhánh của *lý thuyết xác suất* và các *quá trình ngẫu nhiên* (stochastical process).

Lịch sử hình thành (tt)

- Quan trọng và có ý nghĩa nhất là quan hệ liên kết giữa LTTT và *vật lý thống kê*.
- Trong một thời gian dài trước khi LTTT được hình thành, L. Boltzman và sau đó là L.Szilard đã đánh đồng ý nghĩa của thông tin với khái niệm nhiệt động học của entropy. Một mặt khác, D. Gabor chỉ ra rằng "lý thuyết truyền thông phải được xem như một nhánh của vật lý".
- *C. E. Shannon* là cha đẻ của LTTT.

Bài 2 Một số khái niệm cơ bản

2.1 Thông tin (Information)

2.2 Mô hình của các quá trình truyền tin

2.3 Các loại hệ thống truyền tin – Liên tục và rời rạc

2.4 Rời rạc hoá

Thông tin

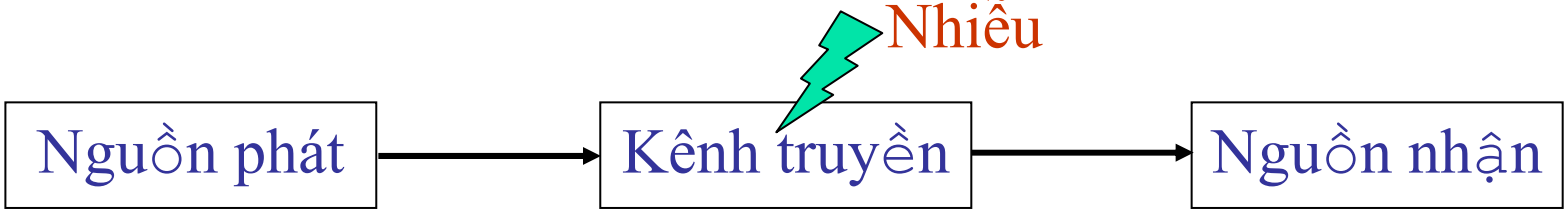
- Thông tin là một khái niệm trừu tượng, phi vật chất và rất khó được định nghĩa chính xác. Hai định nghĩa về thông tin.
- *Thông tin là sự cảm hiểu của con người về thế giới xung quanh thông qua sự tiếp xúc với nó.*
- *Thông tin là một hệ thống những tin báo và mệnh lệnh giúp loại trừ sự không chắc chắn (uncertainty) trong trạng thái của nơi nhận tin. Nói ngắn gọn, thông tin là cái mà loại trừ sự không chắc chắn.*
- Định nghĩa đầu chưa nói lên được bản chất của thông tin. Định nghĩa thứ hai nói rõ hơn về bản chất của thông tin và được dùng để định lượng thông tin trong kỹ thuật.

Thông tin (tt)

- Thông tin là một hiện tượng vật lý, nó thường tồn tại và được truyền đi dưới một dạng vật chất nào đó.
- Những dạng vật chất dùng để mang thông tin được gọi là *tín hiệu*.
- *Lý thuyết tín hiệu* nghiên cứu các dạng tín hiệu và cách truyền thông tin đi xa với chi phí thấp, một ngành mà có quan hệ gần gũi với LTTT.
- Thông tin là một quá trình *ngẫu nhiên*.
- Tín hiệu mang tin tức cũng là tín hiệu ngẫu nhiên và mô hình toán học của nó là các quá trình ngẫu nhiên thực hay phức.
- Và LTTT là *lý thuyết ngẫu nhiên của tin tức*, có nghĩa là nó xét đến tính *bất ngờ* của tin tức đối với nơi nhận tin.

Mô hình của các quá trình truyền tin

- Khái niệm thông tin thường đi kèm với một hệ thống truyền tin.



- Sự truyền tin (transmission)
  - Là sự dịch chuyển thông tin từ điểm này đến điểm khác trong một môi trường xác định.
- Nguồn tin (information source)
  - Là một tập hợp các tin mà hệ thống truyền tin dùng để lập các bảng tin hay thông báo (message) để truyền tin.
  - Bảng tin chính là dãy tin được bên phát truyền đi.
  - Thông tin có thể thuộc nhiều loại như

    (1) một dãy kí tự như trong điện tín (telegraph) của các hệ thống gởi điện tín (teletype system);

Mô hình của các quá trình truyền tin (tt)

(2) một hàm theo chỉ một biến thời gian $f(t)$ như trong radio và điện thoại;

(3) một hàm của thời gian và các biến khác như trong tivi trắng đen – ở đây thông tin có thể được nghĩ như là một hàm $f(x, y, t)$ của toạ độ hai chiều và thời gian biểu diễn cường độ ánh sáng tại điểm (x, y) trên màn hình và thời gian t;

(4) một vài hàm của một vài biến như trong trường hợp tivi màu – ở đây thông tin bao gồm ba hàm $f(x, y, t)$, $g(x, y, t)$, $h(x, y, t)$ biểu diễn cường độ ánh sáng của các ba thành phần màu cơ bản (xanh lá cây, đỏ, xanh dương)

- Thông tin trước khi được truyền đi, tuỳ theo yêu cầu có thể được mã hoá để nén, chống nhiễu, bảo mật, ...

- Kênh tin (channel)
  - Là nơi hình thành và truyền (hoặc lưu trữ) tín hiệu mang tin đồng thời ở đấy xảy ra các tạp nhiễu (noise) phá hủy tin tức.
  - Trong LTTT kênh là một khái niệm trừu tượng đại biểu cho hỗn hợp tín hiệu và tạp nhiễu.

Một số khái niệm (tt)

- Môi trường truyền tin thường rất đa dạng
  - môi trường không khí, tin được truyền dưới dạng âm thanh và tiếng nói, ngoài ra cũng có thể bằng lửa hay bằng ánh sáng;
  - môi trường tầng điện ly trong khí quyển nơi mà thường xuyên xảy ra sự truyền tin giữa các vệ tinh nhân tạo với các trạm rada ở dưới mặt đất;
  - đường truyền điện thoại nơi xảy ra sự truyền tín hiệu mang tin là dòng điện hay đường truyền cáp quang qua biển trong đó tín hiệu mang tin là sóng ánh sáng v.v…
- Nhiễu (noise)
  - Cho dù môi trường nào cũng có nhiễu. Nhiễu rất phong phú và đa dạng và thường đi kèm với môi trường truyền tin tương ứng.
    - Chẳng hạn nếu truyền dưới dạng sóng điện từ mà có đi qua các vùng của trái đất có từ trường mạnh thì tín hiệu mang tin thường bị ảnh hưởng ít nhiều bởi từ trường này. Nên có thể coi từ trường này là một loại nhiễu.
    - Nếu truyền dưới dạng âm thanh trong không khí thì tiếng ồn xung quanh có thể coi là một loại nhiễu.

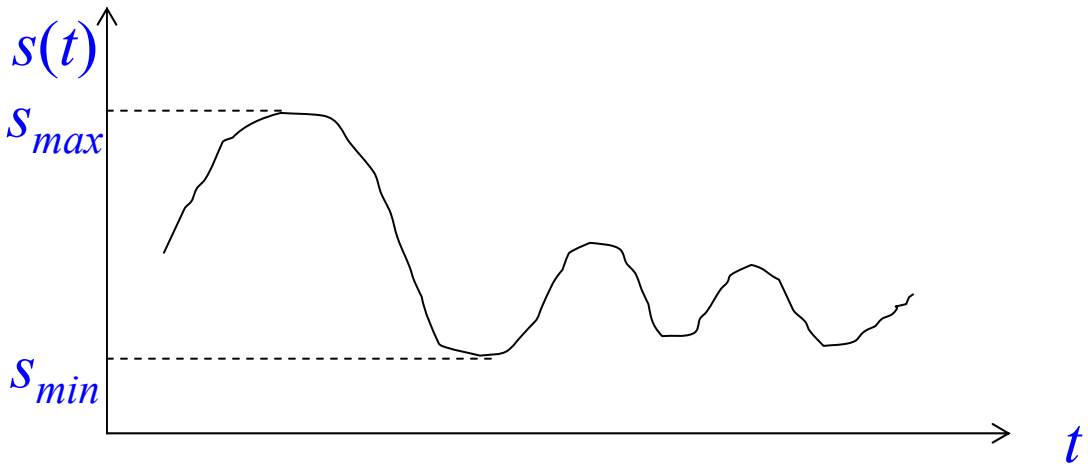
Một số khái niệm (tt)

- Nhiễu có nhiều loại chẳng hạn nhiễu cộng, nhiễu nhân.
- Nhiễu cộng là loại nhiễu mà tín hiệu mang tin bị tín hiệu nhiễu "cộng" thêm vào.
- Nhiễu nhân là loại nhiễu mà tín hiệu mang tin bị tín hiệu nhiễu "nhân" lên.

- Nơi nhận tin (sink)
  - Là nơi tiếp nhận thông tin từ kênh truyền và cố gắng khôi phục lại thành thông tin ban đầu như bên phát đã phát đi.
  - Tin đến được nơi nhận thường không giống như tin ban đầu vì có sự tác động của nhiễu. Vì vậy nơi nhận phải thực hiện việc *phát hiện sai* và *sửa sai*.
  - Nơi nhận còn có thể phải thực hiện việc giải nén hay giải mã thông tin đã được mã hoá bảo mật nếu như bên phát đã thực hiện việc nén hay bảo mật thông tin trước khi truyền

Các loại hệ thống truyền tin

- Các nguồn tin thường thấy trong tự nhiên được gọi là các *nguồn tin nguyên thuỷ*. Đây là các nguồn tin chưa qua bất kỳ một phép biến đổi nhân tạo nào.
- Các tín hiệu âm thanh, hình ảnh được phát ra từ các nguồn tin nguyên thuỷ này thường là các *hàm liên tục theo thời gian và theo mức*, nghĩa là có thể biểu diễn một thông tin nào đó dưới dạng một hàm $s(t)$ tồn tại trong một quãng thời gian T và lấy các trị bất kỳ trong một phạm vi ($smin$, $smax$) nào đó.



Các loại hệ thống truyền tin (tt)

- Các nguồn như vậy được gọi là các *nguồn liên tục* (continuous source), các tin được gọi là *tin liên tục* (continuous information) và kênh tin được gọi là *kênh liên tục* (continuous channel).
- Tuy nhiên vẫn có những nguồn nguyên thuỷ là rời rạc
  - Bảng chữ cái của một ngôn ngữ.
  - Các tin trong hệ thống điện tín, các lệnh điều khiển trong một hệ thống điều khiển, ...
- Trong trường hợp này các nguồn được gọi là *nguồn rời rạc* (discrete source), các tin được gọi là *tin rời rạc* (discrete information) và kênh tin được gọi là *kênh rời rạc* (discrete channel).
- Sự phân biệt về bản chất của tính rời rạc và tính liên tục là số lượng tin của nguồn trong trường hợp rời rạc là *hữu hạn* còn trong trường hợp liên tục là *không đếm được*.

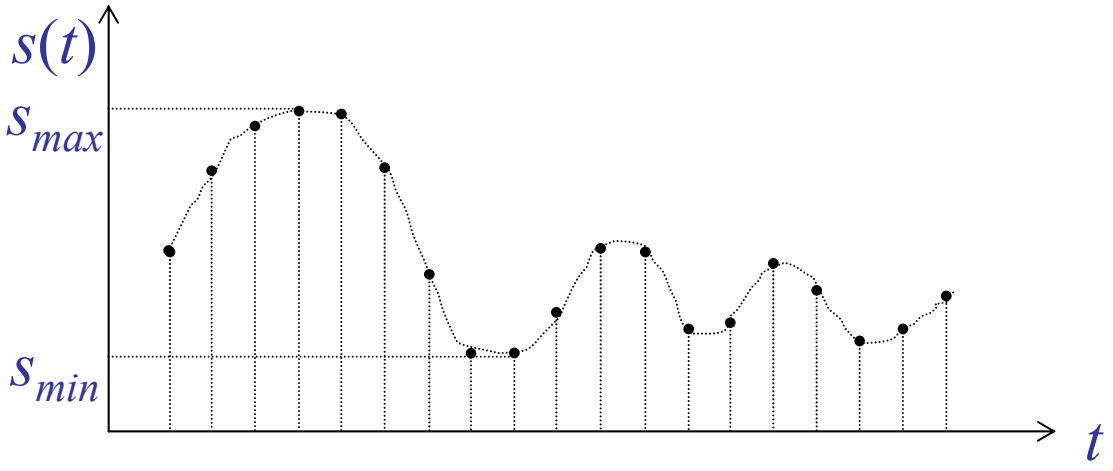
Rời rạc hóa

- Các hệ thống liên tục có nhiều nhược điểm của như cồng kềnh, không hiệu quả, và chi phí cao.
- Các hệ thống truyền tin rời rạc có nhiều ưu điểm hơn, khắc phục được những nhược điểm trên của các hệ thống liên tục và đặc biệt đang ngày càng được phát triển và hoàn thiện dần những sức mạnh và ưu điểm của nó.
- Rời rạc hoá thường bao gồm hai loại: Rời rạc hoá theo trục thời gian, còn được gọi là *lấy mẫu* (sampling) và rời rạc hoá theo biên độ, còn được gọi là *lượng tử hoá* (quantize).

- Lấy mẫu (Sampling)
  - Lấy mẫu một hàm là trích ra từ hàm ban đầu các mẫu được lấy tại những thời điểm xác định.
  - Vấn đề là làm thế nào để sự thay thế hàm ban đầu bằng các mẫu này là một sự thay thế tương đương, điều này đã được giải quyết bằng định lý lấy mẫu nổi tiếng của Shannon.

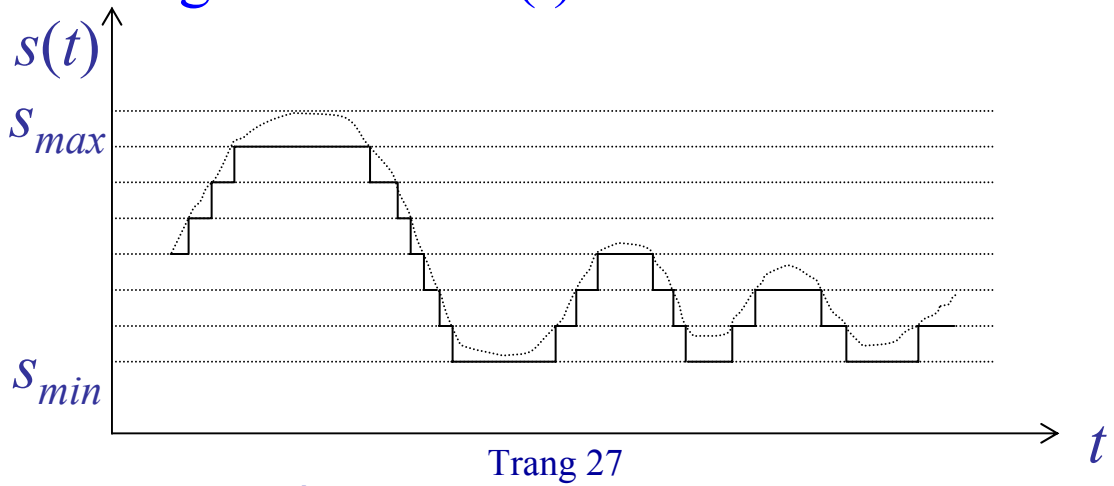
Rời rạc hóa (tt)

- Định lý lấy mẫu của Shannon
  - Một hàm $s(t)$ có phổ hữu hạn, không có thành phần tần số lớn hơn ω_{max} ($= 2\pi f_{\max}$) có thể được thay thế bằng các mẫu của nó được lấy tại những thời điểm cách nhau một khoảng $\Delta t \leq \pi/\omega_{max}$, hay nói cách khác tần số lấy mẫu $F \geq 2f_{max}$.



Rời rạc hóa (tt)

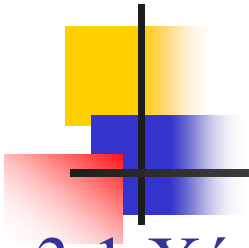
- Lượng tử hoá (Quantize)
  - Biên độ của các tín hiệu thường là một miền liên tục (s_{min}, s_{max}). Lượng tử hoá là phân chia miền này thành một số mức nhất định, chẳng hạn là $s_{min} = s_0, s_1, ..., s_n = s_{max}$ và qui các giá trị biên độ không trùng với các mức này về mức gần với nó nhất.
  - Việc lượng tử hoá sẽ biến đổi hàm $s(t)$ ban đầu thành một hàm $s'(t)$ có dạng hình bậc thang. Sự khác nhau giữa $s(t)$ và $s'(t)$ được gọi là sai số lượng tử. Sai số lượng tử càng nhỏ thì $s'(t)$ biểu diễn càng chính xác $s(t)$.



Nguồn rời rạc

- Nguồn tin liên tục sau khi được lấy mẫu và lượng tử hoá sẽ trở thành *nguồn rời rạc*.
- Chúng ta học chủ yếu các nguồn rời rạc.

- Nguồn rời rạc
  - Một nguồn rời rạc là một bảng chữ cái A gồm m kí hiệu, $A = \{a_1, a_2, ..., a_m\}$, với những xác suất xuất hiện $p(a_i)$, $i = 1, .., m$.
  - Định nghĩa không diễn tả mối quan hệ giữa tin trước và sau trong một bản tin, nên đây được gọi là một *nguồn rời rạc không nhớ* (discrete memoryless source).
- Bảng tin của một nguồn tin rời rạc không nhớ
  - Là một dãy (có thể vô hạn) các kí hiệu liên tiếp từ bảng chữ cái của nguồn tin, $x = (... a_{-2}a_{-1}a_0a_1a_2...)$
  - Trong thực tế bảng tin có bắt đầu và kết thúc cho nên bảng tin là một dãy hữu hạn các kí hiệu, $x^* = (a_1a_2 \ldots a_n)$



Bài 3 Chuẩn bị toán học

3.1 Xác suất (Probability)

3.2 Bất đẳng thức Chebyshev và luật yếu của số lớn

3.3 Tập lồi (Convex sets) và hàm lồi (convex functions), bất đẳng thức Jensen

3.4 Công thức Stirling

Xác suất

- Không gian mẫu (Sample space)
  - Là tập (hay không gian) tất cả các kết quả có thể có của một thí nghiệm. Thường được kí hiệu là E hay S. Nếu không gian mẫu là rời rạc thì E có thể được biểu diễn bằng $E = \{e_1, e_2, ..., e_n\}$
- Sự kiện (Event), sự kiện cơ bản (elementary event)
  - Mỗi tập con của E (không gian mẫu) được gọi là một sự kiện, đặc biệt mỗi phần tử của E được gọi là một sự kiện cơ bản.
- Ví dụ
  - Trong một thí nghiệm tung đồng xu thì E = {U (úp), N (ngửa)}. Nếu đồng tiền là đồng nhất thì xác suất $P(\text{U}) = P(\text{N}) = 1/2$.
  - Trong một thí nghiệm tung con xúc xắc thì E = {1, 2, 3, 4, 5, 6}. Nếu con xúc xắc là đồng nhất thì xác suất $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$, $P(2, 5) = 1/3$, $P(1, 3, 5) = 1/2$.

Xác suất (tt)

- Lấy một văn bản tiếng Anh điển hình và nhặt một kí tự bất kỳ thì $E = \{a, b, c, ..., x, y, z\}$ và xác suất của các kí tự được phân bố như sau $P(a) = 0{,}0642$, ..., $P(e) = 0{,}103$, ..., $P(z) = 0{,}0005$.

- Biến ngẫu nhiên rời rạc (Discrete random variable)
  - Một biến ngẫu nhiên rời rạc x được định nghĩa bằng cách gán một số thực x_i tới mỗi sự kiện cơ bản e_i của không gian mẫu rời rạc E. Xác suất của x_i được định nghĩa là xác suất của sự kiện cơ bản tương ứng và được kí hiệu là $p(x_i)$.
- Trị trung bình (kỳ vọng) (average, expected value), phương sai (variance)
  - Trị trung bình và phương sai của biến ngẫu nhiên rời rạc x lần lượt được kí hiệu và định nghĩa như sau
  - $E(\mathrm{x}) \quad = \bar{\mathrm{x}} = \sum_i \mathrm{x}_i p(\mathrm{x}_i)$

Xác suất (tt)

- $\text{Var}(\mathbf{x}) = E\left(\left(\mathrm{x}-\bar{\mathrm{x}}\right)^2\right) = \sum_i \left(\mathrm{x}_i-\bar{\mathrm{x}}\right)^2 p\left(\mathrm{x}_i\right)$

  $= E\left(\mathrm{x}^2\right) - \bar{\mathrm{x}}^2$

  trong đó $E(\mathbf{x}^2)$ là trị kỳ vọng của $\mathbf{x}^2$.
  - Tổng quát, trị kỳ vọng của một hàm của **x**, chẳng hạn $f(\mathbf{x})$, được định nghĩa bằng

$$E\left(f\left(\mathrm{x}\right)\right) = \sum_i f\left(\mathrm{x}_i\right) p\left(\mathrm{x}_i\right)$$

- Xác suất đồng thời (joint probability), xác suất có điều kiện (conditional probability)
  - Một cặp biến ngẫu nhiên (x, y) liên kết với một thí nghiệm tạo thành một biến ngẫu nhiên nối (joint random variable). Nếu x, y là rời rạc, sự phân bố xác suất nối hay xác suất đồng thời được định nghĩa là

$$p_{ij} = \mathrm{P}(\mathrm{x} = x_i, \mathrm{y} = y_j)$$

Xác suất (tt)

- Xác suất của y trong điều kiện đã biết x được gọi là xác suất có điều kiện và được định nghĩa là

$$p(y_j|x_i) = \frac{p(x_i, y_j)}{p(x_i)}$$

trong đó **xác suất lề** (marginal probability) $p(x_i)$ được giả thiết là khác không.

- Các xác suất lề được định nghĩa như sau:

$$p(x_i) = \sum_j p(x_i, y_j)$$

$$p(y_j) = \sum_i p(x_i, y_j)$$

Ví dụ

- Thí nghiệm tung đồng thời một đồng xu và con xúc xắc.
- Từ kết quả trên ta thấy

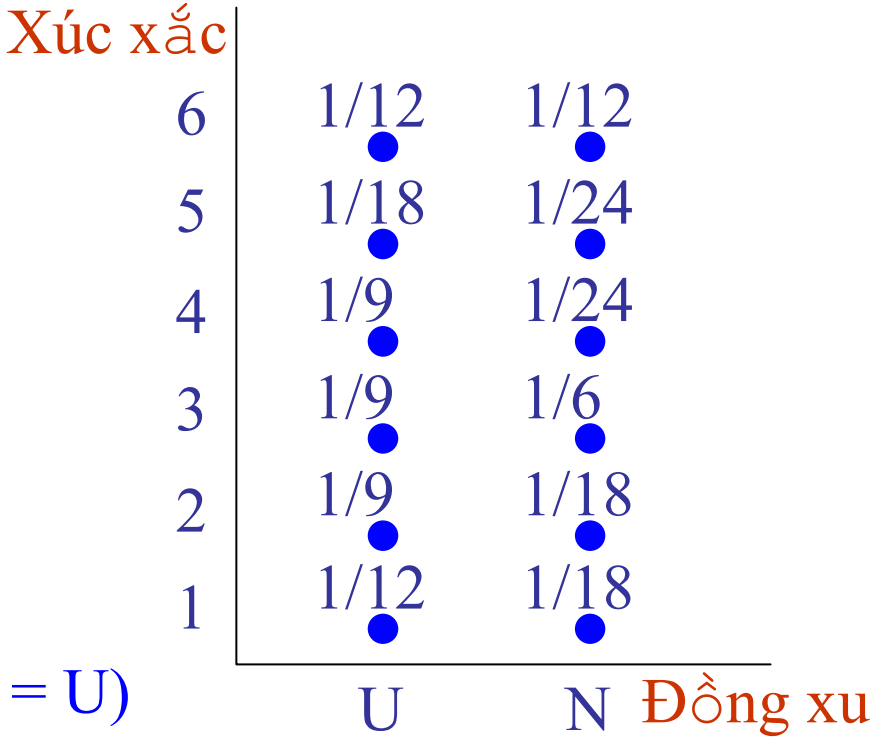
P(U, 5) = 1/18

P(Đồng xu = U) = 5/9

P(Đồng xu = N) = 4/9

P(Xúc xắc = 5) = 7/72

P(Xúc xắc = 5 đã biết Đồng xu = U)



Xác suất (tt)

- Sự độc lập (Independence)
  - Hai biến ngẫu nhiên x và y được gọi là độc lập nếu

$$p(x_i, y_j) = p(x_i)p(y_j) \ \forall \ i, j.$$

  - Chúng ta thấy nếu hai biến x và y độc lập thì

$$p\left(y_j \middle| x_i\right) = \frac{p\left(x_i, y_j\right)}{p\left(x_i\right)} = \frac{p\left(x_i\right)p\left(y_j\right)}{p\left(x_i\right)} = p\left(y_j\right)$$

    có nghĩa là xác suất y_j trong điều kiện có x_i xảy ra hay không xảy ra đều như nhau, không thay đổi, và ngược lại.
  - Cũng từ sự độc lập chúng ta suy ra một kết quả mà hay được sử dụng sau này

$$E(\mathrm{xy}) = \mathrm{E(x)\ E(y)} = \overline{\mathrm{x}}\,\overline{\mathrm{y}}$$

Xác suất (tt)

- Sự tương quan (correlation)
  - Sự tương quan C giữa hai biến x và y được định nghĩa là trị kỳ vọng của $(x - \bar{x})(y - \bar{y})$:

$$C(x, y) = E((x - \bar{x})(y - \bar{y})) =$$
$$= E(xy) - \bar{x}\,\bar{y}$$

  - Trong trường hợp x và y là độc lập chúng ta suy ra $C(x, y) = 0$. Tuy nhiên điều ngược lại thì không đúng.

Bất đẳng thức Chebyshev và luật yếu của số lớn

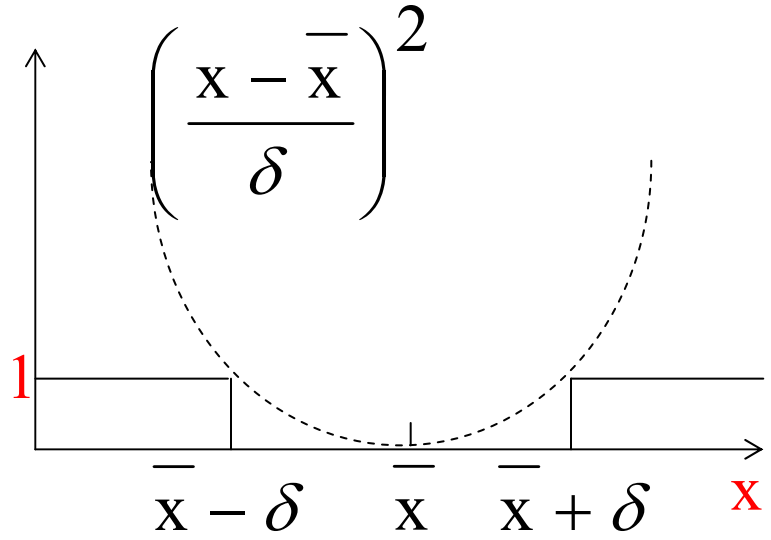
- Bất đẳng thức Chebyshev
  - Cho một biến ngẫu nhiên **x** có trị trung bình là $\overline{x}$ và phương sai là δ_x^2, bất đẳng thức Chebyshev đối với một số dương tuỳ ý δ là

$$P(|\mathbf{x} - \overline{x}| \geq \delta) \leq \frac{\delta_x^2}{\delta^2}$$

- Chứng minh
  - Định nghĩa một hàm $f(x)$ như sau $f(x) = \begin{cases} 1, |x - \overline{x}| \geq \delta \\ 0, |x - \overline{x}| < \delta \end{cases}$
  - Thì

$$P(|\mathbf{x} - \overline{x}| \geq \delta) = \Sigma f(x_i)p(x_i)$$

Bất đẳng thức Chebyshev (tt)



- Dựa trên hình chúng ta có

$$f(x) \le \left(\frac{x-\bar{x}}{\delta}\right)^2$$

- Vì vậy,

$$P\left(\left|x-\bar{x}\right| \ge \delta\right) \le \sum_i \left(\frac{x-\bar{x}}{\delta}\right)^2 p(x_i) = \frac{\delta_x^2}{\delta^2}$$

Luật yếu của số lớn (tt)

- Xét một thí nghiệm nhị phân trong đó các kết quả của thí nghiệm là 0 và 1 với các xác suất tương ứng là p_0 và $1-p_0$.
- Thí nghiệm này được lặp lại N lần một cách độc lập, và kết quả trung bình được định nghĩa là $\mathbf{y}_N$; tức là, $\mathbf{y}_N$ bằng tổng số các số 1 trong N lần thí nghiệm chia cho N.
- Rõ ràng, $\mathbf{y}_N$ là một biến ngẫu nhiên có không gian mẫu là $\{0, 1/N, 2/N, ..., 1\}$.
- Định nghĩa $\mathbf{x}^{(n)}$ là biến ngẫu nhiên tương ứng với kết quả của lần thí nghiệm thứ n, chúng ta có

$$\mathrm{y}_N = \frac{1}{N}\sum_{n=1}^{N} \mathrm{x}^{(n)}$$

Luật yếu của số lớn (tt)

$$\bar{y}_N = \frac{1}{N}\sum_{n=1}^{N} E\left(x^{(n)}\right) = \frac{1}{N}\sum_{n=1}^{N} \bar{x} = \bar{x}$$

$$\delta_y^2 = E\left(\left(y_N - \bar{y}_N\right)^2\right) = E\left(\left[\frac{1}{N}\sum_{n=1}^{N} x^{(n)} - \bar{x}\right]^2\right)$$

$$= E\left(\left(\frac{1}{N}\left[\sum_{n=1}^{N} x^{(n)} - N\bar{x}\right]\right)^2\right) = \frac{1}{N^2} E\left(\left[\sum_{n=1}^{N}\left(x^{(n)} - \bar{x}\right)\right]^2\right)$$

$$= \frac{1}{N^2}\sum_{n=1}^{N} E\left(\left(x^{(n)} - \bar{x}\right)^2\right) = \frac{1}{N^2} N\delta_x^2 = \frac{\delta_x^2}{N}$$

Luật yếu của số lớn (tt)

- Đối với một số nguyên dương tuỳ ý ε, theo bất đẳng thức Chebyshev chúng ta có

$$P\left(\left| \mathrm{y}_N - \bar{\mathrm{y}}_N \right| \geq \varepsilon\right) \leq \frac{\delta_{\mathrm{y}}^2}{\varepsilon^2}$$

từ đây chúng ta dẫn ra được luật yếu của số lớn

$$P\left(\left|\left[\frac{1}{N}\sum_{n=1}^{N} \mathrm{x}^{(n)}\right] - \bar{\mathrm{x}}\right| \geq \varepsilon\right) \leq \frac{\delta_{\mathrm{x}}^2}{N\varepsilon^2}$$

- Chú ý rằng vế phải tiến tới 0 khi N tiến ra vô cùng.
- Luật yếu của số lớn vì vậy khẳng đinh rằng trị trung bình mẫu của **x** tiếp cận trị trung bình thống kê với xác suất cao khi $N \to \infty$.

Tập lồi

- Trong không gian Ơclit, một tập S được gọi là lồi (convex cap ($\cap$)) nếu đối với một cặp điểm P_1, P_2 thuộc S thì mọi điểm thuộc đoạn P_1P_2 cũng thuộc S.

(a) (b)

- Nếu $P_1 = (x_1, x_2, ..., x_n)$ và $P_2 = (y_1, y_2, ..., y_n)$ là các điểm trong không gian Ơclit n chiều, thì đoạn thẳng nối chúng được biểu diễn bằng tập các điểm P, trong đó

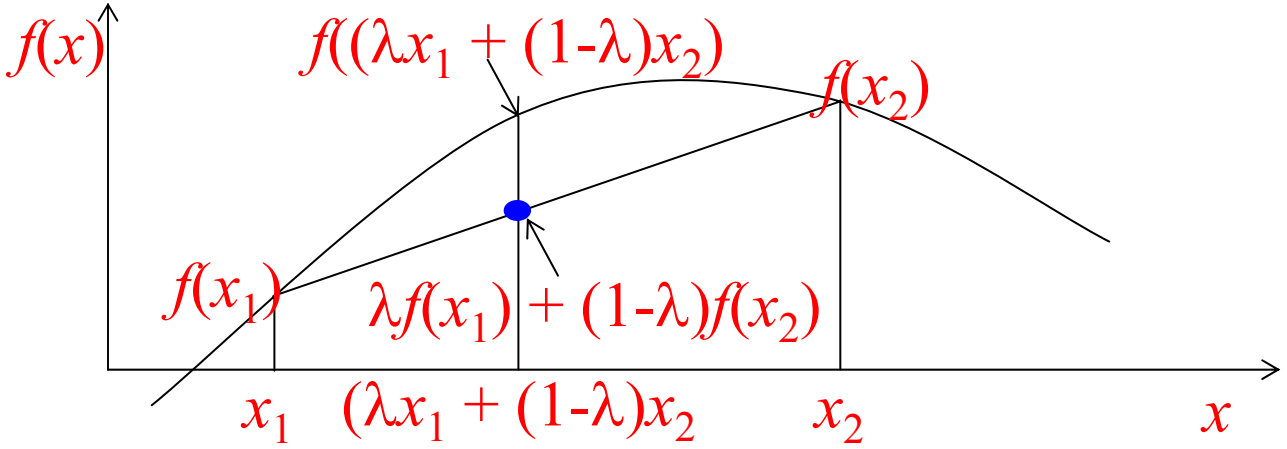
$$P = \lambda P_1 + (1-\lambda)P_2$$

$$= (\lambda x_1 + (1-\lambda)y_1, \lambda x_2 + (1-\lambda)y_2, ..., \lambda x_n + (1-\lambda)y_n) \text{ và } \lambda \in [0, 1].$$

Hàm lồi

- Một ví dụ quan trọng của tập lồi là tập tất cả các điểm (p_1, p_2, ..., p_n) trong đó (p_1, p_2, ..., p_n) là một sự phân bố xác suất (tức là các $p_i \in [0, 1]$ và $\Sigma p_i = 1$).
- Một hàm thực $f(P)$, được định nghĩa trên tập lồi S, được gọi là lồi nếu $\forall$cặp điểm $P_1, P_2 \in S$, và $\forall\ \lambda \in [0, 1]$ bất đẳng thức sau đây đúng:

$$f(\lambda P_1 + (1-\lambda)P_2) \geq \lambda f(P_1) + (1-\lambda)f(P_2)$$



Định lý, bất đẳng thức Jensen

- Nếu $\lambda_1, ..., \lambda_N$ là các số không âm có tổng bằng 1 thì đối với mọi tập điểm $P_1, ..., P_N$ trong miền xác định của hàm lồi $f(P)$ bất đẳng thức sau đây đúng

$$f\left(\sum_{n=1}^{N}\lambda_n P_n\right) \geq \sum_{n=1}^{N}\lambda_n f\left(P_n\right)$$

- Cho biến ngẫu nhiên x lấy các giá trị $x_1, ..., x_n$ với các xác suất $p_1, ..., p_n$. Cho $f(x)$ là một hàm lồi có miền xác định chứa $x_1, ..., x_n$. Chúng ta có $E(\mathrm{x}) = \sum_i p_i x_i$ và $E(f(\mathrm{x})) = \sum_i p_i f\left(x_i\right)$.
- Áp dụng định lý trên chúng ta có

$$f(E(\mathrm{x})) \geq E(f(\mathrm{x}))$$

Đây được gọi là bất đẳng thức Jensen.

Bài 4 Lượng tin

4.1 Lượng tin

4.2 Lượng tin trung bình

Vấn đề cơ bản của truyền thông là việc tái sinh tại một điểm hoặc chính xác hoặc gần đúng một thông báo được chọn tại một điểm khác.

(Claude Shannon 1948)

Lượng tin

- Lượng tin (measure of information) dùng để so sánh định lượng các tin tức với nhau.
- **Một tin đối với người nhận đều mang hai nội dung, một là độ bất ngờ của tin, hai là ý nghĩa của tin**.
- Khía cạnh ngữ nghĩa chỉ có ý nghĩa đối với con người.
- Khía cạnh quan trọng nằm ở chỗ tin thật sự là một cái được chọn từ một **tập các tin** (tập các khả năng) có thể.
- Nếu số tin trong tập tin càng nhiều thì sẽ mang lại một "lượng tin" càng lớn khi nhận được một tin (giả sử các tin là bình đẳng như nhau về khả năng xuất hiện).
- Để sự truyền tin đạt hiệu quả cao chúng ta không thể đối đãi các tin như nhau nếu chúng xuất hiện ít nhiều khác nhau.

Lượng tin

- Xét một tin x có xác suất xuất hiện là $p(x)$, thì chúng ta có thể xem tin này như là một tin trong một tập có $1/p(x)$ tin với các tin có xác suất xuất hiện như nhau.
- Nếu $p(x)$ càng nhỏ thì $1/p(x)$ càng lớn và vì vậy "lượng tin" khi nhận được tin này cũng sẽ càng lớn.
- Vậy "lượng tin" của một tin tỉ lệ thuận với số khả năng của một tin và tỉ lệ nghịch với xác suất xuất hiện của tin đó.
- Xác suất xuất hiện của một tin tỉ lệ nghịch với độ bất ngờ khi nhận được một tin.

"lượng tin" ↑ số khả năng ↑ độ bất ngờ ↓ xác suất

- Một tin có xác suất xuất hiện càng nhỏ thì có độ bất ngờ càng lớn và vì vậy có lượng tin càng lớn.

Lượng tin (tt)

- Xét một nguồn $A = \{a_1, a_2, \ldots, a_m\}$ với các xác suất xuất hiện là $p(a_i)$ $i = 1, ..., m$.
- Kí hiệu lượng tin trong mỗi tin a_i là $I(a_i)$. Vậy hàm f dùng để biểu thị lượng tin phải thoã mãn những điều kiện gì?
- Phản ánh được các tính chất thống kê của tin tức.
  - Ví dụ có hai nguồn K, L với số tin tương ứng là k, l (giả thuyết đều là đẳng xác suất). Nếu $k > l$, thì độ bất ngờ khi nhận một tin bất kỳ của nguồn K phải lớn hơn độ bất ngờ khi nhận một tin bất kỳ của nguồn L, vậy $f(k) > f(l)$
- Hợp lý trong tính toán.
  - Giả thiết hai nguồn độc lập K và L với số tin tương ứng là k và l. Cho việc nhận một cặp k_i và l_j bất kỳ đồng thời là một tin của nguồn hỗn hợp KL. Số cặp $k_i l_j$ mà nguồn này có là $k*l$.

Lượng tin (tt)

- Độ bất ngờ khi nhận được một cặp như vậy phải bằng tổng lượng tin của khi nhận được k_i và l_j. Vì vậy chúng ta phải có:

$$f(kl) = f(k) + f(l)$$

- Khi nguồn chỉ có một tin, lượng tin chứa trong tin duy nhất đó phải bằng không.

$$f(1) = 0$$

- Định nghĩa
  - Lượng đo thông tin của một tin được đo bằng logarit của độ bất ngờ của tin hay nghịch đảo xác suất xuất hiện của tin đó.

$$I(x) = \log \frac{1}{p(x)} = -\log p(x)$$

Lượng tin (tt)

- Lượng tin chứa trong một dãy $x = a_1 a_2 \ldots a_n$ với $a_i \in A$ là

$$I(x) = \log \frac{1}{p(x)} = -\sum_{i=1}^{n} \log p(a_i)$$

- Trong trường hợp m kí hiệu của nguồn đẳng xác suất với nhau tức $p(a_i) = 1/m$ thì

$$I(a_i) = \log \frac{1}{p(a_i)} = \log m$$

Nếu $x = a_1 a_2 \ldots a_n$ với $a_i \in A$

$$I(x) = n \log m$$

Lượng tin trung bình

- Đơn vị của lượng tin
  - Nếu cơ số là 2 thì đơn vị là bits (cho các kí số nhị phân); nếu cơ số là *e* thì đơn vị là nats (cho đơn vị tự nhiên), nếu cơ số là 10 thì đơn vị là Hartley.
- Định nghĩa
  - Lượng tin trung bình của một nguồn tin *A* là lượng tin trung bình chứa trong một kí hiệu bất kỳ của nguồn tin. Nó thường được kí hiệu là *I*(*A*) và được tính bằng công thức sau

$$I(A) = \sum_{a_i \in A} p(a_i) I(a_i) = - \sum_{a_i \in A} p(a_i) \log p(a_i)$$

Ví dụ

- Cho một nguồn tin U bao gồm 8 tin $U = \{u_0, u_1, u_2, u_3, u_4, u_5, u_6, u_7\}$, với các xác suất xuất hiện như sau:

$p(u_0)$	$p(u_1)$	$p(u_2)$	$p(u_3)$	$p(u_4)$	$p(u_5)$	$p(u_6)$	$p(u_7)$
1/4	1/4	1/8	1/8	1/16	1/16	1/16	1/16

Hãy cho biết lượng tin riêng của mỗi tin và lượng tin trung bình của nguồn này trong đơn vị bits.

- Giải
  - Lượng tin riêng của mỗi tin là

$I(u_0)$	$I(u_1)$	$I(u_2)$	$I(u_3)$	$I(u_4)$	$I(u_5)$	$I(u_6)$	$I(u_7)$
2	2	3	3	4	4	4	4

Ví dụ (tt)

- Lượng tin trung bình của nguồn là

  $I(U) = (1/4) \times 2 + (1/4) \times 2 + (1/8) \times 3 + (1/8) \times 3 + (1/16) \times 4 + (1/16) \times 4 + (1/16) \times 4 + (1/16) \times 4 = 2{,}75$ bits.

- Điều này nói lên một ý nghĩa quan trọng rằng, chúng ta có thể biểu diễn mỗi tin trong nguồn U bằng một chuỗi có chiều dài trung bình là 2,75 bits. Nó sẽ tốt hơn so với trong trường hợp chúng ta không chú ý đến cấu trúc thông kê của nguồn. Lúc đó chúng ta sẽ biểu diễn mỗi tin trong 8 tin của nguồn bằng các chuỗi có chiều dài là 3 bits.

Bài 5 Entropy

5.1 Entropy của một biến ngẫu nhiên rời rạc

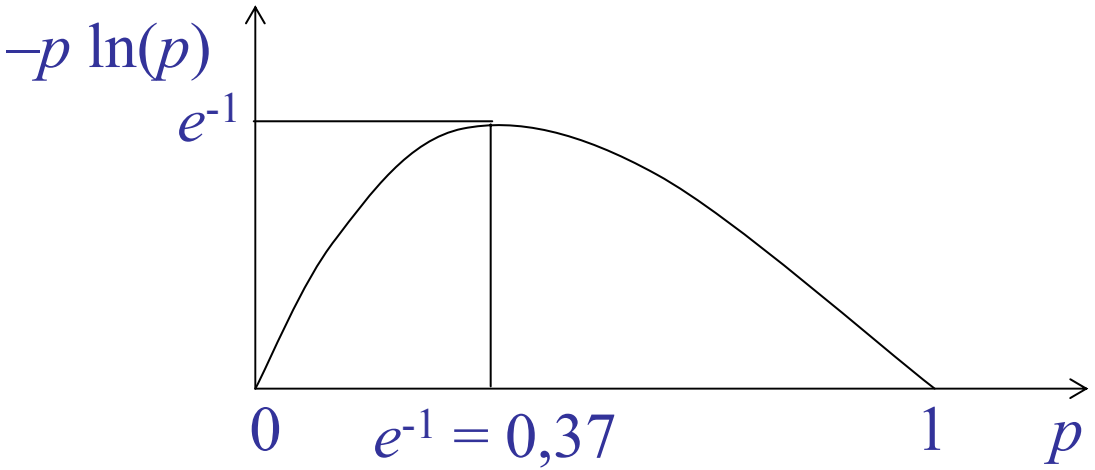
5.2 Các đặc tính của entropy

5.3 Entropy và các dãy của một biến ngẫu nhiên

Entropy của một biến ngẫu nhiên rời rạc

- Định nghĩa
  - Cho x là một biến ngẫu nhiên với không gian mẫu $X = \{x_1, \ldots, x_N\}$ và độ đo xác suất $P(x_n) = p_n$. Entropy của x được định nghĩa là:

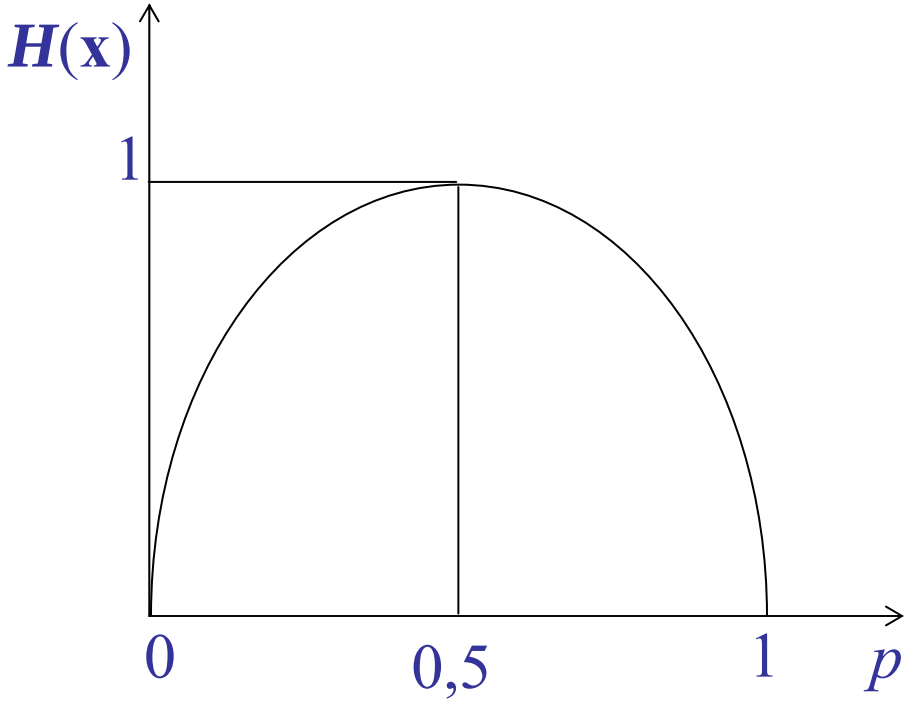
$$H(\mathrm{x}) = -\sum_{n=1}^{N} p_n \log(p_n)$$



Entropy của một biến ngẫu nhiên rời rạc (tt)

- Ví dụ
  - Cho $X = \{0, 1\}$, $P(0) = p$, còn $P(1) = 1-p$. Thì

$$\boldsymbol{H}(\mathbf{x}) = -p\log(p) - (1-p)\log(1-p)$$



Các đặc tính của entropy

1. Entropy là một đại lượng luôn luôn dương hoặc bằng không.
   - $\boldsymbol{H}(\mathbf{x}) = 0 \Leftrightarrow$ có một xác suất $p_i = 1$, còn tất cả các xác suất còn lại bằng 0. Điều này nói lên rằng độ bất ngờ về một thí nghiệm chỉ có một kết quả duy nhất là bằng 0.
2. $H(\mathrm{x}) \le \log N$ và dấu bằng xảy ra $\Leftrightarrow p_1 = p_2 = \ldots = p_N = 1/N$. Hay nói cách khác entropy đạt cực đại khi xác suất xuất hiện của các kí hiệu bằng nhau.

- Chứng minh

$$H(\mathrm{x}) - \ln(N) = -\sum_{n=1}^{N} p_n \ln(p_n) - \sum_{n=1}^{N} p_n \ln(N) = \sum_{n=1}^{N} p_n \ln\left(\frac{1}{Np_n}\right)$$

$$\le \sum_{n=1}^{N} p_n \left(\frac{1}{Np_n} - 1\right) = \sum_{n=1}^{N} \left(\frac{1}{N}\right) - \sum_{n=1}^{N} p_n = 1 - 1 = 0$$

Các đặc tính của entropy (tt)

3. Cho biến ngẫu nhiên **x** có không gian mẫu $X = \{x_1, ..., x_N\}$ và biến ngẫu nhiên **y** có không gian mẫu $Y = \{y_1, ..., y_M\}$. Thì biến ngẫu nhiên nối **z** = (**x**, **y**) có không gian mẫu $Z = \{(x_1, y_1), ..., (x_1, y_M), (x_2, y_1), ..., (x_2, y_M), ..., (x_N, y_1), ..., (x_N, y_M)\}$ gồm NM phần tử. Nếu **x**, **y** độc lập nhau thì $\boldsymbol{H}(\mathbf{z}) = \boldsymbol{H}(\mathbf{x}) + \boldsymbol{H}(\mathbf{y})$.

- Chứng minh

$$H(z) = -\sum_{n=1}^{N}\sum_{m=1}^{M} P(x_n, y_m)\log P(x_n, y_m) = -\sum_{n=1}^{N}\sum_{m=1}^{M} P(x_n)P(y_m)\left[\log P(x_n) + \log P(y_m)\right]$$

$$= -\sum_{n=1}^{N} P(x_n)\log P(x_n)\sum_{m=1}^{M} P(y_m) - \sum_{m=1}^{M} P(x_m)\log P(x_m)\sum_{n=1}^{N} P(y_n)$$

$$= H(\mathrm{x}) + H(\mathrm{y})$$

Các đặc tính của entropy (tt)

4. Xét một biến ngẫu nhiên **x** có không gian mẫu $X = \{x_1, ..., x_n, x_{n+1}, ..., x_N\}$ và các xác xuất $p(x_i) = p_i$. Chúng ta phân X thành hai không gian con, $Y = \{x_1, ..., x_n\}$ và $Z = \{x_{n+1}, ..., x_N\}$. Các xác suất liên kết với Y và Z được cho bởi $P(Y) = \sum_{i=1}^{n} p_i$ và $P(Z) = \sum_{i=n+1}^{N} p_i$. Hơn nữa, chúng ta định nghĩa các biến ngẫu nhiên **y** và **z** bằng $P(y_i) = P(x_i)/P(Y)$, $i = 1, 2, ..., n$ và $P(z_i) = P(x_i)/P(Z)$, $i = n+1, n+2, ..., N$. $\boldsymbol{H}(\mathbf{x})$ bây giờ có thể được viết thành

$$H(\mathrm{x}) = -\sum_{i=1}^{N} p_i \log p_i = -\sum_{i=1}^{n} p_i \log p_i - \sum_{i=n+1}^{N} p_i \log p_i$$

$$= -P(Y)\sum_{i=1}^{n} P(y_i)\left(\log P(y_i) + \log P(Y)\right) - P(Z)\sum_{i=n+1}^{N} P(z_i)\left(\log P(z_i) + \log P(Z)\right)$$

$$= -[P(Y)\log P(Y) + P(Z)\log P(Z)] + [P(Y)H(y) + P(Z)H(z)]$$

Các đặc tính của entropy (tt)

- Trong biểu thức cuối cặp ngoặc vuông đầu biểu diễn độ bất ngờ liên kết với thí nghiệm thứ nhất (là chọn một trong hai không gian mẫu *Y* và *Z*) còn cặp ngoặc vuông thứ hai biểu diễn độ bất ngờ trung bình liên kết với thí nghiệm thứ hai (sau khi đã chọn một trong hai không gian mẫu, sẽ chọn tiếp sự kiện cơ bản nào). Công thức này diễn tả một tính chất của entropy đó là tính chất nhóm.

- Người ta đã chứng minh được rằng công thức định nghĩa của $H(\mathbf{x})$ là công thức duy nhất phù hợp để đo về độ bất ngờ, cái mà phải thoã mãn các tính chất 2,3, 4 và cộng thêm tính liên tục.
- Mặc dầu hai khái niệm lượng tin trung bình và entropy xuất hiện một cách độc lập và ở trong những lĩnh vực khác nhau (entropy vốn xuất phát từ việc nghiên cứu các quá trình nhiệt động) nhưng chúng có cùng công thức giống nhau. Vì vậy chúng ta có thể xem lượng tin trung bình của một nguồn chính là entropy của nguồn đó.

Entropy và các dãy của một biến ngẫu nhiên

- Ví dụ
  - Xét một biến ngẫu nhiên **x** có không gian mẫu $X = \{x_1, x_2\}$, $P(x_1) = p_1 = 1/3$, $P(x_2) = 2/3$. Thì entropy của **x** là

    $$\boldsymbol{H}(\mathbf{x}) = -(1/3)\log(1/3) - (2/3)\log(2/3) = 0.918295834 \text{ bits}$$

  - Chúng ta hãy lặp lại thí nghiệm này N lần để nhận một dãy N phần tử. Tổng quát có đến 2^N dãy có thể. Nếu trong dãy có n phần tử x_1 thì xác suất xuất hiện của dãy là $p_1^n(1-p_1)^{N-n}$
  - Có $\binom{N}{n} = \frac{N!}{n!(N-n)!}$ dãy như vậy, nên tổng xác suất của chúng bằng

    $$\binom{N}{n} p_1^{\,n} (1-p_1)^{N-n}$$

  - Bảng bên dưới trình bày xác suất của các dãy khác nhau đối với $N = 15$

Entropy và các dãy của một biến ngẫu nhiên (tt)

n	Số dãy $\binom{N}{n}$	P mỗi dãy $p_1^{\,n}(1-p_1)^{N-n}$	P tổng cộng $\binom{N}{n}p_1^{\,n}(1-p_1)^{N-n}$	n	Số dãy $\binom{N}{n}$	P mỗi dãy $p_1^{\,n}(1-p_1)^{N-n}$	P tổng cộng $\binom{N}{n}p_1^{\,n}(1-p_1)^{N-n}$
0	1	$2^{-15x0.584962501}$	0.002284	8	6435	$2^{-15x1.118295834}$	0.057404
1	15	$2^{-15x0.651629167}$	0.017127	9	5005	$2^{-15x1.184962501}$	0.022324
2	105	$2^{-15x0.718295834}$	0.059946	10	3003	$2^{-15x1.251629167}$	0.006697
3	455	$2^{-15x0.784962501}$	0.129883	11	1365	$2^{-15x1.318295834}$	0.001522
4	1365	$2^{-15x0.851629167}$	0.194825	12	455	$2^{-15x1.384962501}$	0.000254
5	3003	$2^{-15x0.918295834}$	0.214307	13	105	$2^{-15x1.451629167}$	0.000029
6	5005	$2^{-15x0.984962501}$	0.178589	14	15	$2^{-15x1.518295834}$	0.000002
7	6435	$2^{-15x1.051629167}$	0.114807	15	1	$2^{-15x1.584962501}$	0.000000

Nhận xét

- Những dãy có xác suất lớn (dãy có khả năng) là những dãy mà có n gần với giá trị $Np_1 = 5$, cụ thể là $2 \le n \le 8$. Nói cách khác,
- Xác suất xuất hiện của một dãy mà có n nằm xa giá trị Np_1 là rất nhỏ.
- Xsuất riêng của những dãy có khả năng nằm giữa $2^{-15\times 0.718295834}$ và $2^{-15\times 1.118295834}$, cái mà gần sát với $2^{-N\boldsymbol{H}(\mathbf{x})} = 2^{-15\times 0.918295834}$. Nói cách khác,
- Tất cả những dãy có khả năng là nhiều hay ít đẳng xác suất với xác suất $2^{-NH(\mathrm{x})}$.
- Số lượng tổng cộng các dãy khả năng ($2 \le n \le 8$) là $22803 = 2^{15\times 0.965129067}$ cái mà không xa so với $2^{N\boldsymbol{H}(\mathbf{x})}$. Nói cách khác,
- Số lượng các dãy có khả năng là khoảng $2^{NH(\mathrm{x})}$.

Định lý

- Định lý 5.1
  - Cho các số $\varepsilon > 0$ và $\delta > 0$ nhỏ tuỳ ý, $\exists$ một số nguyên dương N_0 sao cho một dãy có chiều dài bất kỳ $N \geq N_0$ sẽ rơi vào một trong hai lớp sau đây:

    (1) Một tập các dãy mà có tổng xác suất của chúng nhỏ hơn hoặc bằng ε.

    (2) Tập còn lại bao gồm các dãy có xác suất thoã mãn bất đẳng thức

$$2^{-NH-A\sqrt{N}} < p < 2^{-NH+A\sqrt{N}}$$

với A là một số dương nào đó. Hay nói cách khác,

$$\left| \frac{\log p^{-1}}{N} - H \right| \leq \delta$$

Chứng minh định lý

- Chứng minh cho nguồn rời rạc không nhớ $A = \{a_1, a_2, ..., a_K\}$. Gọi **x** là biến ngẫu nhiên gắn với nguồn A. Ta có

$$H(\mathrm{x}) = -\sum_{k=1}^{K} p(a_k) \log p(a_k)$$

- Gọi y là biến ngẫu nhiên bằng cách ánh xạ mỗi a_i tới $\log p(a_i)$.

$$\overline{y} = -\sum_{i=1}^{K} p(a_i) \log p(a_i) = H(\mathrm{x})$$

- Xét các dãy có chiều dài N. Có tất cả K^N dãy như vậy. Ta kí hiệu các dãy này bằng các S_i và xác suất của dãy là $P(S_i)$. Ta có

$$P(S_i) = \prod_{j=1}^{N} p(a^{(j)})$$

trong đó $a^{(j)}$ là kí hiệu thứ j của dãy.

Chứng minh định lý

- Gọi z là biến ngẫu nhiên bằng cách ánh xạ mỗi S_i tới -log $P(S_i)$.
- Chú ý

$$-\log P(S_i) = -\sum_{j=1}^{N} \log p(a^{(j)})$$

- Vì vậy **z** là tổng của N biến ngẫu nhiên **y** độc lập.
- Áp dụng luật yếu của số lớn cho hai số $\varepsilon > 0$ và $\delta > 0$ nhỏ tuỳ ý, tồn tại N_0 sao cho với mọi $N \geq N_0$

$$P\left(\left|\left[\frac{1}{N}\sum_{j=1}^{N} y^{(j)}\right] - \overline{y}\right| \geq \delta\right) \leq \varepsilon$$

hay

$$P\left(\left|-\left[\frac{1}{N}\sum_{j=1}^{N} \log p(a^{(j)})\right] - H(\mathrm{x})\right| \geq \delta\right) \leq \varepsilon$$

Chứng minh định lý (tt)

- Hay $$P\left(\left|-\frac{1}{N}\log P(S_i)-H(x)\right|\geq \delta\right)\leq \varepsilon$$
- Vì vậy chúng ta có thể kết luận rằng với xác suất lớn hơn 1– ε
$$\left|-\frac{1}{N}\log P(S_i)-H(\mathrm{x})\right|\leq \delta$$
đối với mọi $N \geq N_0$.
- Từ đây ta suy ra rằng các dãy được chia thành hai nhóm, một nhóm có tổng xác xuất nhỏ hơn hoặc bằng ε và nhóm thứ hai bao gồm các dãy thoã điều kiện .
$$\left|\frac{1}{N}\log\frac{1}{P(S_i)}-H(\mathrm{x})\right|\leq \delta$$
- Vì vậy định lý được chứng minh.

Bài 6 Mã hiệu

6.1 Giới thiệu

6.2 Mã hiệu và các thông số cơ bản của mã hiệu

6.3 Một số phương pháp biểu diễn mã

6.4 Điều kiện phân tách mã

Giới thiệu

- Trong các hệ thống truyền tin, bên nhận thường biết tập hợp các tin mà bên phát dùng để lập nên các bản tin.
- Các tin thường sẽ được ánh xạ (mã hóa) thành một dạng biểu diễn khác thuận tiện hơn để phát đi.

- Ví dụ
  - Xét một nguồn tin $A = \{a, b, c, d\}$. Chúng ta có thể thiết lập một song ánh như sau từ A vào tập các chuỗi trên bảng chữ cái $\{0, 1\}$

$$a \to 00 \qquad c \to 10$$
$$b \to 01 \qquad d \to 11$$

  - Vậy để phát đi bản tin *baba* chúng ta phát đi chuỗi 01000100. Khi bên nhận nhận được chuỗi này thì xác định được bản tin bên phát đã phát đi là *baba*.

Mã hiệu và những thông số cơ bản

- Mã hiệu (Code), cơ số mã
  - Mã hiệu là một tập hữu hạn các kí hiệu và phép ánh xạ các tin/bản tin của nguồn tin thành các dãy kí hiệu tương ứng. Tập các kí hiệu và phép ánh xạ này thường sẽ phải đáp ứng các yêu cầu tùy theo hệ thống truyền tin đặt ra.
  - Tập các kí hiệu mã dùng để biểu diễn được gọi là bảng kí hiệu mã, còn số các kí hiệu thì được gọi là cơ số mã, và thường kí hiệu là *m*. Nếu mã có cơ số hai thì gọi là mã nhị phân, còn nếu mã có cơ số ba thì gọi là mã tam phân ...
- Mã hoá (Encoding), giải mã (decoding)
  - Mã hoá là quá trình dùng các kí hiệu mã để biểu diễn các tin của nguồn.

Mã hiệu và những thông số cơ bản (tt)

- Nói cách khác mã hoá là một phép biến đổi từ nguồn tin thành mã hiệu, hay mã hoá là phép biến đổi từ một tập tin này thành một tập tin khác có đặc tính thống kê yêu cầu.
- Quá trình ngược lại của quá trình mã hoá được gọi là giải mã.

- Từ mã (Code word), bộ mã
  - Từ mã là chuỗi kí hiệu mã biểu diễn cho tin của nguồn. Tập tất cả các từ mã tương ứng với các tin của nguồn được gọi là bộ mã.
  - Vì vậy có thể nói mã hoá là một phép biến đổi một–một giữa một tin của nguồn và một từ mã của bộ mã.
  - Trong một số trường hợp người ta không mã hoá mỗi tin của nguồn mà mã hoá một bản tin hay khối tin. Lúc này chúng ta có khái niệm mã khối.

Mã hiệu và những thông số cơ bản (tt)

- Các từ mã thường được kí hiệu là u, v, w.

- Chiều dài từ mã, chiều dài trung bình
  - Chiều dài từ mã là số kí hiệu có trong từ mã thường được kí hiệu là l. Chiều dài trung bình của bộ mã thường được kí hiệu là $\bar{l}$ và được cho bằng công thức

$$\bar{l} = \sum_{i=1}^{n} p(x_i) l_i$$

trong đó n là số tin của nguồn còn l_i là chiều dài từ mã tương ứng với tin x_i của nguồn.

- Phân loại mã: mã đều, mã đầy, mã vơi
  - Một bộ mã được gọi là mã đều nếu các từ mã của bộ mã có chiều dài bằng nhau.

Mã hiệu và những thông số cơ bản (tt)

- Một bộ mã đều có cơ số mã là m, chiều dài từ mã là l và số lượng từ mã n bằng với m^l thì được gọi là mã đầy, ngược lại thì được gọi là mã vơi.
- Ngoài ra khái niệm mã đầy còn được dùng theo nghĩa rộng hơn như sau: một bộ mã được gọi là đầy theo một tính chất nào đó (chẳng hạn tính đều hay tính prefix như sau này các bạn sẽ thấy) nếu không thể thêm một từ mã nào vào mà vẫn giữ được tính chất đó.

- Ví dụ
  - Cho bảng kí hiệu mã $A = \{0, 1\}$. Thì bộ mã $X_1 = \{0, 10, 11\}$ là mã không đều, bộ mã $X_2 = \{00, 10, 11\}$ là mã đều nhưng vơi còn bộ mã $X_3 = \{00, 01, 10, 11\}$ là mã đều và đầy.

Một số phương pháp biểu diễn mã

- Bảng đối chiếu mã
  - Là cách liệt kê các tin của nguồn và từ mã tương ứng trong một bảng.

Tin	a_1	a_2	a_3	a_4	a_5	a_6
Từ mã	00	010	011	10	110	111

- Mặt toạ độ mã
  - Là cách biểu diễn mỗi từ mã $w = a_0a_1\ldots a_{l-1}$ bằng một điểm (l, b) trong mặt phẳng toạ độ hai chiều, trong đó l là chiều dài từ mã còn b là trọng số của từ mã được tính như sau với m là cơ số mã

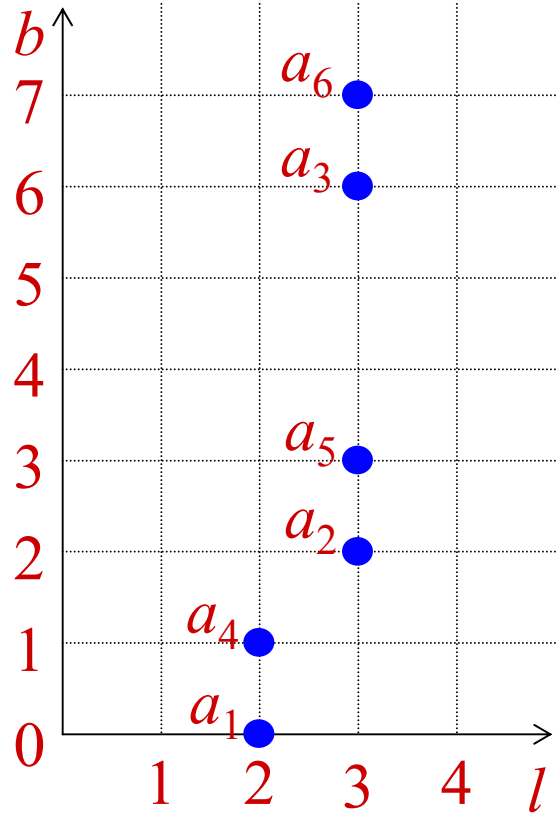
$$b = \sum_{i=0}^{l-1} a_i m^i$$

Một số phương pháp biểu diễn mã (tt)

- Ví dụ

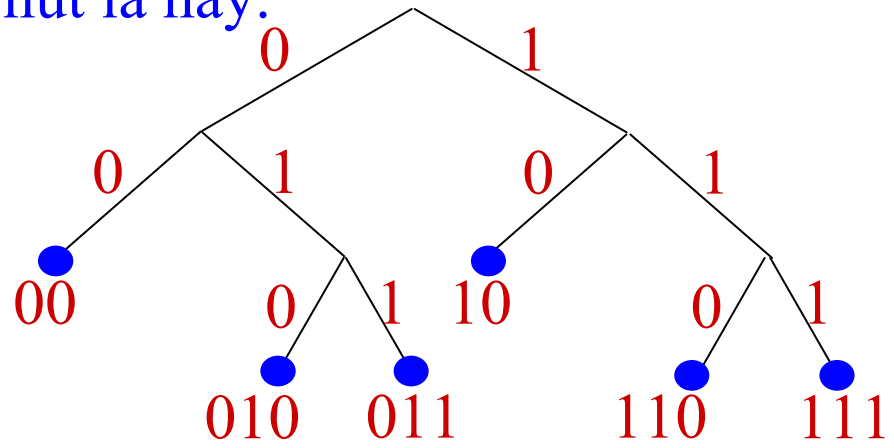
Tin	a_1	a_2	a_3	a_4	a_5	a_6
Từ mã	00	010	011	10	110	111

Tin	a_1	a_2	a_3	a_4	a_5	a_6
Từ mã	00	010	011	10	110	111
Chiều dài l	2	3	3	2	3	3
Trọng số b	0	2	6	1	3	7



Một số phương pháp biểu diễn mã (tt)

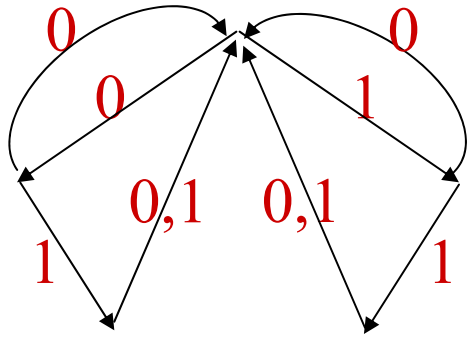
- Cây mã
  - Là cách biểu diễn các từ mã bằng các nút lá của một cây. Mỗi nút lá biểu diễn cho từ mã trùng với nhãn của con đường đi từ nút gốc đến nút lá này.



  - Mã có cơ số *m* thì cây mã tương ứng sẽ là cây *m* phân.
  - Phương pháp cây mã chỉ cho phép biểu diễn những mã prefix, tức là không có từ mã nào trùng với phần đi đầu của một từ mã khác.

Một số phương pháp biểu diễn mã (tt)

- Đồ hình kết cấu mã
  - Là một dạng đặc biệt của cây mã, trong đó các nút lá trùng với nút gốc và ngoài ra mỗi cạnh của đồ hình kết cấu mã đều là cạnh có hướng. Vì vậy một từ mã được biểu diễn bằng một chu trình xuất phát từ nút gốc và quay trở về lại nút gốc.



- Hàm cấu trúc mã
  - Là cách biểu diễn sự phân bố các từ mã theo độ dài của chúng. Phương pháp này biểu diễn bằng một hàm $G(l_i)$ cho biết có bao nhiêu từ mã có chiều dài l_i.

Một số phương pháp biểu diễn mã (tt)

- Ví dụ
  - Bộ mã trong các ví dụ trên được biểu diễn bằng hàm cấu trúc mã sau đây

$$G(l_i) = \begin{cases} 2, \text{ khi } l_i = 2 \\ 4, \text{ khi } l_i = 3 \end{cases}$$

Điều kiện phân tách mã

- Ví dụ
  - Xét bộ mã $X_1 = \{0, 10, 11\}$ mã hoá cho nguồn $A = \{a, b, c\}$. Giả sử bên phát phát đi bảng tin $x = abaac$, lúc đó chuỗi từ mã tương ứng được phát đi là $y = 0100011$.
  - Vấn đề là bên nhận sau khi nhận được chuỗi từ mã y làm sao có thể nhận biết được bảng tin tương ứng mà bên phát đã phát.
  - Để làm được điều này, bên nhận phải thực hiện một quá trình được gọi là tách mã. Chẳng hạn với chuỗi kí hiệu mã nhận được như trên thì bên nhận chỉ có một khả năng để tách mã hợp lý là 0 | 10 | 0 | 0 | 11 và xác định được bảng tin đã được gởi đi là *abaac*.

Điều kiện phân tách mã (tt)

- Xét một bộ mã khác $X_2 = \{0, 10, 01\}$ mã hoá cho nguồn A trên. Giả sử bên nhận nhận được chuỗi kí hiệu là $y = 01010$ và thực hiện quá trình tách mã. Ở đây ta thấy bên nhận có thể thực hiện được ba khả năng tách mã hợp lý sau 0 | 10 | 10, 01 | 0 | 10 và 01 | 01 | 0. Và vì vậy bên nhận sẽ không biết được chính xác bên phát đã phát đi bảng tin nào trong ba bảng tin sau *abb* hay *cab* hay *cca*.
- Một mã như vậy thì không phù hợp cho việc tách mã và được gọi là mã không phân tách được (uniquely undecodable code).
- Vì vậy điều kiện để một bộ mã là phân tách được (uniquely decodable code) là không tồn tại dãy từ mã này trùng với dãy từ mã khác của cùng bộ mã.

Điều kiện phân tách mã (tt)

- Xét một bộ mã khác X_3 = {010, 0101, 10100} mã hoá cho nguồn *A* trên. Giả sử bên nhận nhận được chuỗi kí hiệu là 01010100101 và thực hiện quá trình tách mã. Ở đây ta thấy chỉ có một cách tách mã duy nhất là 0101 | 010 | 0101 nhưng việc tách mã trở nên khó khăn hơn so với bộ mã X_1.
- Chẳng hạn lúc chúng ta gặp chuỗi 010 chúng ta chưa dám chắc đó là một từ mã vì nó có thể là phần đi đầu của từ mã 0101, điều này phụ thuộc vào kí hiệu đi ngay sau chuỗi 010.
- Nếu kí hiệu đi ngay sau là 0 thì chúng ta khẳng định được 010 là từ mã và 0 là phần đi đầu của một từ mã khác sau đó. Còn nếu kí hiệu đi ngay sau là 1 thì chúng ta không khẳng định được, vì có hai khả năng hoặc 010 là một từ mã và 1 là phàn đi đầu của một từ mã khác sau đó, hoặc 0101 là một từ mã.

Điều kiện phân tách mã (tt)

- Nguyên nhân của điều này là do trong bộ mã có một từ mã này là tiếp đầu ngữ của một từ mã khác.
- Và đó cũng chính là nguyên nhân và bản chất của việc một dãy kí hiệu có thể tách thành hai dãy từ mã khác nhau.
- Thật vậy, nếu không có từ mã nào là tiếp đầu ngữ của từ mã khác (hay mã là prefix) thì với mỗi dãy từ mã chỉ có duy nhất một cách tách thành các từ mã thành phần. Vì vậy như sau này chúng ta sẽ thấy các mã thường được sử dụng là các mã prefix.
- Dựa vào tính tiếp đầu ngữ trên, để nhận biết một bộ mã (dĩ nhiên không phải là mã prefix) có phân tách được hay không người ta thường dùng một công cụ được gọi là bảng thử mã.

Bảng thử mã

- Bản chất của bảng thử mã là phân tích những từ mã dài thành những từ mã ngắn đi đầu.
- Chẳng hạn từ mã dài u_1 có thể được phân tích thành $v_{11}v_{12}...v_{1k}w_{11}$ trong đó v_{11}, .., v_{1k} là các từ mã ngắn còn w_{11} là phần còn lại của u_1.
- Nếu w_{11} cũng là một từ mã thì bộ mã này là không phân tách được vì chuỗi $v_{11}v_{12}...v_{1k}w_{11}$ có ít nhất hai cách phân tách thành các từ mã, đó là u_1 và v_{11}, v_{12}, ..., v_{1k}, w_{11}.
- Còn nếu ngược lại w_{11} không là từ mã thì chúng ta dùng nó để xét tiếp. Trong lần xét tiếp theo chúng ta xét xem mỗi w_{11} này có là tiếp đầu ngữ của các từ mã hay không, nếu đúng với một từ mã nào đó, giả sử là u_2, thì từ mã này sẽ có dạng $w_{11}v_{21}...v_{2l}w_{22}$ trong đó v_{21}, ..., v_{2l} là các từ mã ngắn (l có thể bằng 0) còn w_{22} là tiếp vĩ ngữ còn lại.

Bảng thử mã (tt)

- Tương tự nếu w_{22} cũng là một từ mã thì bộ mã là không phân tách được vì chuỗi $v_{11}v_{12}...v_{1k}w_{11}v_{21}...v_{2l}w_{22}$ có ít nhất hai cách phân tách thành các từ mã, đó là $v_{11}v_{12}...v_{1k}w_{11} \mid v_{21} \mid ... \mid v_{2l} \mid w_{22}$, và $v_{11} \mid v_{12} \mid ... \mid v_{1k} \mid w_{11}v_{21}...v_{2l}w_{22}$.
- Nếu ngược lại w_{22} không là từ mã thì chúng ta dùng nó để xét tiếp theo khuôn mẫu tương tự như trên. Vì vậy chúng ta kết luận rằng
- Nếu trong một lần phân tích nào đó, có một từ mã dài, chẳng hạn u, được phân tích thành dãy $w_{ii}v_{(i+1)1}...v_{(i+1)n}$ trong đó w_{ii} là tiếp vĩ ngữ của một từ mã nào đó trong lần phân tích ngay trước đó, còn $v_{(i+1)1}$, ..., $v_{(i+1)n}$ là các từ mã ngắn thì bộ mã là không phân tách được.

Bảng thử mã (tt)

- Thật vậy, lúc đó sẽ tồn tại một dãy kí hiệu sau

  $v_{11}v_{12}\cdots v_{1k}w_{11}v_{21}\cdots v_{2l}w_{22}\cdots w_{(i-1)(i-1)}v_{i1}\cdots v_{im}w_{ii}v_{(i+1)1}\cdots v_{(i+1)n}$

  cái mà có thể phân tách thành hai dãy từ mã khác nhau.
- Cách 1 là

  $v_{11} \mid v_{12} \mid \cdots \mid v_{1k} \mid w_{11}v_{21}\cdots v_{2l}w_{22} \mid \cdots \mid w_{(i-1)(i-1)}v_{i1}\cdots v_{im}w_{ii} \mid v_{(i+1)1} \mid \cdots \mid v_{(i+1)n}$
- Cách 2 là

  $v_{11}v_{12}\cdots v_{1k}w_{11} \mid v_{21} \mid \cdots \mid v_{2l} \mid w_{22}\cdots w_{(i-1)(i-1)} \mid v_{i1} \mid \cdots \mid v_{im} \mid w_{ii}v_{(i+1)1}\cdots v_{(i+1)n}$

Cách xây dựng bảng thử mã

(1) Đem các từ mã xếp thành một cột, theo thứ tự chiều dài của từ mã từ nhỏ đến lớn, đánh dấu là cột 1.

(2) Trong cột này, đối chiếu các từ mã ngắn với các từ mã dài hơn, nếu từ mã ngắn là tiếp đầu ngữ của từ mã dài thì ghi tiếp vĩ ngữ vào cột tiếp theo và đánh dấu là cột 2.

(3) Tiếp tục, đối chiếu các chuỗi trong cột 1 và cột 2 với nhau, nếu có chuỗi nào trong cột này là tiếp đầu ngữ của chuỗi trong cột kia thì tiếp vĩ ngữ sẽ được ghi vào cột tiếp theo là cột 3.

(4) Tiếp tục theo khuôn mẫu này nếu đang xét cột thứ j thì đối chiếu các chuỗi trong cột này với cột 1. Nếu có chuỗi nào trong cột này là tiếp đầu ngữ của chuỗi trong cột kia thì tiếp vĩ ngữ sẽ được ghi vào cột j + 1. Thực hiện cho đến khi không thể điền thêm được nữa hoặc cột mới thêm vào trùng với một cột trước đó hoặc có một chuỗi trong cột mới trùng với một từ mã.

Bảng thử mã (tt)

- Ví dụ
  - Lập bảng thử mã cho bộ mã như đã nói ở trên A = {00, 01, 011, 1100, 00010}

1	2	3	4	5
00	010	0	0	0
01	1	100	1	1
011			11	11
1100			0010	0010
00010				100
				00
				10

Mã là không phân tách được trên chuỗi 000101100 vì có hai cách phân tách khác nhau

00 | 01 | 011 | 00

00010 | 1100

Bảng thử mã (tt)

- Điều kiện cần và đủ để một bộ mã phân tách được là không có phần tử nào trong các cột từ j ≥ 2 trùng với một phần tử trong cột 1.

- Độ chậm giải mã
  - Độ chậm giải mã, thường kí hiệu là T_{ch}, là số kí hiệu cần phải nhận được đủ để có thể phân tách (nhận dạng) được từ mã.
  - Trong trường hợp không có chuỗi nào trong các cột $j \geq 2$ trùng với từ mã nhưng có hai cột k, l nào đó ($k \neq l$, $k, l \geq 2$) trùng nhau thì mã là phân tách được nhưng có độ chậm giải mã vô hạn.

Bảng thử mã (tt)

- Xét bộ mã {01, 10, 011, 100} có bảng thử mã như sau:

1	2	3	4
01	1	0	1
10	0	00	11
011		1	0
100		11	00

- Bảng thử mã này có các cột 3 và 4 trùng nhau về các chuỗi nên bộ mã có độ chậm giải mã trong trường hợp xấu nhất là vô hạn.
- Chẳng hạn với chuỗi có dạng sau đây thì trong quá trình nhận chưa hết chuỗi chúng ta không thể thực hiện được việc tách mã:

0110101010...

Bài tập

- Hãy lập bảng thử mã cho những bộ mã sau. Cho biết mã có phân tách được không, nếu được thì độ chậm giải mã (trong trường hợp xấu nhất) là bao nhiêu.
- $X_1 = \{00, 01, 100, 1010, 1011\}$
- $X_2 = \{00, 01, 101, 1010\}$
- $X_3 = \{00, 01, 110, 111, 1100\}$
- $X_4 = \{00, 01, 110, 111, 1110\}$
- $X_5 = \{00, 01, 110, 111, 0111\}$
- $X_6 = \{00, 01, 110, 111, 1011, 1101\}$

Bất đẳng thức Kraft

- Định lý 6.1
  - Cho $l_1, l_2, ..., l_K$ là các chiều dài của một bộ mã prefix có bảng kí hiệu mã kích thước m (tức gồm m kí hiệu mã). Thì

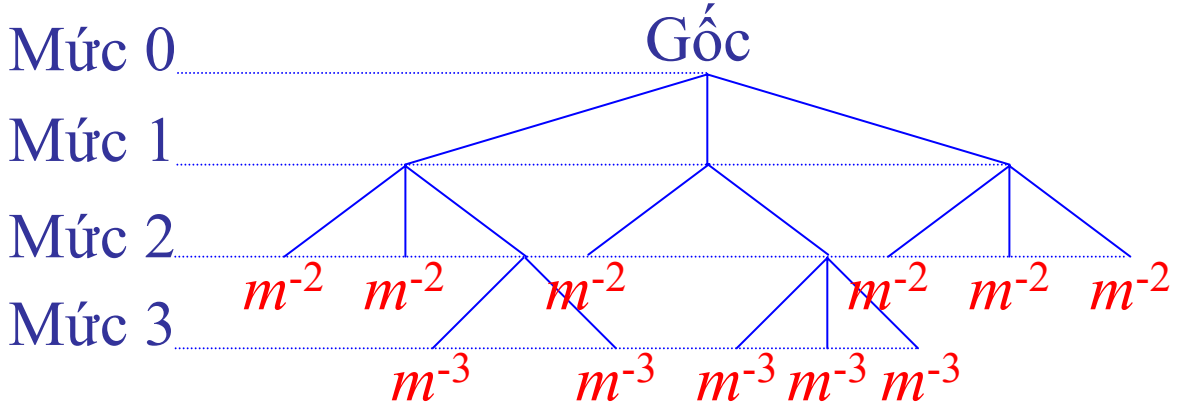
$$\sum_{i=1}^{K} m^{-l_i} \leq 1$$

  - Ngược lại, nếu các số nguyên $l_1, l_2, ..., l_K$ thoã bất đẳng thức trên thì tồn tại một bộ mã prefix với các từ mã có chiều dài là $l_1, l_2, ..., l_K$.
- Chứng minh

Chiều thuận

  - Gọi T là cây mã tương ứng với bộ mã trên

Bất đẳng thức Kraft



- Nút lá ở mức l_i sẽ được gán trọng số là m^{-li}.
- Trọng số của mỗi nút cha được tính bằng tổng trọng số của các nút con.
- Với cách gán này, chúng ta suy ra trọng số của nút cha ở mức h là $\leq m^{-h}$.
- Điều này đúng là vì mỗi nút cha mức h có tối đa m nút con mức $h + 1$.

Bất đẳng thức Kraft (tt)

- Từ đây suy ra, trọng số của nút gốc là ≤ 1.
- Mà trọng số của nút gốc chính là tổng trọng số của các nút lá.
- Vậy suy ra điều cần chứng minh.

Chiều đảo

- Chúng ta chứng minh bằng cách xây dựng một cây mã cho nó.
- Điều này là thực hiện được theo như chứng minh của chiều thuận.

- Ví dụ
  - Tìm bộ mã prefix cho các bộ mã nhị phân có các chiều dài từ mã tương ứng như sau.
  - {2, 2, 3, 4, 4}, {2, 2, 3, 3, 3, 4, 4}, {2, 2, 3, 4, 4, 4, 5, 5}

Định lý

- Định lý 6.2
  - Một mã phân tách được thì có các chiều dài từ mã thoã mãn bất đẳng thức Kraft.

$$\sum_{i=1}^{K} m^{-l_i} \le 1$$

- Chứng minh
  - Gọi $l_1 \le l_2 \le ... \le l_K$ là các chiều dài từ mã với cơ số là m.
  - Với số nguyên N bất kỳ ta có thể viết

$$\left(\sum_{i=1}^{K} m^{-l_i} \right)^N = \sum_{i_1=1}^{K} \cdots \sum_{i_N=1}^{K} m^{-\left(l_{i_1} + \cdots + l_{i_N}\right)}$$

Định lý 6.2 (tt)

- Chú ý $l_{i_1} + \cdots + l_{i_N}$ là chiều dài của một dãy N từ mã và có thể nhận giá trị bất kỳ giữa Nl_1 và Nl_K. Gọi A_j là số dãy N từ mã mà có tổng chiều dài là j. Thì

$$\left(\sum_{i=1}^{K} m^{-l_i}\right)^N = \sum_{j=Nl_1}^{Nl_K} A_j m^{-j}$$

- Vì bộ mã là phân tách được, nên các dãy N từ mã mà có tổng chiều dài là j phải khác nhau.
- Số các dãy có chiều dài j tối đa là m^j. Vì vậy $A_j \le m^j$ và

$$\left(\sum_{i=1}^{K} m^{-l_i}\right)^N \le \sum_{j=Nl_1}^{Nl_K} m^j m^{-j} = N(l_K - l_1) + 1$$

Chứng minh định lý (tt)

- Nếu $\sum_{i=1}^{K} m^{-l_i} > 1$
- Thì với N đủ lớn $\left(\sum_{i=1}^{K} m^{-l_i} \right)^N$ sẽ lớn hơn $N(l_K - l_1) + 1$
- Vì vậy chúng ta có được điều cần chứng minh.

$$\sum_{i=1}^{K} m^{-l_i} \leq 1$$

- Kết hợp hai định lý trên chúng ta rút ra một nhận xét sau.
- Nếu một mã phân tách được thì tồn tại một bộ mã tương đương về chiều dài các từ mã mà có tính prefix.

Bài 7 Mã hóa tối ưu nguồn rời rạc không nhớ

7.1 Các định lý về giới hạn trên và dưới của chiều dài trung bình

7.2 Mã hoá theo Shannon và Fano

7.3 Phương pháp mã hoá tối ưu theo Huffman

Các định lý về giới hạn trên và dưới của chiều dài trung bình

- Định lý 7.1
  - Cho nguồn tin X = $\{a_1, ..., a_K\}$ với các xác suất tương ứng p_1, ..., p_K. Một bộ mã phân tách được bất kỳ cho nguồn này với cơ số mã m, chiều dài trung bình từ mã sẽ thõa (trong đó $H(X)$ là entropy của nguồn với cơ số của logarit là m).

$$\bar{l} \geq \frac{H(X)}{\log m}$$

- Chứng minh

$$H(X) - \bar{l} \ln m = -\sum_{i=1}^{K} p_i \ln p_i - \sum_{i=1}^{K} p_i l_i \ln m = \sum_{i=1}^{K} p_i \ln \frac{m^{-l_i}}{p_i}$$

$$\leq \sum_{i=1}^{K} p_i \left(\frac{m^{-l_i}}{p_i} - 1 \right) = \left(\sum_{i=1}^{K} m^{-l_i} \right) - 1 \leq 1 - 1 = 0$$

Các định lý về giới hạn trên và dưới của chiều dài trung bình (tt)

- Chú ý dấu "=" xảy ra khi và chỉ khi $\frac{m^{-l_i}}{p_i} = 1$, tức là $p_i = m^{-l_i}$

- Định lý 7.2
  - Cho nguồn tin X = $\{a_1, ..., a_K\}$ với các xác suất tương ứng p_1, ..., p_K, có thể xây dựng một mã prefix với cơ số m sao cho

$$\bar{l} < \frac{H(X)}{\log m} + 1$$

- Chứng minh
  - Chọn chiều dài l_i của từ mã cho tin a_i theo qui tắc $l_i = \lceil -\log_m^{p_i} \rceil$
  - Chúng ta có

$$l_i = \lceil -\log_m^{p_i} \rceil \Rightarrow l_i \geq -\log_m^{p_i} \Rightarrow m^{-l_i} \leq p_i$$

$$\Rightarrow \sum_{i=1}^{K} m^{-l_i} \leq \sum_{i=1}^{K} p_i = 1$$

Chứng minh định lý (tt)

- Vì các chiều dài được chọn này thoã bất đẳng thức Kraft nên tồn tại một mã prefix tương ứng có các chiều dài này.
- Tiếp tục chúng ta có

$$l_i = \left\lceil -\log_m^{p_i} \right\rceil \Rightarrow l_i < -\log_m^{p_i} + 1$$

$$\sum_{i=1}^{K} p_i l_i < -\sum_{i=1}^{K} p_i \log_m^{p_i} + \sum_{i=1}^{K} p_i$$

$$= \left(-\sum_{i=1}^{K} \frac{p_i \log p_i}{\log m} \right) + 1 = \frac{H(\mathrm{X})}{\log m} + 1$$

- Điều này hoàn tất chứng minh của chúng ta.

Hệ quả

- Có thể mã hoá một nguồn mà có chiều dài trung bình $\bar{l}$ tiếp cận đến $\frac{H(\mathrm{X})}{\log m}$ với sai số nhỏ tuỳ ý.
- Chúng ta thực hiện điều này bằng cách mã hoá các dãy N tin của nguồn $\mathrm{X} = \{a_1, ..., a_K\}$ theo Định lý 7.2.
- Lúc này chúng ta có nguồn mới với kích thước là K^N, mỗi phần tử là một dãy của N tin được lấy độc lập từ nguồn X.
- Entropy của nguồn mới này là $NH(X)$ và chiều dài trung bình các từ mã của nó theo định nghĩa sẽ là N lần chiều dài trung bình các từ mã của nguồn ban đầu, $\bar{l}$.
- Áp dụng Định lý 7.1 và Định lý 7.2 đối với nguồn mới chúng ta có

Hệ quả (tt)

- Áp dụng Định lý 7.1 và Định lý 7.2 đối với nguồn mới ta có

$$\frac{NH(\mathrm{X})}{\log m} \le N\bar{l} < \frac{NH(\mathrm{X})}{\log m} + 1 \quad \Rightarrow \quad \frac{H(\mathrm{X})}{\log m} \le \bar{l} < \frac{H(\mathrm{X})}{\log m} + \frac{1}{N}$$

- Vì N có thể lớn tuỳ ý, nên $\bar{l}$ tiếp cận đến $H(X) / \log m$ với tốc độ tương đương với $1/N$ tiến đến 0 khi N tiến ra vô cùng.
- Để đánh giá một phương pháp mã hoá nào đó là tốt hay không người ta đưa ra khái niệm hiệu suất lập mã.

Hiệu suất lập mã

- Hiệu suất lập mã h được định nghĩa bằng tỉ số của entropy của nguồn với chiều dài trung bình của bộ mã được lập

$$h = \frac{H(\mathrm{X})}{\bar{l}}$$

Mã hóa tối ưu

- Là phép mã hóa mà kết quả là một bộ mã có chiều dài trung bình là nhỏ nhất trong tất cả các phép mã hóa có thể có cho nguồn.
- Bộ mã của phép mã hóa tối ưu cho nguồn được gọi là bộ mã tối ưu.
- Ba phép mã hóa: Shannon, Fano, Huffman.
- Trong mỗi phép mã hóa chúng ta sẽ mã hóa với cơ số mã $m = 2$ trước (mã hóa nhị phân), sau đó sẽ mở rộng cho trường hợp $m > 2$.

Phương pháp mã hoá Shannon

B1. Sắp xếp các xác suất theo thứ tự giảm dần. Không mất tổng quát giả sử $p_1 \geq ... \geq p_K$.

B2. Định nghĩa $q_1 = 0$, $q_i = \sum_{j=1}^{i-1} p_j$, $\forall\ i = 1, 2, ..., K$.

B3. Đổi q_i sang cơ số 2, (biểu diễn q_i trong cơ số 2) sẽ được một chuỗi nhị phân

B4. Từ mã được gán cho a_i là l_i kí hiệu lấy từ vị trí sau dấu phẩy của chuỗi nhị phân tương ứng với q_i, trong đó $l_i = \left\lceil -\log_2^{p_i} \right\rceil$

Ví dụ

- Hãy mã hoá nguồn $S = \{a_1, a_2, a_3, a_4, a_5, a_6\}$ với các xác suất lần lượt là 0,3; 0,25; 0,2; 0,12; 0,08; 0,05.

Tin a_i	Xác suất p_i	$q_i = \sum_{j=1}^{i-1} q_j$	Biểu diễn nhị phân	$l_i = \lceil -\log_2 p_i \rceil$	Từ mã w_i
a_1	0,3	0	0,00	2	00
a_2	0,25	0,3	0,01001...	2	01
a_3	0,2	0,55	0,10001...	3	100
a_4	0,12	0,75	0,11000...	4	1100
a_5	0,08	0,87	0,11011...	4	1101
a_6	0,05	0,95	0,111100...	5	11110

- H = 2.36, $\bar{l}$ = 2,75, h = 2,36/2,75 = 85,82%

Nhận xét - Bài tập

- Phương pháp Shannon cho kết quả là một mã prefix.
- Phương pháp Shannon có thể mở rộng cho trường hợp m > 2

- Bài tập
  - Hãy mã hoá các nguồn sau bằng phương pháp Shannon. Tính entropy của nguồn, chiều dài trung bình và hiệu suất của phép mã hóa.
  - $S_1 = \{a_1, a_2, a_3, a_4, a_5, a_6\}$ với các xác suất lần lượt là 0,25; 0,21; 0,19; 0,16; 0,14; 0,05.
  - $S_2 = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8\}$ với các xác suất lần lượt là 0,21; 0,18; 0,15; 0,14; 0,12; 0,01; 0,06 ; 0,04.
  - $S_3 = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9\}$ với các xác suất lần lượt là 0,25; 0,19; 0,15; 0,11; 0,09; 0,07; 0,06; 0,04; 0,04.

Phương pháp mã hoá Fano

B1. Sắp xếp các xác suất theo thứ tự giảm dần. Không mất tổng quát giả sử $p_1 \geq ... \geq p_K$.

B2. Phân các xác suất thành 2 nhóm có tổng xác suất gần bằng nhau nhất.

B3. Gán cho hai nhóm lần lượt các kí hiệu 0 và 1 (hoặc ngược lại).

B4. Lặp lại bước 2 cho các nhóm con cho đến khi không thể tiếp tục được nữa.

B5. Từ mã ứng với mỗi tin là chuỗi bao gồm các kí hiệu theo thứ tự lần lượt được gán cho các nhóm có chứa xác suất tương ứng của tin.

Ví dụ

- Hãy mã hoá nguồn $S = \{a_1, a_2, a_3, a_4, a_5, a_6\}$ với các xác suất lần lượt là 0,3; 0,25; 0,2; 0,12; 0,08; 0,05.

Tin	Xác suất	Phân nhóm lần				Từ mã
		1	2	3	4	
a_1	0,3	0	0			00
a_2	0,25	0	1			01
a_3	0,2	1	0			10
a_4	0,12	1	1	0		110
a_5	0,08	1	1	1	0	1110
a_6	0,05	1	1	1	1	1111

- H = 2.36, $\bar{l} = 2{,}38$, $h = 2{,}36/2{,}38 = 99{,}17\%$

Chú ý

- Chú ý, trong nhiều trường hợp có nhiều hơn một cách chia thành các nhóm có tổng xác suất gần bằng nhau, ứng với mỗi cách chia có thể sẽ cho ra các bộ mã có chiều dài trung bình khác nhau.

- Ví dụ
  - Hãy mã hoá nguồn $S = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8\}$ với các xác suất lần lượt là 0,23; 0,2; 0,14; 0,12; 0,1; 0,09; 0,06; 0,06.

Ví dụ

a_i	p_i	1	2	3	4	w_i
a_1	0,23	0	0			00
a_2	0,2	0	1			01
a_3	0,14	1	0	0		100
a_4	0,12	1	0	1		101
a_5	0,1	1	1	0	0	1100
a_6	0,09	1	1	0	1	1101
a_7	0,06	1	1	1	0	1110
a_8	0,06	1	1	1	1	1111

a_i	p_i	1	2	3	4	w_i
a_1	0,23	0	0			00
a_2	0,2	0	1	0		010
a_3	0,14	0	1	1		011
a_4	0,12	1	0	0		100
a_5	0,1	1	0	1		101
a_6	0,09	1	1	0		110
a_7	0,06	1	1	1	0	1110
a_8	0,06	1	1	1	1	1111

- $\bar{l}_1 = 2{,}88,$ $\bar{l}_2 = 2{,}89$

Nhận xét - Bài tập

- Nhận xét
  - Phương pháp Fano thường cho kết quả tốt hơn phương pháp Shannon.
- Bài tập
  - Hãy mã hoá các nguồn sau bằng phương pháp Fano. Tính hiệu suất của phép mã hóa.
  - $S_1 = \{a_1, a_2, a_3, a_4, a_5, a_6\}$ với các xác suất lần lượt là 0,25; 0,21; 0,19; 0,16; 0,14; 0,05.
  - $S_2 = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8\}$ với các xác suất lần lượt là 0,21; 0,18; 0,15; 0,14; 0,12; 0,1; 0,06 ; 0,04.
  - $S_3 = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9\}$ với các xác suất lần lượt là 0,25; 0,19; 0,15; 0,11; 0,09; 0,07; 0,06; 0,04; 0,04.

Phương pháp mã hoá tối ưu Huffman

- Trước hết xét cơ số mã $m = 2$. Trường hợp $m > 2$, chúng ta sẽ có một sự chú ý về sự khác biệt so với trường hợp $m = 2$.

- Bổ đề
  - Cho nguồn $S = \{a_1, ..., a_K\}$ có các xác suất lần lượt là $p_1, ..., p_K$. Gọi $l_1, ..., l_K$ là chiều dài các từ mã tương ứng với bộ mã tối ưu cho S. Nếu $p_i > p_j$ thì $l_i \le l_j$.

- Chứng minh
  - Với $p_i > p_j$, giả sử $l_i > l_j$. Xét bộ mã mới bằng cách hoán đổi hai từ mã có chiều dài l_i và l_j cho nhau. Xét hiệu chiều dài trung bình của bộ mã mới so với bộ mã cũ

$$\Delta\bar{l} = p_i l_j + p_j l_i - p_i l_i - p_j l_j = (p_j - p_i)(l_i - l_j) < 0$$

Điều này mâu thuẫn với định nghĩa của bộ mã tối ưu.

Hai định lý của Huffman

- Bổ đề này thật sự phát biểu một điều rằng, để mã hoá tối ưu cho một nguồn tin thì tin có xác suấ càng lớn phải được mã hoá thành từ mã có chiều dài càng nhỏ.

- Định lý 7.3 (Định lý số 1 của Huffman)
  - Trong bộ mã tối ưu ($m = 2$) cho một nguồn tin, thì hai từ mã tương ứng với hai tin có xác suất nhỏ nhất phải có chiều dài bằng nhau ($l_{K-1} = l_K$) và có thể làm cho chúng chỉ khác nhau duy nhất ở bit cuối (bit tận cùng bên phải).

- Chứng minh
  - Nếu $l_{K-1} < l_K$ thì loại bỏ bit cuối cùng của từ mã w_K chúng ta được một bộ mã mới vẫn có tính prefix nhưng có chiều dài trung bình nhỏ hơn bộ mã cũ.

Hai định lý của Huffman (tt)

- Giả sử w_{K-1} và w_K không thõa điều kiện là khác nhau chỉ ở bit cuối.
- Nếu có một từ mã w_i khác có chiều dài bằng l_K đồng thời khác từ mã w_K chỉ ở bit cuối thì chúng ta có thể hoán đổi w_{K-1} và w_i cho nhau, vì vậy định lý cũng được chứng minh.
- Nếu không tồn tại một từ mã w_i như vậy thì chúng ta có thể tạo ra một bộ mã mới bằng cách bỏ đi bit cuối của từ mã w_K. Bộ mã mới này không vi phạm điều kiện prefix và có chiều dài trung bình nhỏ hơn bộ mã cũ. Vì vậy định lý được chứng minh.

Hai định lý của Huffman (tt)

- Định lý 7.4 (Định lý số 2 của Huffman)
  - Xét một nguồn mới $S' = \{a'_1, ..., a'_{K-1}\}$ với sự phân bố xác suất là $p'_1, ..., p'_{K-1}$ trong đó $p'_i = p_i$ với $1 \le i \le K-2$ còn $p'_{K-1} = p_{K-1} + p_K$. Nếu $\{w'_1, ..., w'_{K-1}\}$ làm một mã tối ưu cho S' thì mã nhận được theo qui tắc sau là mã tối ưu cho S.

$$w_i = w'_i, \qquad 1 \le i \le K-2$$
$$w_{K-1} = w'_{K-1}\mathbf{0}$$
$$w_K = w'_{K-1}\mathbf{1}$$

- Chứng minh
  - Vì $l_K = l_{K-1} = 1 + l'_{K-1}$, nên

$$\bar{l} = p_1 l_1 + ... + p_K l_K = p_1 l'_1 + ... + (p_{K-1} + p_K)(1 + l'_{K-1})$$
$$= \bar{l}' + (p_{K-1} + p_K)$$

Hai định lý của Huffman (tt)

- Sự khác biệt giữa $\bar{l}$ và $\bar{l}'$ là một hằng số.
- Nên nếu mã tối ưu cho nguồn S là tốt hơn mã theo qui tắc đã phát biểu thì mã được dẫn xuất từ mã tối ưu này bằng cách bỏ đi hai từ mã w_K và w_{K-1} và thay vào từ mã mà bỏ đi bit cuối của w_K thì sẽ được một mã tối ưu tốt hơn cho nguồn S', điều này mâu thuẫn.
- Vậy mã nhận được cho S theo qui tắc trên là tối ưu.
- Định lý Định lý 7.3 và 7.4 cho phép qui bài toán tìm mã tối ưu cho nguồn có K tin về bài toán tìm mã tối ưu cho nguồn có K–1 tin. Và quá trình này có thể được lặp lại cho đến khi chỉ còn hai tin. Lúc đó thì mã tối ưu là dễ thấy.

Giải thuật mã hóa Huffman

B1. Sắp xếp các xác suất theo thứ tự giảm dần chẳng hạn $p_1 \geq ... \geq p_K$

B2. Gán **0** tới bit cuối của w_{K-1} và **1** đến bit cuối của w_K hoặc ngược lại. Tuy nhiên chúng ta sẽ qui ước thực hiện theo chiều thứ nhất.

B3. Kết hợp p_K và p_{K-1} để tạo thành một tập xác suất mới $p_1, ..., p_{K-2}, \boldsymbol{p_{K-1} + p_K}$

B4. Lặp lại các bước trên cho tập mới này.

- Ví dụ
  - Hãy mã hoá nguồn $S = \{a_1, a_2, a_3, a_4, a_5, a_6\}$ với các xác suất lần lượt là 0,3; 0,25; 0,2; 0,12; 0,08; 0,05.

Ví dụ

a_i	p_i	Lần 1	Lần 2	Lần 3	Lần 4	w_i
a_1	0,3	0,3	0,3	0,45	0,55 0	00
a_2	0,25	0,25	0,25	0,3 0	0,45 1	01
a_3	0,2	0,2	0,25 0	0,25 1		11
a_4	0,12	0,13 0	0,2 1			101
a_5	0,08 0	0,12 1				1000
a_6	0,05 1					1001

- H = 2.36, $\bar{l}$ = 2,38, h = 2,36/2,38 = 99,17%

Nhận xét

- Nhận xét
  - So sánh với phương pháp Fano ta thấy trong trường hợp trên thì cả hai phương pháp cho hiệu suất bằng nhau.
  - Tuy nhiên trong trường hợp tổng quát phương pháp Fano không phải là phương pháp mã hóa tối ưu.
- Chú ý
  - Trong trường hợp nếu xác suất $p_{K-1} + p_K$ bằng với một xác suất p_i nào đó thì chúng ta có thể đặt $p_{K-1} + p_K$ nằm dưới hoặc nằm trên xác suất p_i thì kết quả chiều dài từ mã trung bình vẫn không thay đổi cho dù các từ mã kết quả có thể khác nhau.

Mở rộng cho cơ số m > 2

- Nếu $K \leq m$ thì việc mã hoá tối ưu là quá tầm thường
- Giả sử $K > m$, tồn tại n sao cho: $m + (n-1)(m-1) < K \leq m + n(m-1)$. Chúng ta sẽ bổ sung vào một số tin "phụ" có xác suất bằng 0 sao cho tổng số tin của nguồn bằng với $m + n(m-1)$. Sau đó thủ tục mã hoá trên được điều chỉnh như sau

B1. Sắp xếp các xác suất theo thứ tự giảm dần chẳng hạn $p_1 \geq ... \geq p_K$

B2. Gán lần lượt các kí hiệu 0, 1, ..., $m-1$ tới các bit cuối của m từ mã có xác suất nhỏ nhất

B3. Kết hợp m xác suất nhỏ nhất lại thành một và tạo với $K-m$ xác suất còn lại thành một tập mới.

B4. Lặp lại các bước trên cho tập mới này.

Ví dụ

- Hãy mã hoá nguồn $S = \{a_1, a_2, a_3, a_4, a_5, a_6\}$ với các xác suất lần lượt là 0,3; 0,25; 0,2; 0,12; 0,08; 0,05 với $m = 3$.

a_i	p_i	Lần 1	Lần 2	w_i
a_1	0,3	0,3	0,45 (0)	1
a_2	0,25	0,25	0,3 (1)	2
a_3	0,2	0,2 (0)	0,25 (2)	00
a_4	0,12	0,13 (1)		02
a_5	0,08 (0)	0,12 (2)		010
a_6	0,05 (1)			011
a_7	0,0 (2)			

- H = 1.49, $\bar{l}$ = 1,58, h = 1,49/1,58 = 94,24%

Bài tập

- Hãy mã hoá các nguồn sau bằng phương pháp Huffman theo các cơ số $m = 2$ và $m = 3$. Tính hiệu suất của phép mã hóa trong mỗi trường hợp.
- $S_1 = \{a_1, a_2, a_3, a_4, a_5, a_6\}$ với các xác suất lần lượt là 0,25; 0,21; 0,19; 0,16; 0,14; 0,05.
- $S_2 = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8\}$ với các xác suất lần lượt là 0,23; 0,2; 0,14; 0,12; 0,1; 0,09; 0,06 ; 0,06.
- $S_3 = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8\}$ với các xác suất lần lượt là 0,21; 0,18; 0,15; 0,14; 0,12; 0,01; 0,06 ; 0,04.
- $S_4 = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9\}$ với các xác suất lần lượt là 0,25; 0,19; 0,15; 0,11; 0,09; 0,07; 0,06; 0,04; 0,04.

Nhận xét

- Xét nguồn $S = \{a_1, a_2, a_3, a_4\}$ có sự phân bố xác suất là {0,4; 0,25; 0,2; 0,15}. Xét nguồn mới $S^2 = \{a_i a_j, 1 \le i, j \le 4\}$ có tập phân bố xác suất là {0,16; 0,1; 0,08; 0,06; 0,1; 0,0625; 0,05; 0,0375; 0,08; 0,05; 0,04; 0,03; 0,06; 0,0375; 0,03; 0,0225}.
- $H(S) = 1,9$ và $H(S^2) = 2H(S) = 3,8$.
- Hai bảng sau đây trình bày kết quả việc mã hoá tối ưu cho S và S^2 theo Huffman.

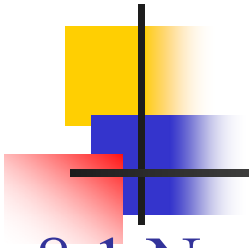
- Nhận xét
  - Việc mã hoá cho một dãy tin (hay khối tin) thì cho hiệu suất cao hơn so với việc mã hoá cho từng tin.

Nhận xét (tt)

Tin	p_i	Từ mã
a_1	0,4	1
a_2	0,25	01
a_3	0,2	000
a_4	0,15	001

- $\overline{l_s} = 1,95$
- $h_1 = 97,63\%$
- $\overline{l_{s^2}} = 3,8375$
- $h_2 = 99,26\%$

Tin	p_{ij}	Từ mã	Tin	p_{ij}	Từ mã
a_1a_1	0,16	000	a_2a_3	0,05	1110
a_1a_2	0,1	101	a_3a_2	0,05	1111
a_2a_1	0,1	110	a_3a_3	0,04	01000
a_1a_3	0,08	0010	a_2a_4	0,0375	01001
a_3a_1	0,08	0011	a_4a_2	0,0375	01010
a_2a_2	0,0625	0110	a_3a_4	0,03	01011
a_1a_4	0,06	0111	a_4a_3	0,03	10010
a_4a_1	0,06	1000	a_4a_4	0,0225	10011



Bài 8 Mã hóa nguồn phổ quát

8.1 Nguồn rời rạc không nhớ với thống kê không biết trước

8.2 Các vectơ tần xuất và tựa–entropy (quasi–entropy)

8.3 Một sơ đồ mã hoá phổ quát cho nguồn rời rạc không nhớ

Giới thiệu

- Vấn đề này không được khởi xướng bởi Shannon mà bởi B. M. Fitingof.
- Lý thuyết của Shannon dựa trên kiến thức về các hàm phân bố xác suất và chứng minh tồn tại phép mã hoá tối ưu.
- Mã hoá nguồn phổ quát tiếp cận theo cách khác bằng việc lợi dụng cấu trúc của các dãy và cũng đi đến được cùng kết quả tối ưu.
- Trong trường hợp mà các hàm phân bố xác suất là không có sẵn hoặc sự thống kê về nguồn là thay đổi theo thời gian, những điều mà thường xảy ra trong thực tế, thì kỹ thuật mã hoá nguồn phổ quát là một lựa chọn thích hợp hơn là dùng kỹ thuật của Shannon.

Nguồn rời rạc không nhớ với thống kê không biết trước

- Xét nguồn $A = \{a_1, ..., a_K\}$ có sự phân bố xác suất là $\{p_1, ..., p_K\}$ sinh ra một dãy các kí hiệu độc lập có phân bố đồng nhất.
- Chúng ta giả thiết rằng sự phân bố xác suất $\{p_1, ..., p_K\}$ là cố định nhưng không được biết trước bởi bộ mã hoá (encoder).
- Thực tế sự phân bố xác suất thường là không được biết trước hoặc chỉ được biết ở mức độ gần đúng, hoặc sự phân bố này thay đổi chậm theo thời gian.
- Vì vậy một sơ đồ mã hoá dựa trên xác suất có thể hiệu quả ở khung thời gian này nhưng sẽ không hiệu quả ở khung thời gian khác.

Nguồn rời rạc không nhớ với thống kê không biết trước (tt)

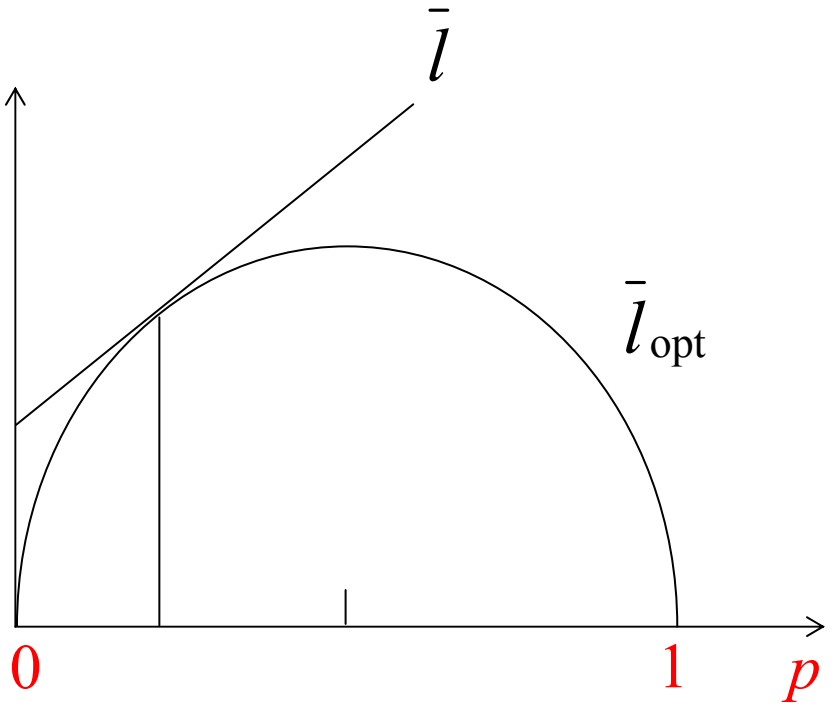
- Đánh giá ảnh hưởng của sự biết không chính xác về thống kê của nguồn đến hiệu quả của việc mã hoá
- Xét nguồn rời rạc không nhớ nhị phân với sự phân bố xác suất là P(0) = p, P(1) = 1– p.
- Nếu bộ mã hoá được cung cấp xác suất gần đúng với p là p_0 thì theo phương pháp của Shannon kí hiệu 0 sẽ được gán với từ mã có chiều dài là $-\log p_0$ còn 1 được gán với từ mã có chiều dài $-\log (1- p_0)$.
- Chiều dài trung bình của các từ mã là

$$\bar{l} = -p \log p_0 - (1-p) \log(1-p_0)$$

- Chiều dài trung bình từ mã tối ưu là

$$\bar{l}_{\text{opt}} = -p \log p - (1-p) \log(1-p)$$

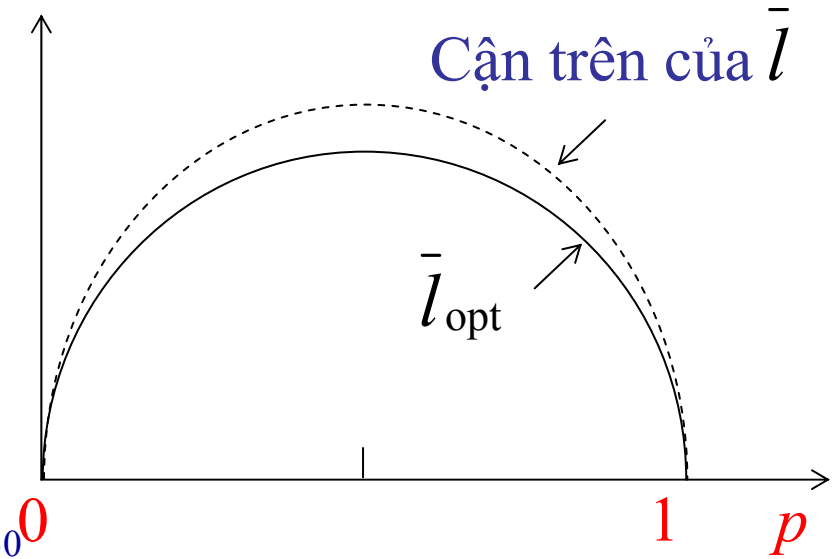
Nguồn rời rạc không nhớ với thống kê không biết trước (tt)



- Chú ý rằng $\bar{l}$ là một tiếp tuyến của $\bar{l}_{\text{opt}}$ tại $p = p_0$, nhưng khi p lệch ra xa p_0 thì khoảng cách giữa hai đồ thị gia tăng khá nhanh.

Nguồn rời rạc không nhớ với thống kê không biết trước (tt)

- Trong bài này chúng ta phát triển các ý tưởng cơ bản về mã hoá phổ quát, một sơ đồ mã hoá không dựa trên xác xuất của các dãy mà lại dựa vào cấu trúc của chúng.
- Chúng ta sẽ chứng minh rằng $\forall$ ε nguyên dương nhỏ tuỳ ý có khả năng mã hoá một nguồn sao cho $\bar{l} \le \boldsymbol{H}(\mathbf{x}) + \varepsilon$ đối $\forall$ sự phân bố xác suất $\{p_1, ..., p_K\}$ của nguồn.
- ε có thể được làm nhỏ tuỳ ý bằng cách chọn chiều dài khối tin cần mã hoá đủ lớn.



Các vectơ tần suất và tựa-entropy

- Xét các dãy nguồn S_i có chiều dài N.
- Có K^N dãy và ta gọi tập K^N dãy này là không gian mẫu S.
- Chúng ta kí hiệu N_{ki} là số kí hiệu a_k có trong dãy S_i và q_{ki} là tần suất của a_k trong S_i

$$q_{ki} = N_{ki} / N$$

- Vectơ $(q_{1i}, ..., q_{Ki})$ (kí hiệu là $Q(S_i)$ hay gọn hơn là Q_i) được gọi là vectơ tần suất ứng với chuỗi Si.
- Gọi các q_k ($k = 1, ..., K$) là các biến ngẫu nhiên trên S bằng cách gán mỗi S_i với q_{ki}. Chúng ta có bổ đề sau.
- Giá trị trung bình của q_k chính là xác suất p_k của a_k.

$$E(q_k) = \sum_{i=1}^{K^N} P(S_i) q_{ki} = p_k$$

Các vectơ tần suất và tựa-entropy (tt)

- Chứng minh
  - Định nghĩa biến ngẫu nhiên $x_k^{(n)}$ bằng $1/N$ nếu nguồn sinh ra kí hiệu a_k tại vị trí thứ n của dãy và bằng 0 nếu ngược lại.

    Vì nguồn là không nhớ, dãy $x_k^{(1)}$, ..., $x_k^{(N)}$ là độc lập và có phân bố đồng nhất.

    Giá trị trung bình của $x_k^{(n)}$ bằng $p_k/N \; \forall \; n$. Mà $\mathbf{q}_k = \sum_{n=1}^{N} \mathbf{x}_k^{(n)}$

    Vì vậy
    $$E(\mathrm{q}_k) = \sum_{n=1}^{N} E(\mathrm{x}_k^{(n)}) = p_k$$
  - Mỗi dãy S_i có tương ứng một vectơ tần suất Q_i, nhưng ngược lại với một vectơ $Q = (q_1, ..., q_K)$ có thể tương ứng với nhiều dãy S_i.

Các vectơ tần suất và tựa-entropy (tt)

- Gọi $\omega(Q)$ là số các dãy S_i mà có cùng vectơ tần suất Q (tức là những dãy mà có số lần xuất hiện của mỗi a_k trong dãy bằng nhau và bằng $N_k = Nq_k \ \forall \ k = 1, ..., K$).

$$\omega(Q) = \frac{N!}{\prod_{k=1}^{K} N_k!}$$

- Gọi $\phi(K, N)$ là số vectơ biểu diễn cho các dãy nguồn có chiều dài N.
- Con số này có thể diễn đạt thành một bài toán tập hợp tương đương khá quen thuộc là: Có bao nhiêu bộ gồm K số nguyên không âm mà có tổng bằng N.

- Bổ đề

$$\Phi(K, N) = \binom{N + K - 1}{N}$$

Các vectơ tần suất và tựa-entropy (tt)

- Chứng minh
  - Xét một hàng gồm $N + K - 1$ khoảng trống. Dùng N đối tượng giống nhau lấp vào N khoảng trống bất kỳ. $K - 1$ khoảng trống còn lại sẽ chia N đối tượng này thành K nhóm. Do đó ứng với mỗi cách lấp N đối tượng vào $N + K - 1$ vị trí chúng ta có một tổng tương ứng. Vì vậy số lượng tổng này bằng

$$\binom{N+K-1}{N}$$

  - Với mỗi dãy S_i chúng ta có tương ứng một vectơ $Q_i = (q_{1i}, \ldots, q_{Ki})$. Chúng ta định nghĩa một biến ngẫu nhiên $\psi(Q)$ gán mỗi dãy S_i với giá trị (kí hiệu là $\psi(S_i)$ hoặc $\psi(Q_i)$)

$$-\sum_{k=1}^{K} q_{ki} \log q_{ki}$$

Các vectơ tần suất và tựa-entropy (tt)

- Chú ý Q_i là một vectơ xác suất và $\psi(Q_i)$ có công thức giống như của entropy nên chúng ta gọi $\psi(Q_i)$ là tựa–entropy.
- Dĩ nhiên $\psi(Q_i)$ có tất cả các tính chất của hàm entropy $H(Q_i)$ cái mà chỉ phụ thuộc duy nhất vào Q_i.
- Chúng ta có định lý sau thiết lập mối quan hệ giữa $\psi(\mathrm{Q})$ (hay viết rõ ra là $\psi(q_1, ..., q_K)$) với entropy của nguồn $H(p_1, ..., p_K)$.

- Định lý 8.1

$$E(\psi(\mathbf{Q})) \le H(p_1, ..., p_K)$$

- Chứng minh

$$E(\psi(Q)) = E\left(-\sum_{k=1}^{K} q_k \log q_k\right) = \sum_{k=1}^{K} E\left(-q_k \log q_k\right)$$

Các vectơ tần suất và tựa-entropy (tt)

- Mà để ý hàm $-x \log x$ là hàm lồi, vì vậy theo bất đẳng thức Jensen chúng ta có

$$E(-q_k \log q_k) \le E(-q_k) \log E(q_k)$$

Theo một bổ đề trước đây chúng ta có $E(q_k) = p_k$. Vì vậy

$$E(\psi(Q)) = \sum_{k=1}^{K} - p_k \log p_k = H(p_1, \ldots, p_K)$$

Một sơ đồ mã hoá phổ quát cho nguồn rời rạc không nhớ

- Một từ mã cho một dãy S_i gồm hai phần: phần đầu là chuỗi mã hoá cho vectơ tuần suất Q_i tương ứng của dãy S_i, phần thứ hai là chuỗi mã hoá cho dãy S_i trong số các dãy có cùng vectơ Q_i.
- Vì tổng các vectơ tần suất khác nhau là $\phi(K, N)$, nên số bit dùng để biểu diễn cho phần đầu là

$$\lceil \log \phi(K, N) \rceil$$

- Tương tự số bit để biểu diễn cho phần thứ hai là

$$\lceil \log \varpi(Q_i) \rceil$$

- Vì vậy từ mã biểu diễn cho dãy S_i có chiều dài là

$$l(S_i) = \lceil \log \phi(K, N) \rceil + \lceil \log \varpi(Q_i) \rceil$$
$$< \log \phi(K, N) + \log \omega(Q_i) + 2.$$

Một sơ đồ mã hoá phổ quát cho nguồn rời rạc không nhớ (tt)

- Chúng ta chứng minh được giá trị trung bình của $l(S_i)$ thoã

$$E(l(S_i)) < NH(p_1,\ldots,p_K) + N\log(1+\frac{K-1}{N}) + (K-1)\log(1+\frac{N}{K-1})$$

- Suy ra chiều dài trung bình trên một kí tự nguồn thoã

$$\bar{l} = \frac{E(l(S_i))}{N} < H(p_1,\ldots,p_K) + \left[\log(1+\frac{K-1}{N}) + \frac{K-1}{N}\log(1+\frac{N}{K-1})\right]$$

- Chú ý thành phần nằm trong dấu móc vuông tiến đến 0 khi $N \to \infty$ với tốc độ bằng với tốc độ của $\frac{\log N}{N} \to 0$

- Điều này nói lên rằng phương pháp này tiếp cận đến entropy của nguồn chậm hơn so với các phương pháp mà biết trước xác suất. Điều này cũng dễ hiểu và cũng là cái giá phải trả nếu chúng ta không biết trước xác suất.

Ví dụ

- Bảng sau đây mô tả việc mã hoá phổ quát cho một nguồn nhị phân cho từng khối có chiều dài 7.
- Có $\phi(2, 7) = 8$ vectơ tần suất và vì vậy cần dùng 3 bit để mã hoá 8 vectơ này; 3 bit này sẽ là 3 bit đầu của mọi từ mã. Các bit còn lại dùng để nhận biết mỗi dãy trong lớp đã cho (là lớp các dãy có cùng vectơ tần suất).

Q_i	$\omega(Q_i)$	S_i	$\psi(S_i)$	w_i
(0/7,7/7)	1	1111111	0	000
(1/7,6/7)	7	0111111 1011111 ... 1111110	0,592	001 000 001 001 001 110

Ví dụ (tt)

Q_i	$\omega(Q_i)$	S_i	$\psi(S_i)$	w_i
(2/7,5/7)	21	0011111 0101111 ... 1111100	0,863	010 00000 010 00001 ... 010 10100
(3/7,4/7)	35	0001111 0010111 ... 1111000	0,985	011 000000 011 000001 011 100010
(4/7,3/7)	35	0000111 0001011 ... 1110000	0,985	100 000000 100 000001 100 100010

Ví dụ (tt)

Q_i	$\omega(Q_i)$	S_i	$\psi(S_i)$	w_i
(5/7,2/7)	21	0000011 0000101 ... 1100000	0,863	101 00000 101 00001 ... 101 10100
(6/7,1/7)	7	0000001 0000010 ... 1000000	0,592	110 000 110 001 110 110
(7/7,0/7)	1	0000000	0	111

Bài 9 Kênh rời rạc không nhớ
Lượng tin tương hỗ

9.1 Kênh rời rạc không nhớ và ma trận kênh

9.2 Entropy điều kiện và lượng tin tương hỗ

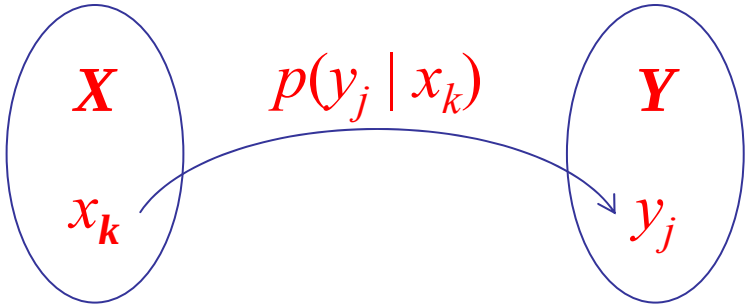
9.3 Một số loại kênh

9.4 Sự nhập nhằng (equivocation) và tốc độ truyền tin

9.5 Dung lượng kênh

Kênh rời rạc không nhớ và ma trận kênh

- Định nghĩa
  - Một kênh rời rạc không nhớ (DMC) được định nghĩa bằng một bảng kí hiệu đầu vào (nguồn phát) $X = \{x_1, ..., x_K\}$, một bảng kí hiệu đầu ra (nguồn nhận) $Y = \{y_1, ..., y_J\}$, và một sự phân bố xác suất có điều kiện $p(y_j \mid x_k)$, với $1 \le k \le K$, $1 \le j \le J$.



  - Bảng kí hiệu đầu ra không nhất thiết giống bảng kí hiệu đầu vào. Điều này có nghĩa là bên nhận có thể nhận những kí hiệu mà không giống với những kí hiệu mà bên phát phát đi.

Nhận xét

- Thuật ngữ không nhớ (memoryless) suy ra rằng

$$p\{y_{j1}\cdots y_{jN} \mid x_{k1}\cdots x_{kN}\} = \prod_{n=1}^{N} p(y_{jn} \mid x_{kn})$$

với N bất kỳ.

- Một kênh rời rạc không nhớ thường được biểu diễn dưới dạng một **ma trận kênh** $[p(y_j \mid x_k)]$ có kích thước $K \times J$.

	y_1	y_2		y_J
x_1	$p(y_1 \mid x_1)$	$p(y_2 \mid x_1)$	...	$p(y_J \mid x_1)$
x_2	$p(y_1 \mid x_2)$	$p(y_2 \mid x_2)$	...	$p(y_J \mid x_2)$
...	...	...	...	...
x_K	$p(y_1 \mid x_K)$	$p(y_2 \mid x_K)$	...	$p(y_J \mid x_K)$

Nhận xét (tt)

- Chúng ta thấy, ma trận kênh chính là cái mà biểu diễn tính chất tạp nhiễu của kênh truyền.
- Chú ý, nếu chúng ta biết sự phân bố xác suất trên X thì sự phân bố xác suất của Y sẽ được xác định như sau

$$p(y_j) = \sum_{k=1}^{K} p(x_k) p(y_j \mid x_k)$$

Entropy điều kiện và lượng tin tương hỗ

- Xét bài toán truyền tin sau

  Cho biết cấu trúc thống kê của nguồn X và ma trận kênh. Hãy xác định kí hiệu x_k nào đã được phát phát đi khi nhận được ở đầu nhận một kí hiệu y_j nào đó?

- Ví dụ
  - Cho nguồn $X = \{x_1, x_2\}$ với các xác suất lần lượt là $p(x_1) = 1/4$, $p(x_2) = 3/4$, nguồn $Y = \{y_1, y_2\}$ và ma trận kênh là

	y_1	y_2
x_1	4/5	1/5
x_2	2/5	3/5

  - Nếu nhận được y_1 thì x_k nào có khả năng đã được phát đi?

Ví dụ

$$p(x_k \mid y_j) = \frac{p(x_k, y_j)}{p(y_j)} = \frac{p(x_k) \times p(y_j \mid x_k)}{\sum_{i=1}^{K} p(x_i, y_j)} = \frac{p(x_k) \times p(y_j \mid x_k)}{\sum_{i=1}^{K} p(x_i) \times p(y_j \mid x_i)}$$

$$p(x_1 \mid y_1) = \frac{p(x_1)p(y_1 \mid x_1)}{p(x_1)p(y_1 \mid x_1) + p(x_2)p(y_1 \mid x_2)} \qquad p(x_2 \mid y_1) = \frac{3}{5}$$

$$= \frac{(1/4) \times (4/5)}{(1/4) \times (4/5) + (3/4) \times (2/5)} = \frac{2}{5}$$

- $p(x_1 \mid y_1) < p(x_2 \mid y_1)$, như vậy chúng ta có thể khẳng định được kí hiệu x_2 có khả năng được phát đi hơn x_1?

Ví dụ (tt)

- Để ý, trong công thức của $p(x_i \mid y_j)$ có chứa thừa số $p(x_i)$, nên $p(x_i \mid y_j)$ đã bị ảnh hưởng bởi xác suất lề $p(x_i)$.
- Vì vậy để công bằng trong việc so sánh chúng ta phải dựa trên tỉ số $p(x_i \mid y_j)/p(x_i)$ cái mà không bị ảnh hưởng nhiều bởi $p(x_i)$.

$$\frac{p(x_1 \mid y_1)}{p(x_1)} = \frac{2/5}{1/4} = \frac{8}{5}$$

$$\frac{p(x_2 \mid y_1)}{p(x_2)} = \frac{3/5}{3/4} = \frac{4}{5}$$

- Như vậy thực sự kí hiệu x_1 mới có khả năng được phát đi hơn kí hiệu x_2.
- Từ xác suất điều kiện chúng ta giới thiệu khái niệm lượng tin có điều kiện.

Lượng tin có điều kiện $I(x_k \mid y_j)$

- Định nghĩa

$$I(y_j \mid x_k) = -\log p(y_j \mid x_k)$$
$$I(x_k \mid y_j) = -\log p(x_k \mid y_j)$$

  - $p(y_j \mid x_k) \to 1$ thì $I(y_j \mid x_k) \to 0$ và ngược lại.
  - Nếu khi phát đi x_k và biết chắc y_j sẽ nhận được thì ở phía nhận chúng ta không cần tốn thêm thông tin gì để giải thích.
  - Nếu $p(y_j \mid x_k) = 1/2$ ($I(y_j \mid x_k) = 1$ bit) thì khi phát đi x_k bên nhận sẽ có hai khả năng và y_j chỉ là một trong hai khả năng đó, có nghĩa là bên nhận cần thêm thông tin (cần thêm 1 bit) để biết chính xác đó là khả năng nào.
  - Xác suất $p(y_j \mid x_k) = 1/2$ chỉ xảy ra khi kênh truyền có nhiễu.

Lượng tin có điều kiện $I(x_k \mid y_j)$

- Vì vậy lượng tin có điều kiện còn được gọi là lượng tin bị mất đi do nhiễu.
- Khi phát đi x_k bên nhận sẽ có một tập các y_j có khả năng được nhận.
- Ngược lại khi nhận được y_j bên phát sẽ có một tập các x_k có khả năng được phát.
- Để đo mức độ "quan hệ" giữa x_k với y_j chúng ta giới thiệu khái niệm lượng tin tương hỗ.

Lượng tin tương hỗ

- Định nghĩa
  - Lượng tin tương hỗ giữa hai tin là lượng tin của của tin này được chứa trong tin kia và ngược lại. Bằng công thức

    Lượng tin tương hỗ = Lượng tin riêng – Lượng tin bị mất đi

    $I(x_k, y_j) = I(x_k) - I(x_k \mid y_j) = I(y_j) - I(x_k \mid y_j)$

    $$= \log \frac{p(x_k|y_j)}{p(x_k)} = \log \frac{p(y_j|x_k)}{p(y_j)}$$

  - Nếu $p(x_k \mid y_j) = 1$ có nghĩa là nếu y_j đã nhận được thì chắc chắn x_k đã được phát đi, điều này có nghĩa là lượng tin của x_k đã được truyền nguyên vẹn thông qua kênh, do đó $I(x_k, y_j) = I(x_k)$.

Lượng tin có điều kiện trung bình

$$I(X \mid y_j) = \sum_{k=1}^{K} p(x_k \mid y_j) I(x_k \mid y_j) = -\sum_{k=1}^{K} p(x_k \mid y_j) \log p(x_k \mid y_j)$$

$$I(Y \mid x_k) = \sum_{j=1}^{J} p(y_j \mid x_k) I(y_j \mid x_k) = -\sum_{j=1}^{J} p(y_j \mid x_k) \log p(y_j \mid x_k)$$

$$I(X \mid Y) = \sum_{j=1}^{J} p(y_j) I(X \mid y_j) = -\sum_{j=1}^{J} p(y_j) \sum_{k=1}^{K} p(x_k \mid y_j) \log p(x_k \mid y_j)$$

$$= -\sum_{k=1}^{K} \sum_{j=1}^{J} p(x_k, y_j) \log p(x_k \mid y_j)$$

$$I(Y \mid X) = -\sum_{k=1}^{K} \sum_{j=1}^{J} p(x_k, y_j) \log p(y_j \mid x_k)$$

Entropy điều kiện

- Định nghĩa
  - Xét hai biến ngẫu nhiên x và y với phân bố xác suất đồng thời $p(x_k, y_j)$, $k = 1, ..., K$, $j = 1, ..., J$. Entropy điều kiện của x đã cho y được định nghĩa là

$$H(\mathbf{x} \mid \mathbf{y}) = -\sum_{k=1}^{K}\sum_{j=1}^{J} p(x_k, y_j) \log p(x_k \mid y_j)$$

  - $H(\mathbf{x} \mid \mathbf{y})$ có thể được diễn dịch như là độ bất ngờ trung bình về x sau khi đã biết y.
- Định lý 9.1
  - $H(\mathrm{x} \mid \mathrm{y}) \le H(\mathrm{x})$, dấu "=" xảy ra $\Leftrightarrow$ x và y là độc lập.

Chứng minh

$$H(\mathrm{x}\,|\,\mathrm{y}) - H(\mathrm{x}) = -\sum_{k=1}^{K}\sum_{j=1}^{J} p(x_k, y_j)\ln p(x_k\,|\,y_j) + \sum_{k=1}^{K} p(x_k)\ln p(x_k)$$

$$= -\sum_{k=1}^{K}\sum_{j=1}^{J} p(x_k, y_j)\ln\frac{p(x_k)}{p(x_k\,|\,y_j)}$$

- Lấy tổng trên những cặp (k, j) mà $p(x_k, y_j) \neq 0$. Vì vậy

$$H(\mathrm{x}\,|\,\mathrm{y}) - H(\mathrm{x}) = \sum_{k}\sum_{j} p(x_k, y_j)\left[\frac{p(x_k)}{p(x_k\,|\,y_j)} - 1\right]$$

$$= \sum_{k}\sum_{j} p(x_k)p(y_j) - p(x_k, y_j)$$

$$= \sum_{k}\sum_{j}\left[p(x_k)p(y_j)\right] - 1 \leq 0$$

Chứng minh (tt)

- Dấu "=" xảy ra $\Leftrightarrow p(x_k) = p(x_k \mid y_j)$ đối với tất cả các cặp (k, j) mà $p(x_k, y_j) \neq 0$ đồng thời tổng $p(x_k)p(y_j)$ trên tất cả những cặp này bằng 1.
- Điều kiện thứ hai tương đương với điều kiện $p(x_k)p(y_j) = 0$ bất kỳ khi nào mà $p(x_k, y_j) = 0$.
- Cả hai điều kiện này cuối cùng tương đương với điều kiện là **x** và **y** độc lập.

- Định lý 9.2
  - $H(\mathrm{x}, \mathrm{y}) = H(\mathrm{x}) + H(\mathrm{y} \mid \mathrm{x}) = H(\mathrm{y}) + H(\mathrm{x} \mid \mathrm{y})$

Chứng minh

$$H(\mathrm{x},\mathrm{y}) = -\sum_k \sum_j p(x_k, y_j) \log p(x_k, y_j)$$

$$= -\sum_k \sum_j p(x_k, y_j) \left[\log p(x_k) + \log p(y_j \mid x_k)\right]$$

$$= -\sum_k p(x_k) \left[\log p(x_k)\right] - \sum_k \sum_j p(y_j, y_k) \log p(y_j \mid x_k)$$

$$= H(\mathrm{x}) + H(\mathrm{y} \mid \mathrm{x})$$

- Phần thứ hai chứng minh hoàn toàn tương tự.
- Kết hợp hai định lý trên chúng ta suy ra rằng

$$H(\mathbf{x}, \mathbf{y}) \le H(\mathbf{x}) + H(\mathbf{y})$$

- dấu "=" xảy ra $\Leftrightarrow$ **x**, **y** là độc lập.

Lượng tin tương hỗ trung bình

$$I(X,Y) = \sum_k \sum_j p(x_k, y_j) I(x_k, y_j)$$

$$= \sum_k \sum_j p(x_k, y_j) \log \frac{p(x_k \mid y_j)}{p(x_k)}$$

$$= \sum_k \sum_j p(x_k, y_j) \log \frac{p(y_j \mid x_k)}{p(y_j)}$$

$$= \sum_k \sum_j p(x_k, y_j) \log \frac{p(x_k, y_j)}{p(x_k) p(y_j)}$$

- Nếu biểu diễn theo entropy thì chúng ta có

$$\boldsymbol{I(x, y) = H(x) - H(x \mid y) = H(y) - H(y \mid x)}$$

Một số loại kênh rời rạc không nhớ

- Kênh đối xứng (Symmetric channel)
  - Là kênh mà mỗi dòng của ma trận kênh chứa cùng tập các số p_1', ..., p_J' và mỗi cột chứa cùng tập các số q_1', ..., q_K'.
- Ví dụ

$[p(y_j \mid x_k)] =$

$j=1$	2	3	4	
0,2	0,2	0,3	0,3	$k=1$
0,3	0,3	0,2	0,2	$k=2$

Các ma trận biểu diễn các kênh đối xứng

$[p(y_j \mid x_k)] =$

0,2	0,3	0,5
0,3	0,5	0,2
0,5	0,2	0,3

Kênh đối xứng nhị phân (binary symmetric channel – BSC).

$[p(y_j \mid x_k)] =$

$1-\varepsilon$	ε	$0 \le \varepsilon \le 1$
ε	$1-\varepsilon$	

Nhận xét

- Kênh đối xứng thì $H(\mathbf{y} \mid \mathbf{x})$ độc lập với sự phân bố xác suất của nguồn phát và được xác định duy nhất bằng ma trận kênh.

- Chứng minh

$$H(\mathrm{y} \mid \mathrm{x}) = -\sum_{k=1}^{K}\sum_{j=1}^{J} p(x_k, y_j)\log p(y_j \mid x_k)$$

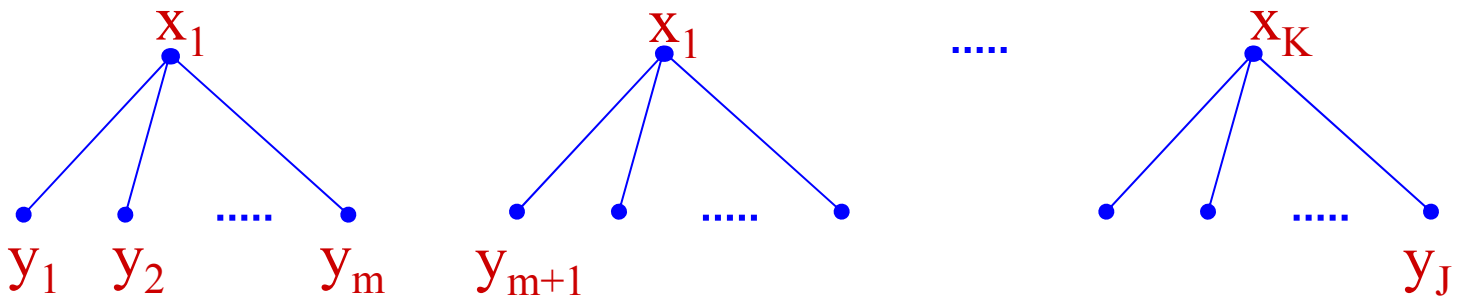
$$= -\sum_{k=1}^{K} p(x_k)\sum_{j=1}^{J} p(y_j \mid x_k)\log p(y_j \mid x_k)$$

$$= -\sum_{k=1}^{K} p(x_k)\sum_{j=1}^{J} p'_j \log p'_j$$

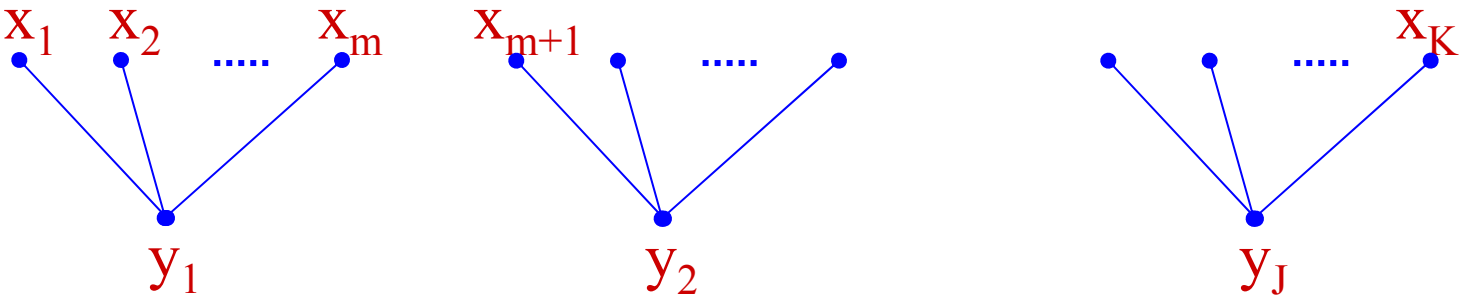
$$= -\sum_{j=1}^{J} p'_j \log p'_j$$

Kênh không mất (Lossless channel)

- Cạnh nối giữa x_k và y_j nghĩa là $p(y_j \mid x_k) \neq 0$. Trong kênh không mất đầu ra xác định duy nhất đầu vào, vì vậy $\boldsymbol{H}(\mathbf{x} \mid \mathbf{y}) = 0$.



- Kênh đơn định (Deterministic channel)



- Trong kênh này đầu vào xác định duy nhất đầu ra, vì vậy $\boldsymbol{H}(\mathbf{y} \mid \mathbf{x}) = 0$

Kênh vô dụng (Useless channel)

- Một kênh là vô dụng nếu và chỉ nếu x và y là độc lập với mọi sự phân bố xác suất của đầu vào (nguồn phát).
- Đối với một kênh vô dụng thì $H(x \mid y) = H(x)$, tức là kiến thức về đầu ra không làm giảm độ bất ngờ về đầu vào. Vì vậy, đối với mục đích xác định đơn định đầu vào, chúng ta có thể phớt lờ đầu ra hoàn toàn. Bây giờ chúng ta sẽ chứng minh rằng.
- Một kênh rời rạc không nhớ là vô dụng nếu và chỉ nếu ma trận kênh của nó có các dòng giống nhau.

- Chứng minh
  - Điều kiện đủ

    Giả sử ma trận có các dòng giống nhau p_1', ..., p_J'. Thì đối với mọi đầu ra y_j

Kênh vô dụng (tt)

$$p(y_j) = \sum_{k=1}^{K} p(x_k, y_j) = \sum_{k=1}^{K} p(x_k)\, p(y_j \mid x_k) = p'_j \sum_{k=1}^{K} p(x_k) = p'_j$$

Đối với mọi cặp đầu vào– ra (x_k, y_j), chúng ta có

$$p(x_k, y_j) = p(x_k)\, p(y_j \mid x_k) = p(x_k)\, p_j' = p(x_k)\, p(y_j)$$

Vì vậy đầu vào và đầu ra độc lập nhau bất chấp sự phân bố xác suất của đầu vào.

- Điều kiện cần

  Giả sử các dòng của ma trận không giống nhau $\Rightarrow \exists$ một cột chẳng hạn j_0 mà có các phần tử không giống nhau.

  Giả sử $p(y_{j0} \mid x_{k0})$ là phần tử lớn nhất trong cột này. Xét sự phân bố đồng nhất (đẳng xác suất) ở đầu vào (đầu phát), chúng ta có

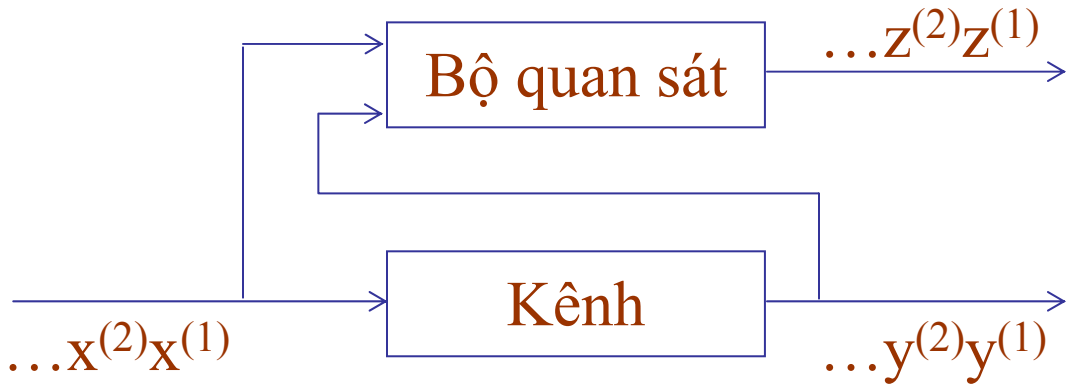
Kênh vô dụng (tt)

$$p(y_{j0}) = \sum_{k=1}^{K} p(x_k)\,p(y_{j0} \mid x_k) = \frac{1}{K}\sum_{k=1}^{K} p(y_{j0} \mid x_k) < p(y_{j0} \mid x_{k0})$$

- Tức là $p(y_{j0}) \neq p(y_{j0} \mid x_{k0})$. Vì vậy $p(x_{k0}, y_{j0}) = p(x_{k0})\,p(y_{j0} \mid x_{k0}) \neq p(x_{k0})\,p(y_{j0})$. Mâu thuẫn với giả thiết là x, y độc lập với mọi sự phân bố xác suất của đầu vào.

Sự nhập nhằng (equivocation) và tốc độ truyền tin

- Xét một kênh nhị phân đối xứng với xác suất chéo ε. Giả sử rằng tại đầu vào $P(0) = P(1) = 1/2$, tốc độ sinh thông tin ở đầu phát là $\boldsymbol{H}(\mathbf{x}) = 1$ bit/kí hiệu.
- Một thiết bị được gọi là bộ quan sát, nhận mỗi cặp kí hiệu vào/ra (x, y) và sinh ra một kí hiệu z
- $z = 0$ nếu $x = y$, $z = 1$ nếu $x \neq y$



Sự nhập nhằng (equivocation) và tốc độ truyền tin (tt)

- Sự phân bố của **z** được tìm thấy như sau:

$$P(z = 1) = P(x = 0)\,P(y = 1 \mid x = 0) + P(x = 1)\,P(y = 0 \mid x = 1)$$
$$= \varepsilon/2 + \varepsilon/2 = \varepsilon$$
$$P(z = 0) = 1 - P(z = 0) = 1 - \varepsilon$$

- Tốc độ sinh thông tin bởi bộ quan sát vì vậy bằng

$$\boldsymbol{H}(\mathbf{z}) = -\varepsilon \log \varepsilon - (1 - \varepsilon) \log(1 - \varepsilon) \text{ bits/kí hiệu}$$

- Đối với một dãy đầu ra đã cho $y^{(1)}y^{(2)}\ldots$, nơi nhận (receiver) có thể xây dựng lại chính xác dãy đầu vào $x^{(1)}x^{(2)}\ldots$ chỉ khi đầu ra của bộ quan sát $z^{(1)}z^{(2)}\ldots$ đã được tạo sẵn.
- Tốc độ truyền thông tin trên kênh, thường kí hiệu là R, là bằng tốc độ sinh thông tin $H(x)$ trừ tốc độ sinh thông tin bổ sung $H(z)$.

$$R = H(x) - H(z)$$

Ví dụ

- Chẳng hạn, nếu dữ liệu đầu vào được sinh ở tốc độ 1000 kí hiệu/giây và $\varepsilon = 0{,}01$, chúng ta có

  $H(x) = 1 \rightarrow$ tốc độ dữ liệu đầu vào $= 1000$ bits/giây

  $H(z) = 0{,}081 \rightarrow$ tốc độ dữ liệu bổ sung $= 81$ bits/giây

  $R = 0{,}919 \rightarrow$ tốc độ truyền thông tin $= 919$ bits/giây

- Một người có thể lý luận rằng trong một dãy dài, vì $\varepsilon = 0{,}01$, nghĩa là chỉ 1% số bit được truyền bị lỗi, và vì vậy tốc độ truyền thông tin phải là 990 bits/giây.
- Câu trả lời là rằng kiến thức về số bit bị lỗi không đủ để xây dựng lại dữ liệu, mà chúng ta cần phải biết thêm về vị trí lỗi nữa, và vì lý do này nên tốc độ truyền thông tin là thực sự bằng một giá trị thấp hơn là 919 bits/giây.

Nhận xét

- Trong trường hợp tốt nhất $\varepsilon = 0$, chúng ta có $H(z) = 0$ và vì vậy $R = 1000$ bits/giây.
- Trong một trường hợp khác nếu $\varepsilon = 1/2$, thì $H(z) = 1$, kết quả là $R = 0$ bits/giây.
- Cả hai kết luận là nhất quán với sự mong đợi của chúng ta.
- Đối với kênh nhị phân đối xứng với đầu vào đẳng xác suất, chúng ta chứng minh được rằng $H(z) = H(x \mid y)$.
- Tổng quát chúng ta chứng minh được rằng
- Sự tái xây dựng chính xác dãy đầu vào từ dãy đầu ra là có thể nếu bộ quan sát có thể sinh ra thông tin bổ sung ở tốc độ lớn hơn hay bằng $H(x \mid y)$.

Nhận xét (tt)

- Để thấy điều này một cách trực quan, quan sát rằng đối với các dãy dài có chiều dài N có khoảng $2^{NH(x \mid y)}$ dãy đầu vào có thể sinh ra một dãy đầu ra cụ thể.
- Chỉ khi thông tin bổ sung được sinh ra tại tốc độ $H(x \mid y)$ hay nhanh hơn mới cho phép phân biệt giữa các khả năng này.
- Đối với lý do này, $H(x \mid y)$ thường được coi như là sự nhập nhằng (equivocation) của kênh. Và chúng ta định nghĩa lại tốc độ truyền thông tin trên kênh là

$$R = H(x) - H(x \mid y) = I(x, y)$$

Dung lượng kênh

- Theo phần trên tốc độ truyền tin trên kênh được định nghĩa là

$$R = H(\mathrm{x}) - H(\mathrm{x} \mid \mathrm{y}) = I(\mathrm{x}, \mathrm{y})$$

- $I(\mathrm{x}, \mathrm{y})$ tổng quát là một hàm của sự phân bố xác suất đầu vào $\{p_1, \ldots, p_K\}$. Vì vậy, có thể tìm thấy một sự phân bố mà cực đại $I(\mathrm{x}, \mathrm{y})$. Giá trị cực đại của $I(\mathrm{x}, \mathrm{y})$ được định nghĩa là dung lượng kênh ***C*** và là một hàm của ma trận kênh.

  C = Cực đại (trên các sự phân bố xác suất đầu vào) của $\boldsymbol{I(\mathrm{x}, \mathrm{y})}$.

- Tổng quát, việc tính dung lượng kênh là một bài toán khó và là một bài toán chưa được giải một cách triệt để.
- Tuy nhiên đối với các kênh đã được giới thiệu ở trên *C* có thể tính toán được như phần sau đây trình bày.

Kênh đối xứng

$$C = \log J + \sum_{j=1}^{J} p'_j \log p'_j$$

trong đó p_1', …, p_J' là các phần tử của các hàng của ma trận.

- Trong trường hợp kênh nhị phân đối xứng với xác suất chéo là p chúng ta có

$$C = 1 - H(p) \text{ với } H(p) = -p\log p - (1-p)\log(1-p)$$

- Kênh không mất
  - $H(\mathbf{x} \mid \mathbf{y}) = 0$, vì vậy

$$C = \text{Max}\ \{H(\mathbf{x}) - H(\mathbf{x} \mid \mathbf{y})\} = \text{Max}\{H(\mathbf{x})\} = \log K$$

trong đó K là kích thước của bảng kí hiệu đầu vào. Dung lượng đạt được trong trường hợp đầu vào có sự phân bố đẳng xác suất.

Kênh đơn định

- Ở đây $H(\mathbf{y} \mid \mathbf{x}) = 0$, vì vậy

$$C = \text{Max}\ \{H(\mathbf{y}) - H(\mathbf{y} \mid \mathbf{x})\} = \text{Max}\{H(\mathbf{y})\} = \log J$$

trong đó J là kích thước của bảng kí hiệu đầu ra.

- Kênh vô dụng
  - Ở đây $H(\mathbf{x} \mid \mathbf{y}) = H(\mathbf{x})$, vì vậy

$$C = \text{Max}\ \{H(\mathbf{x}) - H(\mathbf{x} \mid \mathbf{y})\} = \text{Max}\{H(\mathbf{x}) - H(\mathbf{x})\} = 0$$

  - Một kênh vô dụng thì có dung lượng kênh bằng 0.

Bài 10 Mã hóa chống nhiễu, định lý kênh

10.1 Giới thiệu bài toán chống nhiễu

10.2 Định lý kênh có nhiễu cho kênh nhị phân đối xứng rời rạc (BSC)

10.3 Định lý ngược của kênh truyền có nhiễu

Giới thiệu bài toán chống nhiễu

- Mục tiêu chống nhiễu là bên nhận có thể đoán (giải mã) được càng chính xác càng tốt dãy kí hiệu đã được phát.
- Chẳng hạn xét nguồn nhị phân đối xứng với xác suất chéo ε, đồng thời giả sử nguồn phát đẳng xác suất, tức $P(0) = P(1) = 1/2$.
- Với $\varepsilon < 1/2$, xét cơ chế giải mã ở bên nhận như sau: Nếu $y = 0$ thì đoán $x = 0$ và nếu $y = 1$ thì đoán $x = 1$.
- Xác suất giải mã bị lỗi của cơ chế này là

$$P(\text{lỗi}) = P(y = 0)\, P(x = 1 \mid y = 0) + P(y = 1)\, P(x = 0 \mid y = 1) = \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

- Chú ý trong trường hợp ở đây chúng ta tính được

$$P(y = 0) = P(y = 1) = 1/2 \text{ và } P(x \neq y \mid y) = \varepsilon.$$

- Vấn đề quan trọng là có thể giảm được xác suất giải mã bị lỗi hay không?

Giới thiệu bài toán chống nhiễu (tt)

- Một hướng giải quyết như sau: để gởi 0 chúng ta gởi chuỗi 3 kí hiệu 0 và tương tự để gởi 1 chúng ta gởi 3 kí hiệu 1.
- Cơ chế giải mã của bên nhận như sau: Nếu chuỗi nhận có nhiều kí hiệu 0 hơn 1 thì giải mã thành 0 và ngược lại.
- Chẳng hạn bên nhận nếu nhận được 010 thì giải mã thành 0 còn nếu nhận được 110 thì giải mã thành 1.
- Cơ chế này có xác suất giải mã bị lỗi là

$$P(\text{lỗi}) = 3(1-\varepsilon)\varepsilon^2 + \varepsilon^3 < \varepsilon$$

- Xác suất này nhỏ hơn ε. Tuy nhiên hiệu suất truyền thông tin bị giảm xuống 3 lần.
- Tương tự nếu muốn xác suất giải mã tiến đến 0 chúng ta sẽ mã hoá 0 thành dãy $2n + 1$ kí hiệu 0 và mã hoá 1 thành $2n + 1$ kí hiệu 1, nhưng tương ứng lúc này hiệu suất truyền thông tin giảm xuống $2n + 1$ lần so với ban đầu.

Giới thiệu bài toán chống nhiễu (tt)

- Có một cách có thể giảm xác suất giải mã lỗi xuống gần bằng 0 nhưng không giảm hiệu suất truyền thông tin xuống gần bằng 0 mà chỉ cần nhỏ hơn một *ngưỡng* nào đó là đủ.
- Ngưỡng đó chính là *dung lượng kênh*.
- Cách này cũng khai thác ý tưởng trên ở chỗ thay vì để gởi đi 0 và 1, cái mà có "*khoảng cách Hamming*" giữa chúng là 1 thì chúng ta sẽ mã hoá lần lượt thành 000 và 111, cái mà có "*khoảng cách Hamming*" giữa chúng là 3 và vì vậy giảm xác suất giải mã bị lỗi.

Định lý kênh có nhiễu cho kênh nhị phân đối xứng rời rạc (BSC)

- Xét kênh nhị phân đối xứng với xác suất chéo p.
- Dung lượng kênh trong đơn vị bits/kí hiệu là

$$C = 1 - H(p) \text{ với } H(p) = -p\log p - (1-p)\log(1-p)$$

- Giả sử thời gian truyền 1 kí hiệu là T, số kí hiệu được truyền trong 1 giây là $1/T$, thì dung lượng theo đơn vị bits/giây là

$$C = [1 - H(p)]/T$$

- Xét nguồn X có entropy $H(X)$ bits/ký hiệu, tức là nguồn này tạo ra thông tin ở tốc độ theo đơn vị bits/giây.

$$R = H(X)/T$$

- Định lý 10.1.
  - Chừng nào mà R (bits/giây) còn nhỏ hơn C (bits/giây), thì việc truyền trên kênh với tỉ lệ lỗi nhỏ tuỳ ý là có thể thực hiện được.
  - Để chứng minh định lý này cần một số khái niệm sau.

Các khái niệm

- Trọng số Hamming
  - Trọng số Hamming của một dãy kí hiệu $v = a_1a_2...a_n$, trong đó mỗi $a_i \in \{0, 1, ..., m–1\}$, là số kí hiệu khác 0 của dãy, và thường được kí hiệu là $w(v)$.
- Khoảng cách Hamming
  - Khoảng cách Hamming của hai dãy kí hiệu v_1, v_2 với chiều dài bằng nhau là số vị trí khác nhau của hai dãy, và thường được kí hiệu là $d(v_1, v_2)$.
- Phép cộng cơ số m, $\oplus$
  - Xét $a, b \in \{0, 1, ..., m–1\}$ thì $a \oplus b = (a + b) \bmod m$.
  - Nếu $v_1 = a_1a_2...a_n$, $v_2 = b_1b_2...b_n$ thì $v_1 \oplus v_2 = c_1c_2...c_n$ trong đó $c_i = a_i \oplus b_i$ với $i = 1, 2, ..., n$.

Các khái niệm (tt)

- Ví dụ
  - $w(10100) = 2$, $w(01120) = 3$.
  - $d(10100, 10001) = 2$, $d(011010, 101101) = 5$.
  - Với $m = 2$ thì $1011 \oplus 1101 = 0110$. Với $m = 3$ thì $1021 \oplus 2120 = 0111$.
- Bổ đề

$$d(v_1, v_2) = w(v_1 \oplus v_2)$$
$$d(v_1, v_2) + d(v_2, v_3) \geq d(v_1, v_3)$$

- Nhận xét
  - Bất đẳng thức thứ hai có dạng của bất đẳng thức tam giác: tổng hai cạnh của một tam giác lớn hơn hoặc bằng cạnh còn lại.
  - Định lý 10.1 đúng cho kênh rời rạc không nhớ bất kỳ. Tuy nhiên ở đây chúng ta chỉ chứng minh cho kênh nhị phân đối xứng rời rạc.

Chứng minh định lý

- Ý tưởng chứng minh là mã hoá các dãy dữ liệu thành các từ mã, trong đó các kí hiệu mã lấy từ bảng kí hiệu đầu vào của kênh và xử lý các từ mã này như các đầu vào cơ bản của kênh.
- Xác suất lỗi nhỏ tuỳ ý có thể đạt được dựa trên sự mã hoá như sau:

  (1) chọn chiều dài N của dãy dữ liệu đủ dài

  (2) mã hoá các dãy này thành các từ mã có khoảng cách Hamming xa nhau.
- Nguyên tắc giải mã ở đầu ra được thiết kế như sau: dãy kí hiệu nhận được ở đầu ra $\boldsymbol{v_j}$ sẽ được giải mã thành từ mã $\boldsymbol{w_i}$ mà có khoảng cách Hamming nhỏ nhất đối với $\boldsymbol{v_j}$.
- Với cách chọn này xác suất giải mã lỗi là nhỏ nhất. Thật vậy

$$p(w_i \mid v_j) = p(w_i)p(v_j \mid w_i)/p(v_j)$$

Chứng minh định lý (tt)

- Do đó khi chúng ta không rõ về $p(w_i)$ và dĩ nhiên sẽ kéo theo $p(v_j)$ thì $p(w_i \mid v_j)$ lớn nhất khi $p(v_j \mid w_i)$ là lớn nhất. Mà

$$p(v_j \mid w_i) = p^D(1-p)^{N-D}$$

  trong đó D là khoảng cách Hamming giữa v_j và w_i, N là chiều dài của chúng, p là xác suất chéo.
- Nếu xác suất chéo $p < 0{,}5$ thì $p(v_j \mid w_i)$ sẽ lớn nhất khi D là nhỏ nhất.
- Chứng minh rằng $\forall\ \theta > 0$ nhỏ tuỳ ý, với N đủ lớn tồn tại cách mã hoá các dãy dữ liệu thành các từ mã sao cho với nguyên tắc giải mã trên có xác suất giải mã lỗi là nhỏ hơn θ.
- Thật vậy số dãy dữ liệu có chiều dài N là vào khoảng

$$M = 2^{NH(X)} = 2^{NRT}$$

  trong khi đó tổng số dãy có chiều dài N là 2^N.

Chứng minh định lý (tt)

- Gọi $\{w_1, w_2, \ldots, w_M\}$ là một tập từ mã bất kỳ, P_e là xác suất giải mã lỗi đối với tập này.
- Nếu chúng ta chứng minh được rằng $\forall\ \theta > 0$ nhỏ tuỳ ý, với N đủ lớn giá trị trung bình của P_e, $\overline{P_e}$, nhỏ hơn θ thì sẽ tồn tại một tập từ mã mà có xác suất giải mã lỗi P_e nhỏ hơn θ.
- Với xác suất lỗi trên đường truyền là p, một dãy có chiều dài N sẽ có trung bình Np vị trí lỗi.
- Với hai số dương ε, δ nhỏ tuỳ ý, theo luật yếu của số lớn với N đủ lớn thì xác suất để số vị trí của chuỗi nhận v_j khác với chuỗi phát w_i lớn hơn $N(p + \varepsilon)$ là nhỏ hơn δ. Hay nói theo ngữ cảnh của khoảng cách Hamming là

$$P[d(w_i, v_j) > N(p + \varepsilon)] < \delta$$

- Vì vậy bộ mã mà chúng ta mong muốn sẽ như sau: Khoảng cách Hamming giữa hai từ mã bất kỳ là $\geq 2N(p + \varepsilon) + 1$

Chứng minh định lý (tt)

- Như vậy với mỗi v_j nhận được theo bất đẳng thức tam giác tồn tại một từ mã w_i mà có

$$d(w_i, v_j) \le N(p + \varepsilon)$$

còn các từ mã w_k khác có

$$d(w_k, v_j) \ge N(p + \varepsilon) + 1$$

- Vì vậy chúng ta sẽ giải mã được duy nhất v_j thành w_i.
- Với ý tưởng này, chúng ta sẽ đưa ra cơ chế giải mã lỏng hơn cho một tập từ mã bất kỳ $\{w_1, w_2, \ldots, w_M\}$, nhưng cũng sẽ đảm bảo xác suất giải mã lỗi là nhỏ hơn θ.
- Với mỗi dãy v_j nhận được, định nghĩa một tập kiểm tra A_j bao gồm tất cả những dãy có chiều dài N và có khoảng cách Hamming so với v_j nhỏ hơn hay bằng $N(p + \varepsilon)$.
- Nếu từ mã được truyền w_i là từ mã duy nhất thuộc tập A_j thì giải mã v_j thành w_i. Ngược lại thông báo một lỗi đã xảy ra.

Chứng minh định lý (tt)

- Một lỗi xảy ra thuộc vào một trong hai trường hợp sau đây

  (1) Từ mã được truyền w_i không thuộc A_j, tức là

  $$d(w_i, v_j) > N(p + \varepsilon)$$

  Lỗi này xảy ra với xác suất nhỏ hơn δ.

  (2) Tồn tại một từ mã w_k khác cũng thuộc A_j. Lúc này chúng ta không biết nên giải mã v_j thành w_i hay w_k.
- Chúng ta chứng minh rằng theo cách này xác suất giải mã lỗi trung bình sẽ nhỏ hơn θ với θ nhỏ tuỳ ý cho trước.
- Chúng ta có

$$P_e \leq \delta + \sum_{\substack{i=1 \\ i \neq j}}^{M} P(w_i \in A_j)$$

Chứng minh định lý (tt)

- Để tính $\overline{P_e}$ chúng ta sẽ tính giá trị trung bình của $P(w_i \in A_j)$.
- Giá trị trung bình này sẽ bằng số dãy thuộc tập A_j chia cho tổng số dãy

$$\overline{P}(W_i \in A_j) = \frac{\sum_{k=0}^{N(p+\varepsilon)} \binom{N}{k}}{2^N}$$

- Suy ra

$$\overline{P_e} < \delta + (M-1)\frac{\sum_{k=0}^{N(p+\varepsilon)} \binom{N}{k}}{2^N}$$

Chứng minh định lý (tt)

- Mà chúng ta có một bất đẳng thức nổi tiếng sau

$$\sum_{k=0}^{N\alpha} \binom{N}{k} \le 2^{NH(\alpha)}$$

với $H(\alpha) = -\alpha \log\alpha - (1-\alpha)\log(1-\alpha)$.

- Áp dụng vào bất đẳng thức trên chúng ta có

$$\overline{P_e} \le \delta + M \times 2^{-N[1-H(p+\varepsilon)]} = \delta + 2^{NRT}\, 2^{-N[1-H(p+\varepsilon)]}$$
$$= \delta + 2^{-N[1-H(p+\varepsilon)-RT]}$$

- Vì ε và δ có thể nhỏ tuỳ ý, nên chừng nào $R < [1 - H(p)]/T = C$ (bits/giây) thì có thể làm cho $\overline{P_e}$ nhỏ tuỳ ý bằng cách tăng N.
- Chứng minh được hoàn tất.

Ví dụ

- Xét ví dụ trước đây, một kênh đối xứng nhị phân có xác suất chéo $\varepsilon = 0{,}01$. Tốc độ truyền kí hiệu $f = 1000$ kí hiệu/giây (tức $T = 0{,}001$ giây). Chúng ta có $C = 0{,}919$ bits/kí hiệu hay $C = 919$ bits/giây.
- Định lý kênh cho phép chúng ta kết luận, với xác suất đúng tiến tới 1, rằng với N khá lớn chẳng hạn $N = 1000$, thì trong 2^{1000} dãy có chiều dài 1000 chúng ta có thể chọn được 2^K dãy với $K < 919$ sao cho khoảng cách Hamming giữa các dãy là $\geq 2N\varepsilon + 1 = 21$.

- Khoảng cách Hamming của bộ mã
  - Khoảng cách Hamming của một bộ mã A, với điều kiện A là mã đều, kí hiệu là $d(A)$, là khoảng cách Hamming nhỏ nhất trong tất cả các khoảng cách giữa hai từ mã bất kỳ của A.

Định lý

- Định lý 10.2
  - Một bộ mã nhị phân có khoảng cách Hamming d thì có thể
    - Phát hiện sai được t bit nếu $d \geq t + 1$.
    - Sửa sai được t bit nếu $d \geq 2t + 1$.
- Chứng minh
  - Gọi w_i là từ mã phát, v_i là dãy nhận tương ứng. Nếu sai tối đa $t > 0$ bit thì $d(w_i, v_i) \leq t$. Do đó tổ hợp nhận sẽ không thể trùng với bất kỳ từ mã nào vì khoảng cách Hamming giữa hai từ mã bất kỳ là $\geq t + 1$. Vì vậy bên nhận có thể phát hiện được sai.
  - Tương tự nếu $d(w_i, w_j) \geq 2t + 1$, theo bất đẳng thức tam giác $\Rightarrow$ $d(w_j, v_i) \geq t + 1$ $\forall$ từ mã $w_j \neq w_i$. Vì vậy bên nhận có thể giải mã đúng v_i thành w_i dựa trên sự khác biệt này.

Định lý ngược của kênh truyền có nhiễu

- Định lý 10.2
  - Nếu tốc độ truyền tin *R* (bits/giây) lớn hơn dung lượng kênh *C* (bits/giây), thì sự truyền thông trên kênh với tỉ lệ lỗi nhỏ tuỳ ý là không thể thực hiện được. Hay nói cách khác xác suất giải mã lỗi tiến đến 1 khi chiều dài của dãy cần truyền gia tăng.
  - Định lý này nói cách khác nếu tốc độ truyền tin lớn hơn dung lượng kênh thì việc truyền không được đảm bảo có nghĩa là chúng ta không thể giải mã đúng được.

Bài 11 Cơ sở toán học của mã chống nhiễu

- Bài này trình bày các cơ sở toán học của mã khối tuyến tính.
- Các kiến thức này là rất quan trọng để hiểu được cách xây dựng các loại mã khối tuyến tính.
- Các khái niệm được trình bày bao gồm các cấu trúc đại số như nhóm, trường và đặc biệt là các trường $GF(2)$ và $GF(2^m)$, đây là các trường có ứng dụng đặc biệt vào trong việc xây dựng các mã khối tuyến tính chống nhiễu.

Bài 11 Cơ sở toán học của mã chống nhiễu

11.1 Một số khái niệm cơ bản

11.2 Trường $GF(2)$ và các đa thức trên trường $GF(2)$

11.3 Trường $GF(2^m)$

Một số khái niệm cơ bản

- Phép toán đóng
  - Cho G là một tập hợp, một phép toán hai ngôi f được gọi là đóng trên G nếu f có dạng

    $$f : G \times G \to G$$

    tức là nếu $a, b \in G$ thì $f(a, b) \in G$.
- Chú ý
  - $f(a, b)$ có một cách viết tương đương là afb và ngược lại $f(b, a)$ còn được viết là bfa. Chẳng hạn nếu f là phép cộng thì thay vì viết $+(a, b)$ chúng ta thường viết là $a + b$.
  - Kể từ đây trở về sau khi nói đến một phép toán nếu chúng ta không nói gì thêm thì có nghĩa là phép toán này có tính đóng.

Một số khái niệm cơ bản (tt)

- Tính kết hợp
  - Một phép toán hai ngôi f trên G được gọi là có tính kết hợp nếu $\forall\ a, b, c \in G$ thì

$$(afb)fc = af(bfc)$$

- Tính giao hoán
  - Một phép toán hai ngôi f trên G được gọi là có tính giao hoán nếu $\forall\ a, b \in G$ thì

$$afb = bfa$$

- Ví dụ
  - Trên tập số thực khác 0, phép cộng và phép nhân có tính kết hợp và giao hoán nhưng phép trừ và phép chia không có tính kết hợp và giao hoán.

Nhóm

- Tính phân phối
  - Phép toán f_1 được gọi là có tính phân phối đối với phép toán f_2 nếu $\forall\ a, b, c \in G$ thì

$$af_1(bf_2c) = (af_1b)f_2(af_1c)$$

  - Chẳng hạn trên tập số thực, phép nhân có tính phân phối đối với phép cộng vì $\forall\ a, b, c \in R$

$$a\times(b+c) = (a\times b)+(a\times c)$$

- Nhóm
  - Một tập $G \neq \varnothing$, với một phép toán hai ngôi f được gọi là một nhóm nếu thoã 3 điều kiện sau:

    (1) f có tính kết hợp.

Nhóm (tt)

(2) G chứa phần tử e, sao cho $\forall\ a \in G$ thì $afe = efa = a$. e được gọi là phần tử trung hoà (đối với một số phép toán e còn được gọi là phần tử đơn vị)

(3) Mọi phần tử đều có phần tử đối xứng, tức là $\forall\ a \in G$, tồn tại phần tử $b \in G$ sao cho

$$afb = bfa = e$$

- Chẳng hạn, trên tập R nếu f là phép cộng thì phần tử trung hoà là số 0, còn trên tập số thực khác 0 nếu f là phép nhân thì phần tử trung hoà là 1 và còn được gọi là phần tử đơn vị.

- Nhóm giao hoán
  - Một nhóm mà phép toán f có tính giao hoán thì được gọi là nhóm giao hoán.

Nhóm (tt)

- Nhóm hữu hạn, nhóm vô hạn
  - Một nhóm có số phần tử hữu hạn được gọi là nhóm hữu hạn, một nhóm có số phần tử vô hạn được gọi là nhóm vô hạn.
- Nhóm con
  - Cho G là một nhóm. Một tập H con của G được gọi là một nhóm con nếu H đóng với phép toán hai ngôi của G và thoã điều kiện của một nhóm.
  - Tập các số chẵn ≥ 0 là một nhóm con của tập số tự nhiên với phép cộng thông thường.

Phép cộng và nhân modulo

- Phép cộng modulo và phép nhân modulo
  - Cho một số nguyên dương m xác định. Xây dựng một tập số nguyên sau $G = \{0, 1, \ldots, m-1\}$. Với + là phép cộng thông thường. Định nghĩa phép toán mới $\oplus$ như sau và gọi là phép cộng modulo

$$\forall\ a, b \in G \text{ thì } a \oplus b = (a + b) \bmod m$$

  - Tương tự với $\times$ là phép nhân thông thường. Định nghĩa phép toán mới $\otimes$ như sau và gọi là phép nhân modulo

$$\forall\ a, b \in G \text{ thì } a \otimes b = (a \times b) \bmod m$$

Ví dụ

- Tập R là một nhóm giao hoán đối với phép cộng và là một nhóm vô hạn.
- Tập $R - \{0\}$ là một nhóm giao hoán đối với phép nhân và là một nhóm vô hạn.
- Với m là một số nguyên dương xác định, tập $G = \{0, 1, \ldots, m - 1\}$ với phép cộng modulo là một nhóm giao hoán và là một nhóm hữu hạn.
- Hai bảng sau đây trình bày lần lượt trường hợp $m = 5$ và $m = 6$.

Ví dụ (tt)

$m = 5$

$\oplus$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$m = 6$

$\oplus$	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

- Tương tự tập $G = \{1, \ldots, m-1\}$ với phép nhân modulo và m nguyên tố là một nhóm giao hoán hữu hạn.

Bổ đề

- Bổ đề 11.1
  - Nếu m là một số nguyên tố thì $G = \{0, 1, \ldots, m-1\}$ là một nhóm giao hoán với phép nhân modulo $\otimes$. Ngược lại nếu m không nguyên tố thì G không là một nhóm.

$m = 5$

$\otimes$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

$m = 6$

$\otimes$	1	2	3	4	5
1	1	2	3	4	5
2	2	4	0	2	4
3	3	0	3	0	3
4	4	2	0	4	2
5	5	4	3	2	1

Trường

- Trường
  - Một tập G với hai phép toán đóng hai ngôi bất kỳ, chẳng hạn kí hiệu là + và *, được gọi là một trường nếu thoã 3 điều kiện sau

    (1) G là nhóm giao hoán đối với phép +. Phần tử trung hoà trong phép + thường được kí hiệu là 0.

    (2) Tập các phần tử khác 0 là một nhóm đối với phép *. Phần tử trung hoà trong phép * thường được gọi là phần tử đơn vị và kí hiệu là 1.

    (3) Phép * có tính phân phối đối với phép +.
- Chú ý
  - Phép + và phép * ở trên không nhất thiết là phép cộng và phép nhân thông thường mà chúng có thể là bất kỳ phép nào. Chúng ta kí hiệu như vậy để dễ trình bày.

Trường (tt)

- Các phần tử của một trường không nhất thiết là các số nguyên hay thực mà có thể là bất kỳ cái gì, chẳng hạn có thể là các số phức, vectơ, ma trận hay đa thức ...
- Từ định nghĩa của trường chúng ta suy ra một trường bao gồm tối thiểu hai phần tử: phần tử trung hoà của phép + (kí hiệu là 0) và phần tử trung hoà của phép * (kí hiệu là 1). Các phần tử 0 và 1 không nhất thiết là số 0 và số 1 theo nghĩa thông thường mà có thể là bất kỳ cái gì chẳng hạn là ma trận 0 và ma trận đơn vị, ...

- Trường giao hoán
  - Một trường mà phép * có tính giao hoán thì được gọi là trường giao hoán.

Trường (tt)

- Chẳng hạn trong ví dụ ở slide 198 với $m = 5$ chúng ta thấy G là một trường giao hoán.
- Tổng quát chúng ta có bổ đề sau và để dành việc chứng minh cho các bạn sinh viên.

- Bổ đề 11.2
  - Cho p là một số nguyên tố bất kỳ, $G = \{0, 1, ..., p - 1\}$ thì G là một trường giao hoán đối với phép cộng modulo $\oplus$ và phép nhân modulo $\otimes$.
  - Sau đây là một số tính chất của trường
- Tính chất 1
  - Mọi phần tử a của trường đều thoã $a * 0 = 0$.

Trường Galois

- Tính chất 2
  - Nếu a, b là hai phần tử khác 0 của trường thì $a * b \neq 0$.
- Tính chất 3
  - Nếu $a \neq 0$ và $a * b = a * c$ thì $b = c$. Hay nói cách khác nếu $a \neq 0$ và $b \neq c$ thì $a * b \neq a * c$.
- Bậc của một trường, trường hữu hạn, trường vô hạn.
  - Số phần tử của một trường được gọi là bậc của một trường. Một trường có số phần tử hữu hạn được gọi là trường hữu hạn, một trường có số phần tử vô hạn được gọi là trường vô hạn.
- Trường $GF(q)$
  - Một trường có số phần tử hữu hạn được gọi là trường Galois. Nếu bậc của trường Galois là q thì trường được kí hiệu là $GF(q)$.

Trường Galois

- Đối với các trường hữu hạn tức là trường Galois chúng ta có định lý sau.

- Định lý 11.1
  - Một trường hữu hạn thì số phần tử của nó phải có dạng p^m trong đó p là một số nguyên tố còn m là một số nguyên dương. Hay nói cách khác các trường Galois đều có dạng $GF(p^m)$ trong đó p là một số nguyên tố còn m là một số nguyên dương.
  - Đối với các trường $GF(p)$ với p nguyên tố thì đó chính là tập $\{0, 1, 2, ..., p-1\}$ với hai phép toán cộng modulo $\oplus$ và nhân modulo $\otimes$ như đã biết.
  - Đối với các trường $GF(p^m)$, vì tính phức tạp của chúng, chúng ta sẽ giới thiệu sau. Chú ý lúc này các phần tử của trường $GF(p^m)$ không đơn thuần là các số mà sẽ có dạng khá đặc biệt.

Trường Galois (tt)

- Kí hiệu các phần tử đối xứng
  - Phần tử đối xứng của a trong phép + được kí hiệu là $-a$, phần tử đối xứng của a trong phép * được kí hiệu là a^{-1}.
- Phép – và phép /
  - Đối với một trường giao hoán, từ hai phép + và phép * chúng ta định nghĩa thêm hai phép – và phép / như sau (không nhất thiết là phép trừ và phép chia bình thường)

$$a - b = a + (-b)$$

$$a / b = a * b^{-1}$$

trong đó $-b$ là phần tử đối xứng của b qua phép +, còn b^{-1} là phần tử đối xứng của b qua phép *.

  - Vậy một trường giao hoán G có bốn phép toán +, –, *, /. Phép + và – đóng trên G, phép * và / đóng trên $G - \{0\}$.

Trị riêng của một trường

- Xét một trường $GF(q)$. Xét các dãy tổng của các phần tử đơn vị

$$\sum_{i=1}^{k} 1 = 1 + 1 + \cdots + 1 \quad (k \text{ lần, với } k = 1, 2, 3, \ldots)$$

- Vì trường đóng với phép cộng nên kết quả của những tổng này cũng là các phần tử của trường. Vì k có thể nhận vô hạn giá trị mà trường chỉ có q phần tử nên tồn tại hai giá trị k_1 và k_2 khác nhau (giả sử $k_1 > k_2$) sao cho

$$\sum_{i=1}^{k_1} 1 = \sum_{i=1}^{k_2} 1$$

- Từ đây suy ra

$$\sum_{i=1}^{k_1 - k_2} 1 = 0$$

Trị riêng của một trường

- Trị riêng của một trường kí hiệu là số nguyên dương nhỏ nhất λ sao cho

$$\sum_{i=1}^{\lambda} 1 = 0$$

- Dễ thấy đối với các trường $GF(p) = \{0, 1, 2, ..., p-1\}$ với p là một số nguyên tố thì trị riêng $\lambda = p$. Tổng quát chúng ta có định lý sau.

- Định lý 11.2
  - Trị riêng λ của một trường $GF(q)$ là một số nguyên tố.
- Chứng minh
  - Giả sử λ không nguyên tố $\Rightarrow \lambda = k \times l$ (k, l nguyên > 1). Từ qui tắc phân phối của phép nhân đối với phép cộng suy ra

Trị riêng của một trường (tt)

$$\sum_{i=1}^{k} 1 \times \sum_{i=1}^{l} 1 = \sum_{i=1}^{k \times l} 1 = \sum_{i=1}^{\lambda} 1 = 0$$

Suy ra $$\sum_{i=1}^{k} 1 = 0 \qquad \sum_{i=1}^{l} 1 = 0$$

Mà $k, l < \lambda$, điều này mâu thuẫn với định nghĩa của λ.

- Chu kỳ của một phần tử
  - Xét một phần tử a bất kỳ khác 0 của trường $GF(q)$. Xét các luỹ thừa a^k của a với $k = 1, 2, 3, \ldots$ Vì trường đóng với phép nhân nên các a^k cũng là các phần tử của trường. Vì k có thể nhận vô hạn giá trị mà trường chỉ có q phần tử nên tồn tại hai giá trị k_1 và k_2 khác nhau (giả sử $k_1 > k_2$) sao cho

$$a^{k1} = a^{k2} \Rightarrow a^{k1-k2} = 1$$

Chu kỳ của một phần tử

- Chu kỳ của một phần tử a của một trường $GF(q)$ là số nguyên dương nhỏ nhất n sao cho $a^n = 1$.

- Ví dụ
  - Xét trường $GF(7) = \{0, 1, 2, 3, 4, 5, 6\}$ với hai phép $\oplus$ và $\otimes$. Trị riêng của trường này là 7. Còn chu kỳ của các phần tử khác 0 của trường được trình bày trong bảng sau

Phần tử	1	2	3	4	5	6
Chu kỳ	1	3	6	3	6	2

  - Từ định nghĩa trên chúng ta thấy dãy các luỹ thừa của a

$$a^1, a^2, ..., a^k, ..., a^n = 1, a^{n+1} = a, ...$$

sẽ lặp lại sau n phần tử.

Nhóm tuần hoàn

- Bổ đề 11.3
  - Dãy $a^1, a^2, ..., a^k, ..., a^n = 1$ tạo nên một nhóm con đóng với phép nhân trên trường $GF(q)$.
- Nhóm tuần hoàn
  - Một nhóm (không chứa phần tử 0) với phép nhân * được gọi là tuần hoàn nếu tồn tại một phần tử trong nhóm mà các luỹ thừa của nó tạo nên mọi phần tử trong nhóm.
  - Từ định nghĩa này suy ra một nhóm hữu hạn được gọi là tuần hoàn nếu tồn tại một phần tử trong nhóm có chu kỳ đúng bằng số phần tử của nhóm.
- Định lý 11.3
  - Nếu a là một phần tử khác 0 của một trường $GF(q)$ thì

$$a^{q-1} = 1$$

Nhóm tuần hoàn (tt)

- Chứng minh
  - Gọi $b_1, b_2, ..., b_{q-1}$ là $q - 1$ phần tử khác nhau và khác 0 của trường. Theo tính chất 3 và tính chất 2 của trường chúng ta có $a*b_1, a*b_2, ..., a*b_{q-1}$ cũng là $q - 1$ phần tử khác nhau và khác 0 của trường. Vì vậy chúng ta có

$$a*b_1*a*b_2* \ ... \ *a*b_{q-1} = b_1*b_2* \ ... \ *b_{q-1}$$

  - Từ đây suy ra $a^{q-1} = 1$. Hoàn tất chứng minh.
- Định lý 11.4
  - Chu kỳ của một phần tử bất kỳ khác 0 của một trường $GF(q)$ là ước số của $q - 1$.

Phần tử cơ sở

- Chứng minh
  - Gọi n là chu kỳ của phần tử a khác 0 của trường $GF(q)$. Giả sử $q-1$ không chia hết cho n. Do đó $q-1 = kn + r$, trong đó r là số dư của phép chia $q-1$ cho n, $0 < r < n$. Chúng ta có

$$a^{q-1} = a^{kn+r} = (a^n)^k * a^r$$

    Do $a^{q-1} = 1$ và $a^n = 1$ suy ra $a^r = 1$. Mà $0 < r < n$ điều này mâu thuẫn với định nghĩa chu kỳ của a. Vậy $q-1$ chia hết cho n.
- Phần tử cơ sở
  - Một phần tử a khác 0 của một trường $GF(q)$ được gọi là phần tử cơ sở nếu chu kỳ của a bằng $q-1$.
  - Từ định nghĩa này $\Rightarrow$ nếu a là một phần tử cơ sở thì các luỹ thừa của a gồm $a^0 = 1, a^1 = a, a^2, \ldots, a^{q-2}$ hình thành nên $q-1$ phần tử khác 0 của trường.

Ví dụ

- Xét trường $GF(7)$ như trong ví dụ ở slide 209. Chu kỳ của các phần tử khác 0 của trường đều là ước số của 6. Đặc biệt các phần tử 3 và 5 có chu kỳ bằng 6 nên chúng là các phần tử cơ sở của trường $GF(7)$.

$3^1 = 3$	$3^2 = 2$	$3^3 = 6$	$3^4 = 4$	$3^5 = 5$	$3^6 = 1$
$5^1 = 5$	$5^2 = 4$	$5^3 = 6$	$5^4 = 2$	$5^5 = 3$	$5^6 = 1$

- Trong các trường Galois thì trường $GF(2)$ và trường $GF(2^m)$ là những trường có nhiều ứng dụng đặc biệt trong lý thuyết mã, nên chúng ta sẽ chỉ trình bày hai trường này.

Trường *GF*(2)

- Trường *GF*(2)
  - Trường *GF*(2) bao gồm hai phần tử {0, 1} với hai phép cộng + và nhân * như sau

+	0	1
0	0	1
1	1	0

*	0	1
0	0	0
1	0	1

  - Phần tử đối xứng của 0 và 1 qua phép cộng cũng chính là 0 và 1. Phần tử đối xứng của 1 qua phép nhân cũng chính là 1.
  - Trong trường *GF*(2) thì phép trừ giống với phép cộng, phép chia cho một số khác 0 cũng giống với phép nhân.

Các đa thức trên trường $GF(2)$

- Đa thức trên trường $GF(2)$
  - Một đa thức trên trường $GF(2)$, chẳng hạn kí hiệu là $f(x)$, là đa thức có dạng

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

    trong đó các hệ số $a_i \in GF(2)$.
- Bậc của đa thức
  - Là bậc lớn nhất của đa thức.
- Ví dụ
  - Đa thức $f(x) = 1 + x + x^3$ có bậc 3 đa thức $g(x) = x + x^2 + x^5$ có bậc 5.

Các đa thức trên trường $GF(2)$ (tt)

- Phép cộng đa thức và nhân đa thức
  - Với $f(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$, $g(x) = b_0 + b_1x + b_2x^2 + \ldots + b_nx^n$ với các hệ số a_i và b_j thuộc trường $GF(2)$ chúng ta định nghĩa các phép cộng đa thức và nhân đa thức như sau

$$f(x) + g(x) = \sum_{i=0}^{n} (a_i + b_i)x^i$$

$$f(x) * g(x) = \sum_{i,j=0}^{n} (a_i * b_j)x^{i+j}$$

trong đó $a_i + b_i$ và $a_i * b_j$ được thực hiện trên trường $GF(2)$.

Các đa thức trên trường $GF(2)$ (tt)

- Ví dụ
  - Cho $f(x) = 1 + x + x^3$, $g(x) = x + x^2$ thì

$$f(x) + g(x) = (1 + x + x^3) + (x + x^2) = 1 + x^2 + x^3$$

$$f(x) * g(x) = (1 + x + x^3) * (x + x^2) = x + x^3 + x^4 + x^5$$

  - Nếu $g(x)$ có bậc khác 0 thì chúng ta có thể chia $f(x)$ cho $g(x)$ và có thể viết như sau

$$f(x) = q(x) * g(x) + r(x)$$

    trong đó $q(x)$ là đa thức thương còn $r(x)$ là đa thức dư có bậc nhỏ hơn đa thức chia $g(x)$.

- Ví dụ
  - $f(x) = 1 + x + x^4 + x^5 + x^6$ chia cho $g(x) = 1 + x + x^3$

Các đa thức trên trường *GF*(2) (tt)

- $1 + x + x^4 + x^5 + x^6 = (x^2 + x^3) * (1 + x + x^3) + (1 + x + x^2)$
- Để phân tích một đa thức ra thành các thừa số trong đại số Euclid chúng ta có
- Nếu $f(a) = 0$ thì $f(x)$ chia hết cho $(x - a)$.
- Điều này cũng đúng trên trường *GF*(2).

- Ví dụ
  - $f(x) = 1 + x + x^3 + x^5$ có $f(1) = 0$, nên $f(x)$ chia hết cho $(x - 1)$ mà trong trường *GF*(2), phép trừ cũng chính là phép cộng tức là $f(x)$ chia hết cho $(x + 1)$.

$$1 + x + x^3 + x^5 = (1 + x)(1 + x^3 + x^4)$$

Đa thức tối giản

- Một đa thức trên $GF(2)$ được gọi là tối giản nếu nó không phân tích được thành tích của hai đa thức có bậc nhỏ hơn.

1, 2, 3, 4	5	6
x	$1 + x^2 + x^5$	$1 + x + x^6$
$1 + x$	$1 + x^3 + x^5$	$1 + x^3 + x^6$
$1 + x + x^2$	$1 + x + x^2 + x^3 + x^5$	$1 + x + x^2 + x^4 + x^6$
$1 + x + x^3$	$1 + x + x^2 + x^4 + x^5$	$1 + x + x^3 + x^4 + x^6$
$1 + x^2 + x^3$	$1 + x + x^3 + x^4 + x^5$	$1 + x^5 + x^6$
$1 + x + x^4$	$1 + x^2 + x^3 + x^4 + x^5$	$1 + x + x^2 + x^5 + x^6$
$1 + x^3 + x^4$		$1 + x^2 + x^3 + x^5 + x^6$
$1 + x + x^2 + x^3 + x^4$		$1 + x + x^4 + x^5 + x^6$
		$1 + x^2 + x^4 + x^5 + x^6$

Bổ đề

- Cho $f(x)$ là một đa thức trên trường $GF(2)$, thì

$$f(x)^{2^n} = f(x^{2^n})$$

- Chứng minh
  - Đặt $f(x) = a_0 + a_1 x + \ldots + a_n x^n$.

$$\begin{aligned}
[f(x)]^2 &= (a_0 + a_1 x + \ldots + a_n x^n)^2 \\
&= a_0^2 + a_0*(a_1 x + \ldots + a_n x^n) + a_0*(a_1 x + \ldots + a_n x^n) + \\
&\quad (a_1 x + \ldots + a_n x^n)^2 \\
&= a_0^2 + (a_1 x + \ldots + a_n x^n)^2 \\
&= a_0^2 + (a_1 x)^2 + \ldots + (a_n x^n)^2 \\
&= f(x^2) \text{ (vì trong } GF(2) \; a_i^2 = a_i \text{)}
\end{aligned}$$

  - Điều này cũng giúp chúng ta suy ra điều phải chứng minh.

Trường $GF(2^m)$

- Trước hết chúng ta kí hiệu trường $GF(2^m)$ như sau

$$GF(2^m) = \{0, 1, a_1, a_2, ..., a_{2^m-2}\}$$

trong đó 0 và 1 $\in GF(2)$. Trường $GF(2)$ là một trường con của $GF(2^m)$ và được gọi là trường cơ sở của $GF(2^m)$.

- Chú ý
  - Nếu a là một phần tử $\in GF(2^m)$, $f(x)$ là một đa thức trên trường $GF(2)$, thì $f(a)$ cũng là một phần tử của $GF(2^m)$.
  - Có vô hạn đa thức $f(x)$ trên trường $GF(2)$ mà chỉ có hữu hạn (2^m) phần tử $\in GF(2m)$, nên $\forall\ a \neq 0$ của $GF(2^m)$ tồn tại hai đa thức $f_1(x)$ và $f_2(x)$ khác nhau sao cho $f_1(a) = f_2(a)$. Từ đây nếu đặt $f(x) = f_1(x) - f_2(x)$ (chú ý trong trường $GF(2)$ thì phép – giống với phép cộng +) thì $f(a) = 0$.

Đa thức tối thiểu

- Đa thức tối thiểu (minimal polinomial)
  - Cho a là một phần tử khác 0 của trường $GF(2^m)$. Đa thức tối thiểu của a là đa thức $f(x)$ khác 0 trên trường $GF(2)$ và có bậc nhỏ nhất sao cho $f(a) = 0$.
  - Một lần nữa ta phải chú ý rằng khi chúng ta viết $f(\alpha) = 0$ hoặc $f(\alpha) = 1$ thì các kí hiệu 0 và 1 không nhất thiết là các số 0 và 1, mà sẽ được hiểu tuỳ theo ngữ cảnh.
  - Chẳng hạn nếu phần tử α là một ma trận thì 0 chính là ma trận 0 còn 1 chính là ma trận đơn vị.

Đa thức tối thiểu (tt)

- Ví dụ
  - Chẳng hạn nếu *a* là ma trận 4 × 4 bên $T_{4\times4} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$

    trong đó các phép cộng và nhân trên ma trận vẫn thực hiện như bình thường với chú ý rằng các phép cộng và nhân các phần tử của ma trận được thực hiện trên trường GF(2).
  - Chúng ta có thể kiểm tra rằng

    $$1 + T + T^4 = 0$$

    với chú ý rằng 1 là ma trận đơn vị, còn 0 là ma trận 0.
  - Và $f(x) = 1 + x + x^4$ là đa thức tối thiểu của *a*

Định lý

- Hơn nữa chúng ta cũng có

$$T^{15} = 1$$

và chúng ta có thể kiểm tra rằng 15 chính là chu kỳ của a.

- Định lý 11.5
  - Cho a là một phần tử khác 0 của trường $GF(2^m)$ có bậc của đa thức tối thiểu của a là k. Gọi Z là tập tất cả các phần tử có dạng

$$b_0 + b_1 a + \ldots + b_{k-1} a^{k-1}$$

trong đó $b_i \in GF(2)$. Thì Z là một tập con của $GF(2^m)$ và hình thành nên một trường có 2^k phần tử.

Chứng minh

- Đầu tiên chúng ta chứng minh các phần tử được hình thành từ $b_0 + b_1 a + \ldots + b_{k-1} a_{k-1}$ là khác nhau bằng cách chứng minh các phần tử $1, a, a^2, \ldots, a^{k-1}$ là độc lập tuyến tính.

  Thật vậy nếu
  $$\sum_{i=0}^{k-1} b_i a^i = \sum_{i=0}^{k-1} c_i a^i$$

  thì
  $$p(a) = \sum_{i=0}^{k-1} (b_i - c_i) a^i = 0$$

  Vậy chúng ta có đa thức $p(x)$ có bậc nhỏ hơn k thoã $p(a) = 0$. Mà bậc của đa thức tối thiểu của a bằng k. Vậy $p(x) = 0$, suy ra $b_i = c_i \; \forall \; i = 0, 1, ..., k-1$.

Chứng minh (tt)

- Thứ hai, rõ ràng Z là một nhóm giao hoán đối với phép +.

  Thật vậy nếu

$$\sum_{i=0}^{k-1} b_i a^i \in Z, \sum_{i=0}^{k-1} c_i a^i \in Z \Rightarrow \sum_{i=0}^{k-1} (b_i + c_i) a^i = \sum_{i=0}^{k-1} (c_i + b_i) a^i \in Z$$

  Để chứng minh tập $Z_0 = Z - \{0\}$ là một nhóm đối với phép nhân * chúng ta chứng minh nếu

$$\sum_{i=0}^{k-1} b_i a^i \in Z_0, \sum_{i=0}^{k-1} c_i a^i \in Z_0 \Rightarrow \sum_{i=0}^{k-1} b_i a^i * \sum_{i=0}^{k-1} c_i a^i \in Z_0$$

  Gọi $f(\mathrm{x}) = \sum_{i=0}^{k} d_i \mathrm{x}^i$ là đa thức tối thiểu của a, trong đó hệ số cao nhất $d_k = 1$.

Chứng minh (tt)

Từ đây suy ra $$x^k = \sum_{i=0}^{k-1} d_i x^i$$

Do đó mọi a^n với $n \geq k$ đều có thể biểu diễn thông qua một đa thức $g(a)$ nào đó có bậc $\leq k-1$. Vì vậy $$\sum_{i=0}^{k-1} b_i a^i * \sum_{i=0}^{k-1} c_i a^i$$

cũng vậy. Suy ra $$\sum_{i=0}^{k-1} b_i a^i * \sum_{i=0}^{k-1} c_i a^i \in Z$$

Và rõ ràng nếu

$$\sum_{i=0}^{k-1} b_i a^i \neq 0, \sum_{i=0}^{k-1} c_i a^i \neq 0 \Rightarrow \sum_{i=0}^{k-1} b_i a^i * \sum_{i=0}^{k-1} c_i a^i \neq 0$$

Tính chất này được kế thừa từ trường $GF(2^m)$.

Hệ quả

- Cuối cùng tính phân phối của phép nhân * đối với phép cộng + chúng ta cũng kế thừa từ trường $GF(2^m)$. Chứng minh hoàn tất.

- Hệ quả 11.1
  - Nếu đa thức tối thiểu của phần tử $a \in GF(2^m)$ có bậc bằng m thì trường Z trùng với trường $GF(2^m)$ và mỗi phần tử của trường có thể được biểu diễn như một vectơ m thành phần

$$(b_0 b_1 \ldots b_{m-1})$$

- Định lý 11.6
  - Gọi $f(x)$ là đa thức tối thiểu của phần tử $a \neq 0$ của trường $GF(2^m)$ thì $f(x)$ là đa thức tối giản trên trường GF(2).

Chứng minh

- Giả sử $f(x) = g(x) * h(x)$ trong đó $g(x)$ và $h(x)$ có bậc lớn hơn 0 và nhỏ hơn bậc của $f(x)$. Chúng ta có $f(a) = g(a) * h(a) = 0$, suy ra $g(a) = 0$ hoặc $h(a) = 0$. Điều này mâu thuẫn với định nghĩa về đa thức tối thiểu của a.

- Bổ đề 11.5
  - Cho $f(x)$ là đa thức tối thiểu của phần tử $a \neq 0$ của trường $GF(2^m)$ và $p(x)$ là đa thức bất kỳ trên trường $GF(2)$ sao cho $p(a) = 0$. Thì $p(x)$ chia hết cho $f(x)$.
- Chứng minh
  - Chia $p(x)$ cho $f(x)$ chúng ta được

$$p(x) = g(x) * f(x) + r(x)$$

trong đó bậc của $r(x)$ nhỏ hơn bậc của $f(x)$.

Thay $x = a \Rightarrow r(a) = 0$, nên từ định nghĩa của đa thức tối thiểu $\Rightarrow r(x) = 0 \Rightarrow p(x)$ chia hết cho $f(x)$.

Định lý

- Định lý 11.7
  - Cho $f(x)$ là đa thức tối thiểu của phần tử $a \neq 0$ của trường $GF(2^m)$ và $p(x)$ là đa thức tối giản trên trường $GF(2)$ sao cho $p(a) = 0$. Thì $f(x) = p(x)$.
- Chứng minh
  - Theo Bổ đề 11.5 trên chúng ta có $p(x)$ chia hết cho $f(x)$ tức là chúng ta có thể viết

    $$p(x) = g(x) * f(x)$$

    Do $p(x)$ là đa thức tối giản nên $f(x) = 1$ hoặc $f(x) = p(x)$. Tuy nhiên $f(x)$ không thể bằng 1 nên suy ra $f(x) = p(x)$.
- Hệ quả 11.2
  - $2^m - 1$ phần tử khác 0 của trường $GF(2^m)$ đều là nghiệm của phương trình

    $$x^{2^m - 1} + 1 = 0$$

Hệ quả

- Hệ quả 11.3
  - $x^{2^m-1}+1$ chia hết cho các đa thức tối thiểu của các phần tử khác 0 của trường $GF(2^m)$.
  - Chúng ta sẽ dẫn ra một hệ quả mạnh hơn như sau. Trước hết chúng ta phân tích

    $$x^{2^m-1}+1$$

    thành tích của các đa thức tối giản trên trường $GF(2)$

    $$x^{2^m-1}+1 = p_1(x) * p_2(x) * \ldots * p_l(x)$$

    Vì $x^{2^m-1}+1$ có các nghiệm là các phần tử của trường $GF(2^m)$ nên các phần tử của trường $GF(2^m)$ sẽ là nghiệm của một $p_i(x)$ nào đó và ngược lại một $p_i(x)$ bất kỳ sẽ có các nghiệm là các phần tử của trường $GF(2^m)$.

    Hơn nữa nếu $p_i(x)$ có bậc t thì sẽ có t nghiệm trong trường $GF(2^m)$.

Hệ quả

- Hệ quả 11.4
  - Trong việc triển khai $x^{2^m-1}+1$ thành tích các đa thức tối giản, thì mỗi đa thức tối giản sẽ là đa thức tối thiểu của một phần tử khác 0 nào đó của trường $GF(2^m)$.
- Định lý 11.8
  - Cho a là một phần tử khác 0 của trường $GF(2^m)$ và $f(x)$ là một đa thức trên trường $GF(2)$. Nếu $f(a) = 0$ thì

$$f(a^{2^l}) = 0 \ \forall \ l = 0, 1, 2, \ldots$$

- Hệ quả 11.5
  - Nếu $f(x)$ là đa thức tối thiểu của phần tử $a \neq 0$ của trường $GF(2^m)$ thì $f(x)$ cũng là đa thức tối thiểu của các phần tử

$$a^{2^l} \text{ với } l = 0, 1, 2, \ldots \text{ của trường } GF(2^m).$$

Hệ quả (tt)

- Hay nói cách khác các phần tử a^{2^l} với $l = 0, 1, 2, \ldots$ là các nghiệm của đa thức tối thiểu $f(x)$ của phần tử a.
- Hơn nữa chúng ta sẽ chứng minh rằng ngoài chúng ra $f(x)$ không còn nghiệm nào khác thuộc trường $GF(2^m)$.
- Vì vậy nếu có bao nhiêu phần tử a^{2^l} khác nhau thì $f(x)$ có bậc bấy nhiêu.
- Để làm rõ điều này gọi k là số nguyên dương nhỏ nhất sao cho

$$a^{2^k} = a$$

Số k chắc chắn tồn tại vì chúng ta đã có

$$a^{2^m - 1} = 1 \text{ hay } a^{2^m} = a$$

Và số k biểu diễn chu kỳ của dãy

$$a^{2^l} \text{ với } l = 0, 1, 2, \ldots$$

Bổ đề

- Bổ đề 11.6
  - Cho a là một phần tử khác 0 của trường $GF(2^m)$ và k là số nguyên dương nhỏ nhất sao cho $a^{2^k} = a$ thì k là một ước số của m.
- Chứng minh
  - Chia m cho k, $m = n \times k + r$, trong đó r là số dư và $r < k$

$$a^{2^k} = a \Rightarrow \left(a^{2^k}\right)^{2^k} = a^{2^k} = a \text{ hay } a^{2^{2k}} = a$$

Tiếp tục theo kiểu này chúng ta có $a^{2^{nk}} = a$. Mặt khác chúng ta có

$$a = a^{2^m} = a^{2^{n\times k+r}} = a^{2^{n\times k}\times 2^r} = \left(a^{2^{n\times k}}\right)^{2^r} = (a)^{2^r}$$

Bổ đề (tt)

Do định nghĩa của $k \Rightarrow r = 0$. Hoàn tất chứng minh.

- Phần tử liên hợp
  - Cho a là một phần tử khác 0 của trường $GF(2^m)$ và k là số nguyên dương nhỏ nhất sao cho

$$a^{2^k} = a$$

thì các phần tử a^{2^l} với $l = 0, 1, 2, ..., k - 1$ được gọi là các phần tử liên hợp của a và k được gọi là số phần tử liên hợp của a.

  - Từ định nghĩa chúng ta thấy tập các phần tử liên hợp của a là tập các phần tử khác nhau được sinh ra bởi

$$a^{2^l} \text{ với } l = 0, 1, 2, ...$$

- Bổ đề 11.7
  - Nếu a_1 và a_2 là các phần tử liên hợp bất kỳ của a thì a_1 là phần tử liên hợp của a_2 và ngược lại a_2 là phần tử liên hợp của a_1.

Bổ đề (tt)

- Thật vậy giả sử (với k là số phần tử liên hợp của a)

$$a_1 = a^{2^{l_1}}, a_2 = a^{2^{l_2}}, 0 \le l_1 < l_2 < k$$

Thì

$$a_1{}^{2^{l_2 - l_1}} = (a^{2^{l_1}})^{2^{l_2 - l_1}} = a^{2^{l_1} \times 2^{l_2 - l_1}} = a^{2^{l_2}} = a_2$$

và

$$a_2{}^{2^{k + l_1 - l_2}} = (a^{2^{l_2}})^{2^{k + l_1 - l_2}} = a^{2^{l_2} \times 2^{k + l_1 - l_2}}$$

$$= a^{2^{k + l_1}} = (a^{2^k})^{2^{l_1}} = a^{2^{l_1}} = a_1$$

Hoàn tất chứng minh.

- Chú ý bổ đề này nói lên rằng các phần tử liên hợp của a là liên hợp với nhau.

Bổ đề (tt)

- Vì các phần tử a^{2^l} với $l = 0, 1, 2, ..., k - 1$ là các nghiệm của đa thức tối thiểu $f(x)$ của a, nên ta sẽ chứng minh $f(x)$ có dạng

$$f(x) = (x + a) * (x + a^2) * \ldots * (x + a^{2^{k-1}}) = \prod_{i=0}^{k-1} (x + a^{2^i})$$

- Để chứng minh điều này chúng ta sẽ chứng minh

$$p(x) = \prod_{i=0}^{k-1} (x + a^{2^i})$$

là một đa thức tối giản và do $p(a) = 0$ nên theo Định lý 11.7 chúng ta suy ra $f(x) = p(x)$.

Bổ đề (tt)

- Bổ đề 11.7
  - Cho a là một phần tử khác 0 của trường $GF(2^m)$ và k là số nguyên dương nhỏ nhất sao cho

$$a^{2^k} = a$$

thì

$$p(\mathrm{x}) = \prod_{i=0}^{k-1} (\mathrm{x} + a^{2^i})$$

là một đa thức tối giản trên trường $GF(2)$.

- Chứng minh

$$[p(\mathrm{x})]^2 = \left[\prod_{i=0}^{k-1} (\mathrm{x} + a^{2^i})\right]^2 = \prod_{i=0}^{k-1} (\mathrm{x} + a^{2^i})^2$$

Chứng minh (tt)

$$(\mathrm{x}+a^{2^i})^2 = \mathrm{x}^2 + (a^{2^i} + a^{2^i})\mathrm{x} + (a^{2^i})^2 = \mathrm{x}^2 + a^{2^{i+1}}$$

$$\Rightarrow [p(\mathrm{x})]^2 = \prod_{i=0}^{k-1}(\mathrm{x}^2 + a^{2^{i+1}}) = \prod_{i=1}^{k}(\mathrm{x}^2 + a^{2^i})$$

$$= \prod_{i=1}^{k-1}(\mathrm{x}^2 + a^{2^i})(\mathrm{x}^2 + a^{2^k})$$

$$= \prod_{i=1}^{k-1}(\mathrm{x}^2 + a^{2^i})(\mathrm{x}^2 + a) = \prod_{i=0}^{k-1}(\mathrm{x}^2 + a^{2^i})$$

$$= p(\mathrm{x}^2)$$

Chứng minh (tt)

Mặt khác $p(x)$ là một đa thức của x và có thể biểu diễn

$$p(x) = b_0 + b_1 x + \ldots + b_k x^k$$

trong đó các b_i với $i = 0, 1, 2, \ldots, k$ là các đa thức trên trường $GF(2)$ của a. Vì vậy các b_i là các phần tử của trường $GF(2^m)$.

$$[p(x)]^2 = (b_0 + b_1 x + \ldots b_k x^k)^2$$

$$= \sum_{i=0}^{k} b_i{}^2 x^{2i} + (1+1) \sum_{i=0}^{k} \sum_{\substack{j=0 \\ i \neq j}}^{k} b_i b_j x^{i+j} = \sum_{i=0}^{k} b_i{}^2 x^{2i}$$

Do $[p(x)]^2 = p(x^2)$ suy ra

$$b_i = b_i^2 \; \forall \; i = 0, 1, 2, \ldots, k$$

Điều này chỉ đúng nếu các b_i bằng phần tử 0 hoặc phần tử 1 tức là các $b_i \in GF(2)$ hay $p(x)$ là một đa thức trên trường $GF(2)$.

Chứng minh (tt)

Nếu $p(x)$ không tối giản tức $p(x)$ có thể phân tích thành

$$p(x) = q(x) * h(x)$$

trong đó bậc của $q(x)$ và $h(x)$ nhỏ hơn bậc của $p(x)$ là k. Nhưng do $p(a) = 0 \Rightarrow q(a) = 0$ hoặc $h(a) = 0$.

Giả sử $q(a) = 0$, theo Định lý 12.8 $\Rightarrow q(x)$ có các nghiệm là a^{2^l} với $l = 0, 1, 2, ..., k - 1$, $\Rightarrow q(x)$ có bậc tối thiểu là k, mẫu thuẫn. Từ đây suy ra điều phải chứng minh.

- Định lý 11.9
  - Cho a là một phần tử khác 0 của trường $GF(2^m)$ và k là số nguyên dương nhỏ nhất sao cho

$$a^{2^k} = a$$

thì đa thức tối thiểu của a là $f(x) = \prod_{i=0}^{k-1} (x + a^{2^i})$ và có bậc $= k$.

Hệ quả

- Hệ quả 11.6
  - Bậc của một đa thức tối thiểu của một phần tử a khác 0 của trường $GF(2^m)$ là một ước số của m.
- Định lý 11.10
  - Cho a là một phần tử khác 0 của trường $GF(2^m)$ có chu kỳ bằng n thì các phần tử liên hợp của a cũng có chu kỳ bằng n.
- Chứng minh
  - Gọi k là số thành phần liên hợp của a. $\forall\ i = 0, 1, ..., k$

$$\left(a^{2^i}\right)^n = a^{2^i \times n} = (a^n)^{2^i} = 1$$

  - Chúng ta chứng minh rằng không $\exists$ số nguyên dương $l < n$

$$\left(a^{2^i}\right)^l = 1$$

a^{2^l}

Chứng minh

Thật vậy giả sử tồn tại, suy ra $a^{2^i \times l} = 1$

Chia $2i \times l$ cho n

$$2^i \times l = h \times n + r$$

trong đó $0 \le r < n$,

$$\Rightarrow 1 = a^{2^i \times l} = a^{h \times n + r} = (a^n)^h \times a^r = a^r$$

Từ định nghĩa khái niệm chu kỳ, $\Rightarrow r = 0$.

Từ Định lý 11.4 $\Rightarrow n$ là một ước số của $2m - 1$, $\Rightarrow n$ lẽ. Kết hợp với $2^i \times l = h \times n \Rightarrow n$ là một ước số của l, $\Rightarrow n \le l$ vô lý.

- Định lý 11.11
  - $\forall\ m \ge 1$ đều tồn tại một đa thức tối giản bậc m trên trường $GF(2)$.

Định lý

- Định lý 11.12
  - Với một đa thức tối giản $p(x)$ bất kỳ có bậc m,

$$p(x) = b_0 + b_1 x + \ldots + b_m x^m$$

trong đó $b_m = 1$, chúng ta luôn xây dựng được một trường $GF(2^m)$ trong đó $p(x)$ là đa thức tối thiểu của một phần tử của trường.

Để xây dựng trường $GF(2m)$ chúng ta cho phần tử a là một ma trân $m \times m$ như bên

$$T_{m\times m} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ b_0 & b_1 & b_2 & b_3 & \cdots & b_{m-2} & b_{m-1} \end{bmatrix}$$

Định lý (tt)

- Trên ma trận định nghĩa phép cộng và nhân ma trận như bình thường, với chú ý rằng việc cộng hoặc nhân hai phần tử trong 2 ô của hai ma trận được thực hiện như trên trường $GF(2)$.
- Chúng ta công nhận rằng ma trận này có đa thức tối thiểu là $p(x)$. Từ đây chúng ta có thể dẫn ra được các phần tử còn lại của trường $GF(2^m)$ nhờ Định lý 11.5.
- Chú ý, phần tử 0 chính là ma trận 0 và phần tử 1 chính là ma trận đơn vị.

- Định lý 11.13
  - $\forall\ m \geq 2$, các đa thức tối giản bậc m trên trường GF(2) đều là ước số của

$$x^{2^m - 1} + 1$$

  - Chúng ta có thể quay trở về bảng liệt kê các đa thức tối giản để kiểm tra rằng các đa thức tối giản bậc 3 là ước số của $x^7 - 1$, các đa thức tối giản bậc 4 là ước số của $x^{15} - 1$, …

Định lý (tt)

- Đa thức căn bản
  - Một đa thức căn bản là một đa thức tối giản, đồng thời không tồn tại số nguyên dương $n < 2^m - 1$ sao cho $x^n + 1$ chia hết cho nó.
  - Ví dụ, không tồn tại số nguyên dương $n < 15$ sao cho $x^n + 1$ chia hết cho $1 + x + x^4$ nên $1 + x + x^4$ là đa thức căn bản.
  - Còn đa thức $1 + x + x^2 + x^3 + x^4$ là tối giản nhưng không căn bản vì $x^5 + 1$ chia hết cho nó.

1, 2, 3	4, 5	6, 7, 8
$1 + x$	$1 + x + x^4$	$1 + x + x^6$
$1 + x + x^2$	$1 + x^3 + x^4$	$1 + x^3 + x^7$
$1 + x + x^3$	$1 + x^2 + x^5$	$1 + x^2 + x^3 + x^4 + x^8$
$1 + x^2 + x^3$	$1 + x^3 + x^5$	

Định lý (tt)

- Định lý 11.14
  - Cho a là một phần tử khác 0 của trường $GF(2^m)$ có đa thức tối thiểu là $f(x)$. Nếu $f(x)$ là một đa thức căn bản trên trường $GF(2)$ và có bậc bằng m thì a có chu kỳ là $2^m - 1$ và a là một phần tử cơ sở của $GF(2^m)$.
- Chứng minh
  - Gọi n là chu kỳ của a. Đặt $p(x) = x^n + 1$, thì $p(a) = 0$.

    Bổ đề 11.5 $\Rightarrow$ $p(x)$ chia hết cho $f(x)$. Kết hợp điều này với định nghĩa của khái niệm đa thức căn bản, $\Rightarrow$ $n = 2^m - 1$.
  - Định lý này gợi ý cho chúng ta cách xây dựng trường $GF(2^m)$ dựa trên một phần tử cơ sở có đa thức tối thiểu là một đa thức căn bản bậc m.

a^{2^l}

Tóm tắt

- Tóm tắt
  - a là một phần tử của trường $GF(2^m)$ thì $a^{2^m-1}=1$
  - Chu kỳ của một phần tử là một ước số của $2^m - 1$.
  - Các đa thức tối thiểu của trường $GF(2^m)$ là các đa thức tối giản và là ước số của

    $$x^{2^m-1}+1$$

    Hơn nữa bậc của chúng là ước của m.
  - Số phần tử liên hợp khác nhau của a, kể cả a, là ước số của m. Các phần tử liên hợp của nhau có cùng đa thức tối thiểu, hơn nữa chúng là các nghiệm của đa thức tối thiểu này và bậc của đa thức tối thiểu này bằng số các phần tử liên hợp khác nhau. Các phần tử liên hợp thì có cùng chu kỳ.
  - Các đa thức tối giản bậc k là ước của $x^{2^k-1}+1$

Tóm tắt (tt)

- Một phần tử a có đa thức tối thiểu bậc m thì các tổ hợp tuyến tính (với $b_i \in GF(2)$)

$$b_0 1 + b_1 a + \ldots + b_{k-1} a^{k-1}$$

sẽ sinh ra toàn bộ các phần tử của trường $GF(2^m)$

- Một phần tử a có đa thức tối thiểu bậc m và cũng là đa thức căn bản thì các lũy thừa của nó sẽ sinh ra toàn bộ các phần tử của trường $GF(2^m)$.

Ví dụ

- Xây dựng trường $GF(2^m)$ với $m = 4$.
- Chúng ta kí hiệu 0 là ma trận 0, kí hiệu 1 là ma trận đơn vị (có kích thước là 4×4). Lấy phần tử a là ma trận sau

$$T_{4\times4} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

  Chúng ta có đa thức tối thiểu của a là $f(x) = 1 + x + x^4$
- Đây là một đa thức căn bản bậc 4. Vì vậy theo Định lý 11.14, 15 phần tử của $GF(2^4)$ không tính phần tử 0 sẽ có dạng a^i, $i = 0, 1, \ldots, 14$ với chú ý $a^0 = 1$.
- Còn theo Định lý 12.12 chúng sẽ có dạng $b_0 + b_1a + b_2a^2 + b_3a^3$ trong đó các $b_i = 0$ hoặc 1.

Ví dụ (tt)

- Vậy có hai cách để xây dựng trường $GF(2^4)$ như trên.
- Các bảng sau đây biểu diễn các phần tử khác 0 và khác 1 của trường $GF(2^4)$ theo các dạng: lũy thừa của a (a^i), đa thức của a, vectơ, dạng ma trận.

a	a^2	a^3	a^4	a^5
a	a^2	a^3	$1 + a$	$a + a^2$
0100	0010	0001	1100	0110
$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$

Ví dụ (tt)

a^6	a^7	a^8	a^9	a^{10}
$a^2 + a^3$	$1 + a + a^3$	$1 + a^2$	$a + a^3$	$1 + a + a^2$
0011	1101	1010	0101	1110
$\begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

Ví dụ (tt)

a^{11}	a^{12}	a^{13}	a^{14}
$a + a^2 + a^3$	$1 + a + a^2 + a^3$	$1 + a^2 + a^3$	$1 + a^3$
0111	1111	1011	1001
$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

- Chu kỳ, đa thức tối thiểu của các phần tử liên hợp

0	1	a, a^2, a^4, a^8	a^3, a^6, a^{12}, a^9	a^5, a^{10}	$a^7, a^{14}, a^{13}, a^{11}$
		15	5	3	15
x	$1 + x$	$1 + x + x^4$	$1 + x + x^2 + x^3 + x^4$	$1 + x + x^2$	$1 + x^3 + x^4$

Bài 12 Mã khối tuyến tính

12.1 Giới thiệu

12.2 Các khái niệm và nguyên lý hoạt động

12.3 Vấn đề phát hiện sai và sửa sai

12.4 Một số giới hạn

Giới thiệu

- Mã khối tuyến tính được xây dựng dựa trên các kết quả của đại số tuyến tính là một lớp mã được dùng rất phổ biến trong việc chống nhiễu.

- Định nghĩa
  - Một mã khối có chiều dài n gồm 2^k từ mã được gọi là mã tuyến tính $C(n, k)$ nếu và chỉ nếu 2^k từ mã hình thành một không gian vectơ con k chiều của không gian vectơ n chiều gồm tất cả các vectơ n thành phần trên trường $GF(2)$.
  - Mã tuyến tính $C(n, k)$ có mục đích mã hoá những khối tin (hay thông báo) k bit thành những từ mã n bit. Hay nói cách khác trong n bit của từ mã có chứa k bit thông tin.
  - Qui ước viết dấu + thay cho dấu $\oplus$ và dấu + sẽ được hiểu theo ngữ cảnh.

Cách biểu diễn mã – Ma trận sinh

- Mã tuyến tính $C(n, k)$ là một không gian con k chiều của không gian vectơ n thành phần, $\Rightarrow \exists\ k$ từ mã độc lập tuyến tính, chẳng hạn $(g_0, g_1, ..., g_{k-1})$ sao cho mỗi từ mã trong C là một tổ hợp tuyến tính của k từ mã này (với $a_i \in \{0, 1\}\ \forall\ i = 0, 1, ..., k-1$)

$$w = a_0 g_0 + a_1 g_1 + ... + a_{k-1} g_{k-1}$$

- k từ mã này tạo thành một ma trận cấp $k \times n$ như sau

$$G_{k\times n} = \begin{bmatrix} g_0 \\ g_1 \\ \vdots \\ g_{k-1} \end{bmatrix} = \begin{bmatrix} g_{00} & g_{01} & \cdots & g_{0(n-1)} \\ g_{10} & g_{11} & \cdots & g_{1(n-1)} \\ \vdots & \vdots & & \vdots \\ g_{(k-1)0} & g_{(k-1)1} & \cdots & g_{(k-1)(n-1)} \end{bmatrix}$$

- Với $g_i = (g_{i0}, g_{i1}, \ldots, g_{i(n-1)})$, với $i = 0, 1, \ldots, k-1$.

Cách mã hóa

- Nếu $u = (a_0, a_1, \ldots, a_{k-1})$ là thông tin cần được mã hoá thì từ mã w tương ứng với u được ta bằng cách lấy u nhân với G

$$w = u \times G = (a_0, a_1, \ldots, a_{k-1})$$

hay

$$w = a_0 g_0 + a_1 g_1 + \ldots + a_{k-1} g_{k-1}$$

- Vì các từ mã tương ứng với các thông báo được sinh ra bởi G theo cách trên nên G được gọi là ma trận sinh của bộ mã.

Ví dụ

- Cho ma trận sinh của một mã tuyến tính $C(7, 4)$ sau

$$G_{4\times 7} = \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ g_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

- Nếu $u = (1101)$ là thông tin cần mã hoá thì từ mã tương ứng là

$$w = 1.g_0 + 1.g_1 + 0.g_2 + 1.g_3 \quad = (1100101)$$

- Bất kỳ k từ mã độc lập tuyến tính nào cũng có thể được dùng để làm ma trận sinh cho bộ mã.
- Một bộ mã tuyến tính (hay còn gọi là không gian mã) có thể có nhiều ma trận sinh khác nhau cùng biểu diễn.
- Mỗi ma trận sinh tương ứng với một cách mã hóa khác nhau.

Cách giải mã

- Lấy ma trận sinh như ở trong ví dụ trên.
- $u = (a_0, a_1, a_2, a_3)$ là thông báo, $w = (b_0, b_1, b_2, b_3, b_4, b_5, b_6)$ là từ mã tương ứng.
- Chúng ta có hệ phương trình sau liên hệ giữa u và w.

$$w = u \times G \quad \Leftrightarrow$$

$$G_{4\times 7} = \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ g_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$b_0 = a_0 + a_1 + a_3 \quad (1)$$
$$b_1 = a_0 + a_2 \quad (2)$$
$$b_2 = a_1 + a_3 \quad (3)$$
$$b_3 = a_0 + a_1 \quad (4)$$
$$b_4 = a_1 \quad (5)$$
$$b_5 = a_2 \quad (6)$$
$$b_6 = a_2 + a_3 \quad (7)$$

Cách giải mã (tt)

- Chọn bốn phương trình đơn giản nhất để giải các a_i theo các b_j. Chẳng hạn các phương trình (4), (5), (6), (7) chúng ta giải được

$$G_{4\times 7} = \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ g_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad \begin{aligned} a_0 &= b_3 + b_4 \\ a_1 &= b_4 \\ a_2 &= b_5 \\ a_3 &= b_5 + b_6 \end{aligned}$$

- Hệ phương trình trên được gọi là hệ phương trình giải mã.
- Có thể có nhiều hệ phương trình giải mã khác nhau nhưng sẽ cho kết quả như nhau.

$$w = 1001011 \quad \Rightarrow \quad u = ?$$
$$w = 0101110 \quad \Rightarrow \quad u = ?$$

Mã tuyến tính hệ thống

- Một mã tuyến tính $C(n, k)$ được gọi là mã tuyến tính hệ thống nếu mỗi từ mã có một trong hai dạng sau
- Dạng 1: Từ mã bao gồm phần thông tin k bit đi trước và phần còn lại (gồm $n - k$ bit) đi sau (phần này còn được gọi là phần dư thừa hay phần kiểm tra).

k bit thông tin	$n - k$ bit kiểm tra

- Dạng 2: Ngược của dạng 1, từ mã bao gồm phần kiểm tra đi trước và phần thông tin đi sau.

$n - k$ bit kiểm tra	k bit thông tin

Ma trận sinh hệ thống

$$G_{k\times n} = \left[I_{kk} \mid P_{k(n-k)} \right] = \left[\begin{array}{cccc|ccc} 1 & 0 & \cdots & 0 & P_{00} & P_{01} & \cdots & P_{0(n-k-1)} \\ 0 & 1 & \cdots & 0 & P_{10} & P_{11} & \cdots & P_{1(n-k-1)} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & P_{(k-1)0} & P_{(k-1)1} & \cdots & P_{(k-1)(n-k-1)} \end{array} \right]$$

$k\times k$; $k\times(n-k)$

- Ví dụ

$$G_{ht(4\times 7)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

Mã hóa

$u = 1101 \Rightarrow w = u \times G_{ht} = 1101000$

Giải mã

$w = 0110100 \Rightarrow u = 0110$

Ví dụ

- Dùng các phép biến đổi sơ cấp biến đổi các ma trận sinh sau thành ma trận sinh hệ thống.

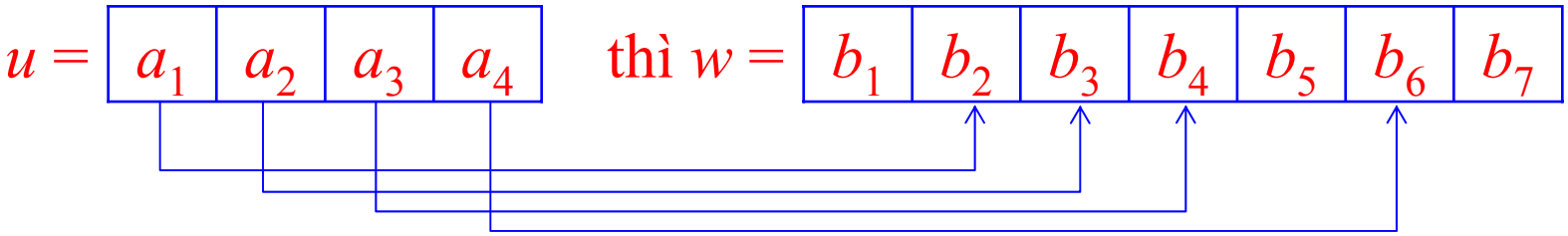
$$G_{4\times 7} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \qquad G_{4\times 7} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

- Không phải mọi ma trận sinh đều có thể biến đổi thành ma trận sinh hệ thống.

Ví dụ (tt)

- Hãy thực hiện phép mã hóa và giải mã trên ma trận sinh sau.

$$G_{4\times 7} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$



- Mã hóa $u = (1101) \Rightarrow w = (1110010)$
- Giải mã $w = (1011000) \Rightarrow u = (0110)$

Phát hiện sai và sửa sai

- Nguyên lý phát hiện sai: Kiểm tra xem tổ hợp nhận có phải là từ mã hay không, nếu không thì tổ hợp nhận là sai.
- Nguyên lý sửa sai: Kiểm tra xem tổ hợp nhận có khoảng cách Hamming gần với từ mã nào nhất, thì đó chính là từ mã đúng đã được phát đi.
- Nguyên lý này được gọi là nguyên lý khoảng cách Hamming tối thiểu.

- Không gian bù trực giao
  - Cho S là một không gian con k chiều của không gian V n chiều. Gọi S_d là tập tất cả các vectơ v trong V sao cho $\forall\ u \in S, u \times v = 0$ (phép nhân vô hướng của hai vectơ). S_d được chứng minh là một không gian con của V và có số chiều là $n - k$. S_d được gọi là không gian bù trực giao của S và ngược lại.

Cách phát hiện sai

- Hệ quả
  - Mỗi ma trận G bất kỳ kích thước $k \times n$ với k hàng độc lập tuyến tính luôn tồn tại ma trận H kích thước $(n-k) \times n$ với $(n-k)$ hàng độc lập tuyến tính sao cho $G \times H^T = 0$, trong đó H^T là ma trận chuyển vị của ma trận H.
  - Nói cách khác các vectơ hàng của H đều trực giao với các vectơ hàng của G.
- Cách phát hiện sai
  - Nếu v là một từ mã được sinh ra từ ma trận sinh G có ma trận trực giao tương ứng là H thì

$$v \times H^T = 0$$

  - Ngược lại nếu

$$v \times H^T = 0$$

    thì v là một từ mã.

Ma trận kiểm tra

- Ma trận kiểm tra
  - Ma trận kiểm tra của một bộ mã có ma trận sinh $G_{k\times n}$ là ma trận H có kích thước $(n-k)\times n$ sao cho

$$G \times H^T = 0$$

- Syndrome – vectơ sửa sai (corrector)
  - $v \times H^T$ được gọi là syndrome hay vectơ sửa sai của v và được kí hiệu là $s(v)$. v là từ mã khi và chỉ khi $s(v) = 0$.
- Ví dụ
  - Tìm ma trận kiểm tra ứng với ma trận sinh sau.

$$G_{4\times 7} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Ma trận kiểm tra (tt)

- H có kích thước 3×7.
- Gọi $h = (a_0, a_1, a_2, a_3, a_4, a_5, a_6)$ là một hàng bất kỳ của H. h trực giao với mọi hàng của G nên chúng ta có hệ bốn phương trình sau

$$G_{4\times 7} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{aligned} a_0 + a_1 + a_3 &= 0 \\ a_0 + a_2 + a_3 + a_4 &= 0 \\ a_1 + a_5 + a_6 &= 0 \\ a_0 + a_2 + a_6 &= 0 \end{aligned}$$

- Vấn đề là tìm được 3 vectơ h độc lập tuyến tính là nghiệm của hệ phương trình trên.
- Chú ý, hệ phương trình trên có thể cho phép chúng ta giải bốn biến theo ba biến còn lại. Chẳng hạn chúng ta giải a_3, a_4, a_5, a_6 theo a_0, a_1, a_2 như sau.

Ma trận kiểm tra (tt)

$$a_3 = a_0 + a_1$$
$$a_4 = a_1 + a_2$$
$$a_5 = a_0 + a_1 + a_2$$
$$a_6 = a_0 + a_2$$

- Cho (a_0, a_1, a_2) lần lượt các giá trị $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ (độc lập tuyến tính với nhau), ta xác định được (a_3, a_4, a_5, a_6) lần lượt như sau $(1, 0, 1, 1)$, $(1, 1, 1, 0)$, $(0, 1, 1, 1)$.
- Vậy H là

$$H_{3\times 7} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

- Chú ý
  - Có thể tồn tại nhiều ma trận kiểm tra khác nhau của cùng một bộ mã và chúng đều có khả năng kiểm tra như nhau.

Ma trận kiểm tra (tt)

- Bổ đề 12.1
  - Nếu ma trận sinh hệ thống của một mã tuyến tính hệ thống có dạng $G_{k\times n} = [I_{kk} \mid P_{k(n-k)}]$

    thì $H_{(n-k)\times n} = [P_{k(n-k)}{}^T \mid I_{(n-k)(n-k)}]$

    là một ma trận kiểm tra của mã.
  - Tương tự nếu ma trận sinh có dạng

    $$G_{k\times n} = [P_{k(n-k)} \mid I_{kk}]$$

    thì ma trận kiểm tra có dạng

    $$H_{(n-k)\times n} = [I_{(n-k)(n-k)} \mid P_{k(n-k)}{}^T]$$

    trong đó $I_{(n-k)(n-k)}$ là ma trận đơn vị kích thước $(n-k)\times(n-k)$, còn $P_{k(n-k)}{}^T$ là ma trận chuyển vị của ma trận $P_{k(n-k)}$.

Chứng minh

$$G_{k\times n} = \left[\begin{array}{cccc|ccc} 1 & 0 & \cdots & 0 & P_{00} & P_{01} & \cdots & P_{0(n-k-1)} \\ 0 & 1 & \cdots & 0 & P_{10} & P_{11} & \cdots & P_{1(n-k-1)} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & P_{(k-1)0} & P_{(k-1)1} & \cdots & P_{(k-1)(n-k-1)} \end{array}\right]$$

$k\times k$ | $k\times(n-k)$

$$H_{(n-k)\times n} = \left[\begin{array}{cccc|ccc} P_{00} & P_{10} & \cdots & P_{(k-1)0} & 1 & 0 & \cdots & 0 \\ P_{01} & P_{11} & \cdots & P_{(k-1)1} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ P_{0(n-k-1)} & P_{1(n-k-1)} & \cdots & P_{(k-1)(n-k-1)} & 0 & 0 & \cdots & 1 \end{array}\right]$$

$(n-k)\times k$ | $(n-k)\times(n-k)$

Chứng minh (tt)

- Ta chứng minh

$$G \times H^T = 0$$

- Chứng minh điều này $\Leftrightarrow$ việc chứng minh

$$g_i \times h_j = 0 \ \forall \ i = 0, \ldots, k-1, j = 0, \ldots, n-k-1$$

trong đó

$g_i = (g_{i0}, \ldots, g_{i(n-1)})$ là hàng i của G còn
$h_j = (h_{j0}, \ldots, h_{j(n-1)})$ là hàng j của ma trận H.

Thật vậy

$$g_i \times h_j = \sum_{s=0}^{n-1} g_{is} h_{js} = \sum_{s=0}^{k-1} g_{is} h_{js} + \sum_{s=k}^{n-k-1} g_{is} h_{js}$$

$$= h_{ji} + g_{i(k+j)} = P_{ij} + P_{ij} = 0$$

Ví dụ

- Tìm ma trận H cho các ma trận sinh sau

$$G_{ht(4\times7)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix} \qquad G_{4\times7} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

$$G_{4\times7} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Khả năng chống nhiễu tương đương

- Hai mã tuyến tính $C(n, k)$ được gọi là có khả năng chống nhiễu tương đương nếu chúng có cùng khoảng cách Hamming.

- Bổ đề 12.2
  - Nếu hoán vị hai cột của một ma trận sinh sẽ tạo ra một bộ mã mới có khả năng chống nhiễu tương đương với bộ mã cũ. Nói cách khác việc hoán vị hai cột của ma trận sinh không làm thay đổi khả năng chống nhiễu.
- Bổ đề 12.3
  - Khoảng cách Hamming của một mã tuyến tính bằng trọng số nhỏ nhất khác 0 của bộ mã.

Bổ đề

- Bổ đề 12.4
  - Gọi H là ma trận kiểm tra của một mã tuyến tính, nếu một từ mã có trọng số d thì tồn tại d cột của H có tổng bằng 0.
- Hệ quả
  - Nếu trong ma trận kiểm tra H của một mã tuyến tính số cột phụ thuộc tuyến tính nhỏ nhất là d thì khoảng cách Hamming của bộ mã đó bằng d.
- Ví dụ 12.5

$$H_{3\times 7} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

$d = 3$ (3, 4, 6)

Cách sửa sai

- Vectơ lỗi
  - Là vectơ biểu diễn các vị trí lỗi giữa từ mã truyền và tổ hợp nhận, mỗi vị trí lỗi được biểu diễn bằng bit 1, còn lại là 0.
  - Nếu từ mã truyền là w, vectơ lỗi là e và vectơ nhận là v thì

$$v = w + e$$

$$e = v + w$$

$$w = e + v$$

- Ví dụ
  - $w = 1011011, e = 0010100 \Rightarrow v = w + e = 1001111.$
  - $w = 0110010, v = 0010011 \Rightarrow e = w + v = 0100001.$
  - $v = 1011001, e = 0010010 \Rightarrow w = v + e = 1001011.$

Tập giải mã - coset

- Cho S là một không gian con các từ mã của không gian V, coset của một phần tử $z \in V$ đối với S được kí hiệu là $z + S$ và được định nghĩa như sau

$$z + S = \{z + w: w \in S\}$$

- Bổ đề 12.5
  - Tập coset $z + S$ có các tính chất sau.

(1) $z \in z + S$.

(2) Nếu $z \in S$ thì $z + S = S$.

(3) Nếu $v \in z + S$ thì $v + S = z + S$.

(4) Nếu $v \notin z + S$ thì $v + S$ và $z + S$ rời nhau.

Sơ đồ giải mã

- Với mỗi vectơ nhận v chúng ta sẽ có một tập coset tương ứng là $v + S$.
- Trong tập này chọn phần tử có trọng số nhỏ nhất, chẳng hạn là z. Phần tử này thường được gọi là coset leader.
- Thông báo từ mã được truyền chính là $w = v + z$.

- Bổ đề 12.6
  - Các phần tử của một tập coset có cùng một syndrome như nhau. Các tập coset khác nhau có các syndrome khác nhau.
  - $e = (a_1, a_2, ..., a_n)$, các cột của H lần lượt bằng $h_1, h_2, ..., h_n$ thì

$$s(e) = e \times H^T = \sum_{i=1}^{n} a_i h_i = \sum_{a_i \neq 0} a_i h_i$$

Sơ đồ giải mã (tt)

- Nghĩa là $s(e)$ bằng tổng những cột ở những vị trí tương ứng với những vị trí bằng 1 của e.
- Nếu vị trí lỗi sai là 3 thì syndrome của vectơ nhận sẽ là cột số 3 của H.

$$G_{4\times 7} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad H_{3\times 7} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

- Tìm vị trí lỗi sai của các vectơ nhận sau đây

$v = 0010011 \Rightarrow s(v) = ? \Rightarrow e = ? \Rightarrow w = ?$

$v = 0101101 \Rightarrow s(v) = ? \Rightarrow e = ? \Rightarrow w = ?$

Mã tuyến tính Hamming

- Mã tuyến tính Hamming là mã có ma trận H có tính chất giá trị của cột h_i bằng i ($i = 1, 2, ...$)

$$H_{3\times 7} = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

- Bổ đề 12.7
  - Các mã tuyến tính Hamming đều có khoảng cách Hamming $d = 3$. Vì vậy có thể phát hiện sai 2 bit và sửa sai 1 bit.

Ma trận sinh của mã tuyến tính Hamming

- Xét mã tuyến tính Hamming $C(7, 4)$ có các bit thông tin nằm ở các vị trí 3, 5, 6, 7. Hãy xác định ma trận sinh G của bộ mã.
- Gọi $w = (a_1, a_2, a_3, a_4, a_5, a_6, a_7)$ là một từ mã. Chúng ta có hệ phương trình sau được dẫn ra từ công thức $w \times H^T = 0$.

$$a_4 + a_5 + a_6 + a_7 = 0$$

$$a_2 + a_3 + a_6 + a_7 = 0$$

$$a_1 + a_3 + a_5 + a_7 = 0$$

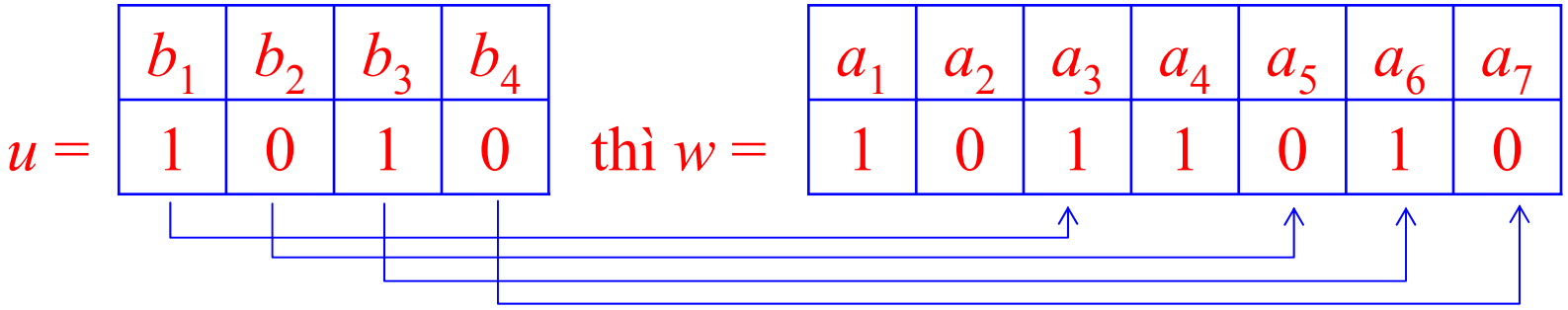
- Từ đây suy ra công thức tính các bit kiểm tra a_1, a_2, a_4 theo các bit thông báo a_3, a_5, a_6, a_7 như sau

$$a_1 = a_3 + a_5 + a_7$$

$$a_2 = a_3 + a_6 + a_7$$

$$a_4 = a_5 + a_6 + a_7$$

Ma trận sinh của mã tuyến tính Hamming



$$G_{4\times 7} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

- Ví dụ
  - Xét mã tuyến tính Hamming $C(7, 4)$ có các bit thông tin nằm ở các vị trí 1, 2, 3, 4. Hãy xác định ma trận sinh G của bộ mã.

Bài 13 Mã vòng

13.1 Giới thiệu

13.2 Các tính chất của mã vòng

13.3 Ma trận sinh và ma trận kiểm tra của mã

13.4 Mã BCH

Giới thiệu

- Định nghĩa
  - Một mã tuyến tính $C(n, k)$ được gọi là mã vòng nếu $w = a_0a_1\ldots a_{n-2}a_{n-1}$ là một từ mã thì $v = a_{n-1}a_0a_1\ldots a_{n-2}$ cũng là một từ mã.
  - Nghĩa là dịch vòng (sang trái hay phải) một từ mã thì kết quả cũng là một từ mã. Ở đây qui ước dịch phải.
- Đa thức mã
  - Nếu $w = a_0a_1\ldots a_{n-2}a_{n-1}$ là một từ mã thì $w(x) = a_0 + a_1x + \ldots + a_{n-2}x^{n-2} + a_{n-1}x^{n-1}$ là đa thức mã tương ứng với từ mã w.
- Ví dụ
  - Bảng sau đây trình bày một mã vòng $C(7, 4)$.

Ví dụ

m	*w*	*w*(x)	*m*	*w*	*w*(x)
0000	0000000	0	0001	0001101	$x^3 + x^4 + x^6$
1000	1101000	$1 + x + x^3$	1001	1100101	$1 + x + x^4 + x^6$
0100	0110100	$x + x^2 + x^4$	0101	0111001	$x + x^2 + x^3 + x^6$
1100	1011100	$1 + x^2 + x^3 + x^4$	1101	1010001	$1 + x^2 + x^6$
0010	0011010	$x^2 + x^3 + x^5$	0011	0010111	$x^2 + x^4 + x^5 + x^6$
1010	1110010	$1 + x + x^2 + x^5$	1011	1111111	$1 + x + x^2 + x^3 + x^4 + x^5 + x^6$
0110	0101110	$x + x^3 + x^4 + x^5$	0111	0100011	$x + x^5 + x^6$
1110	1000110	$1 + x^4 + x^5$	1111	1001011	$1 + x^3 + x^5 + x^6$

Giới thiệu (tt)

- $w^{(i)}$, $w^{(i)}(x)$
  - $w^{(i)}$ là từ mã do dịch từ mã w i bit, và $w^{(i)}(x)$ là đa thức mã tương ứng của $w(i)$. $w^{(0)}$ chính là w.

i	$w^{(i)}$	$w^{(i)}(x)$
0	1101000	$1 + x + x^3$
1	0110100	$x + x^2 + x^4 = x * (1 + x + x^3) = x * w(x)$
2	0011010	$x^2 + x^3 + x^5 = x^2 (1 + x + x^3) = x^2 * w(x)$
3	0001101	$x^3 + x^4 + x^6 = x^3 (1 + x + x^3) = x^3 * w(x)$
4	1000110	$1 + x^4 + x^5 = x^4 + x^5 + x^{7 \bmod 7}$
5	0100011	$x + x^5 + x^6 = x^5 + x^6 + x^{8 \bmod 7}$
6	1010001	$1 + x^2 + x^6 = x^6 + x^{7 \bmod 7} + x^{9 \bmod 7}$

Giới thiệu (tt)

- $w^{(i)}(x) = x^i * w(x)$ tuy nhiên nếu $w^{(i)}(x)$ có x^p với $p \geq n$ thì x^p được thay bằng $x^{p \bmod n}$.
- Mặc khác trên trường $GF(2)$ chúng ta có

$$x^{n+j} = x^j * (x^n + 1) + x^j \text{ hay } x^{n+j} \bmod (x^n + 1) = x^j$$

- Bổ đề 13.1

$$w^{(i)}(x) = [x^i * w(x)] \bmod (x^n + 1)$$

Các tính chất của mã vòng

- Định lý 13.1
  - Đa thức mã khác 0 có bậc nhỏ nhất là duy nhất. Hay nói cách khác không tồn tại hai đa thức mã khác 0, khác nhau và cùng có bậc nhỏ nhất.
- Chứng minh
  - Giả sử ∃ hai đa thức mã khác nhau, cùng có bậc nhỏ nhất là r, $0 < r < n$.

$$g(x) = g_0 + g_1 x + \ldots + g_{r-1} x^{r-1} + x^r$$

$$f(x) = f_0 + f_1 x + \ldots + f_{r-1} x^{r-1} + x^r$$

  - Từ đây suy ra đa thức mã $g(x) + f(x)$ có bậc nhỏ hơn r, mâu thuẫn. Chứng minh hoàn tất.

Các tính chất của mã vòng (tt)

- Kí hiệu đa thức mã có bậc nhỏ nhất là g(x)

$$g(x) = g_0 + g_1 x + \ldots + g_{r-1} x^{r-1} + x^r$$

- Định lý 13.2
  - Hệ số tự do g_0 của g(x) phải bằng 1.
- Chứng minh
  - Giả sử $g_0 = 0$. Suy ra

$$g(x) = x * (g_1 + \ldots + g_{r-1} x^{r-2} + x^{r-1})$$

  - Đặt $f(x) = (g_1 + \ldots + g_{r-1} x^{r-2} + x^{r-1})$, suy ra f(x) cũng là một đa thức mã. Vì f(x) tương ứng với từ mã được dịch trái 1 bit hay dịch phải $(n-1)$ bit từ từ mã ứng với g(x).
  - Mà bậc của f(x) bằng $r - 1 < r$ mâu thuẫn với định nghĩa của g(x).

Các tính chất của mã vòng (tt)

- Định lý 13.3
  - Một đa thức $v(x)$ trên trường $GF(2)$ có bậc $\leq n-1$ là đa thức mã nếu và chỉ nếu nó là một bội số của $g(x)$. Tức là nó có thể viết $v(x) = q(x) * g(x)$.
- Chứng minh
  - **Chiều thuận**
  - Nếu $v(x) = q(x) * g(x)$ và có bậc $\leq n-1$ thì $v(x)$ là đa thức mã.

$$v(x) = q(x) * g(x) = \left(\sum_{i=0}^{p} q_i x^i \right) * g(x) = \sum_{i=0}^{p} q_i \left(x^i * g(x) \right)$$

với p là bậc của $q(x)$ và $p + r \leq n-1$. Do $x^i * g(x)$ với $0 \leq i \leq p$ là đa thức mã, nên $v(x)$ là đa thức mã vì nó là một tổ hợp tuyến tính của các đa thức mã.

Các tính chất của mã vòng (tt)

- **Chiều ngược**
- Nếu $v(x)$ là đa thức mã thì chia $v(x)$ cho $g(x)$

$$v(x) = q(x) * g(x) + r(x)$$

trong đó $r(x)$ là đa thức dư và có bậc nhỏ hơn bậc của $g(x)$.

Đối với các đa thức trên trường $GF(2)$ chúng ta có thể suy ra

$$r(x) = q(x) * g(x) + v(x)$$

Nên $r(x)$ là một đa thức mã. Theo định nghĩa của $g(x)$ suy ra $r(x) = 0$. Chứng minh hoàn tất.

- Từ định lý này chúng ta gọi $g(x)$ là **đa thức sinh**, vì từ $g(x)$ có thể sinh ra tất cả các đa thức mã khác.

Các tính chất của mã vòng (tt)

- Định lý 13.4
  - Đa thức sinh của một mã vòng $C(n, k)$ có bậc $r = n - k$.
- Chứng minh
  - Mỗi đa thức mã $w(x)$ là một bội số của $g(x)$

$$w(x) = q(x) * g(x)$$

  - Có 2^k từ mã nên có 2^k đa thức $q(x)$. Suy ra bậc của $q(x)$ là $\leq k - 1$. Suy ra bậc của g(x) là $n - k$.
  - Từ định lý này đa thức sinh $g(x)$ có thể được biểu diễn như sau

$$g(x) = g_0 + g_1 x + \ldots + g_{n-k} x^{n-k}$$

trong đó $g_0 = g_{n-k} = 1$.

Các tính chất của mã vòng (tt)

- Định lý 13.5
  - Đa thức sinh của mã vòng $C(n, k)$ là một ước số của $x^n + 1$.
- Chứng minh
  - Bổ đề 13.1 suy ra

$$g^{(i)}(x) = [x^i * g(x)] \bmod (x^n + 1)$$

$$\Leftrightarrow x^i * g(x) = q(x) * (x^n + 1) + g^{(i)}(x)$$

Chọn $i = k \Rightarrow q(x) = 1$ tức

$$x^k * g(x) = (x^n + 1) + g^{(i)}(x)$$

$$\Rightarrow x^n + 1 = x^k * g(x) + g^{(i)}(x)$$

Do $g^{(i)}(x)$ là một đa thức mã nên $g^{(i)}(x)$ là một bội của $g(x)$, $\Rightarrow$ $x^n + 1$ là một bội của $g(x)$. Chứng minh hoàn tất.

Các tính chất của mã vòng (tt)

- Định lý 13.6
  - Nếu $g(x)$ là một đa thức có bậc $(n-k)$ và là ước số của $(x^n + 1)$ thì $g(x)$ sinh ra mã vòng $C(n, k)$, hay nói cách khác $g(x)$ là đa thức sinh của một mã vòng $C(n, k)$ nào đó.
- Chứng minh
  - Xét k đa thức $g(x)$, $x * g(x)$, …, $x^{k-1} * g(x)$.

    Các đa thức này đều có bậc $\leq n - 1$.

    Gọi $v(x)$ là một tổ hợp tuyến tính của k đa thức này với các hệ số $a_i \in GF(2)$.

    $$v(x) = a_0 g(x) + a_1 x * g(x) + \ldots + a_{k-1} x^{k-1} * g(x)$$

    $v(x)$ là một đa thức có bậc $\leq n - 1$ và là bội số của $g(x)$.

Các tính chất của mã vòng (tt)

- Có tất cả 2^k tổ hợp tuyến tính $v(x)$ khác nhau và tạo nên một không gian tuyến tính của các đa thức mã với $g(x)$, $x * g(x)$, …, $x^{k-1} * g(x)$ là các đa thức làm cơ sở.

  Chúng ta chứng minh rằng bộ mã tương ứng với không gian này là mã vòng.

  Gọi

$$w(x) = b_0 + b_1 x + \ldots + b_{n-1} x^{n-1}$$

  là một đa thức của không gian.

  Chúng ta chứng minh

$$w^{(1)}(x) = b_{n-1} + b_0 x + b_1 x^2 + \ldots + b_{n-2} x^{n-1}$$

  cũng là một đa thức của không gian.

Các tính chất của mã vòng (tt)

- Theo Bổ đề 13.1 chúng ta có

$$w^{(1)}(x) = [x * w(x)] \bmod (x^n + 1)$$

Dựa vào biểu diễn của $v(x)$ và $w^{(1)}(x)$ chúng ta suy ra

$$x * w(x) = b_{n-1}(x^n + 1) + w^{(1)}(x)$$

Do $v(x)$ và $(x^n + 1)$ đều là bội của $g(x)$ nên $w^{(1)}(x)$ cũng là bội của $g(x)$. Suy ra $w^{(1)}(x)$ cũng là đa thức mã. Hoàn tất chứng minh.

Ma trận sinh

$$G_{k\times n} = \begin{bmatrix} \overbrace{g_0 & g_1 & g_2 & \cdots & g_{n-k}}^{n-k+1} & \overbrace{0 & 0 & \cdots & 0}^{k-1} \\ 0 & g_0 & g_1 & \cdots & g_{n-k-1} & g_{n-k} & 0 & \cdots & 0 \\ 0 & 0 & g_0 & \cdots & g_{n-k-2} & g_{n-k-1} & g_{n-k} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & g_0 & g_1 & g_2 & \cdots & g_{n-k} \end{bmatrix}$$

- Ví dụ
  - Tìm một mã vòng $C(7, 4)$.
  - Theo các tính chất của mã vòng suy ra đa thức sinh của mã có bậc bằng 3 và là một ước số của $x^7 + 1$. Phân tích đa thức này ra thừa số chúng ta được

Ví dụ

$$x^7 + 1 = (1 + x)(1 + x + x^3)(1 + x^2 + x^3)$$

- Chọn chẳng hạn

$$g(x) = (1 + x + x^3)$$

$$G_{4\times 7} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

Mã vòng dạng hệ thống

- Từ dạng hệ thống loại 1 chúng ta có thể dịch vòng k bit để biến đổi sang dạng hệ thống loại 2 và ngược lại.

$$G_{4\times 7} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix} \Rightarrow G_{ht(4\times 7)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

- Mã hóa thành từ mã hệ thống
  - $u(x)$ là thông báo, $w(x)$ là từ mã hệ thống loại 2 tương ứng.

$$x^{n-k} * u(x) = q(x) * g(x) + a(x)$$

$$w(x) = x^{n-k} * u(x) + a(x) = q(x) * g(x)$$

Ví dụ

- Cho mã vòng $C(7, 4)$ có ma trận sinh là $g(x) = (1 + x + x^3)$. Hãy mã hoá thông báo $u = 1010$ thành từ mã hệ thống dạng 2.

$$u(x) = 1 + x^2.$$

Nhân $u(x)$ với $x^{n-k} = x^3$ rồi chia cho $g(x)$ chúng ta được.

$$x^3 * (1 + x^2) = x^3 + x^5 = x^2 * (1 + x + x^3) + x^2$$

- Từ đây suy ra

$$w(x) = x^2 + x^3 + x^5$$

$$w = 0011\mathbf{1010}$$

là từ mã hệ thống dạng 2 tương ứng với u.

Ma trận kiểm tra của mã vòng

- Có một cách khác để tìm ma trận kiểm tra của mã vòng

$$x^n + 1 = g(x) * h(x)$$

- $h(x)$ được gọi là đa thức đối ngẫu của $g(x)$. $h(x)$ có bậc k

$$h(x) = h_0 + h_1 x + \ldots + h_k x^k$$

- Ma trận sau là một ma trận kiểm tra của mã vòng

$$H_{(n-k)\times n} = \begin{bmatrix} \overbrace{h_k & h_{k-1} & h_{k-2} & \cdots & h_0}^{k+1} & \overbrace{0 & 0 & \cdots & 0}^{n-k-1} \\ 0 & h_k & h_{k-1} & \cdots & h_1 & h_0 & 0 & \cdots & 0 \\ 0 & 0 & h_k & \cdots & h_2 & h_1 & h_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & h_k & h_{k-1} & h_{k-2} & \cdots & h_0 \end{bmatrix}$$

Ví dụ

- Cho mã vòng $C(7, 4)$ có ma trận sinh là $g(x) = (1 + x + x^3)$. Từ đây suy ra

$$h(x) = (1 + x + x^2 + x^4)$$

- Ma trận kiểm tra của bộ mã là

$$H_{3\times 7} = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

Ứng dụng trường $GF(2^m)$ để xây dựng mã vòng

- Định lý 13.7
  - Cho a là một phần tử khác 0 của trường $GF(2^m)$ có chu kỳ là n, đa thức tối thiểu $f(x)$ của a có bậc là m. Thì mã có ma trận sau làm ma trận kiểm tra là một mã vòng $C(n, n-m)$, trong đó mỗi phần tử trong ma trận bên dưới được thay thế bằng vectơ m thành phần tương ứng của nó.

$$H_{m\times n} = [1\ a\ a^2\ \ldots\ a^{n-2}\ a^{n-1}]$$

  Hơn nữa mã vòng này có đa thức sinh chính là $f(x)$.

- Ví dụ
  - Xét trường $GF(2^4)$ và a có đa thức tối thiểu là

$$f(x) = 1 + x + x^4$$

Ứng dụng trường $GF(2^m)$ để xây dựng mã vòng (tt)

- Từ đây suy ra ma trận kiểm tra của mã vòng (15, 11).

$$H_{4\times 15} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

- Nếu đa thức tối thiểu của a là $f(x) = 1 + x + x^2 + x^3 + x^4$ thì a có chu kỳ là 5 và các phần tử $1, a, ..., a^4$ được biểu diễn như sau.

$1 = (1000)$ $\quad$ $a^3 = (0001)$

$a = (0100)$ $\quad$ $a^4 = (1111)$

$a^2 = (0010)$

Ứng dụng trường $GF(2^m)$ để xây dựng mã vòng (tt)

- Từ đây suy ra ma trận kiểm tra của mã vòng (5, 1)

$$H_{4\times5} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Mã BCH nhị phân

- Do Bose, Chaudhuri và Hocquenghem sáng lập ra.
- Là mã vòng có khả năng sửa được nhiều lỗi.
- Đối với các số nguyên dương m và t bất kỳ chúng ta sẽ xây dựng một mã BCH nhị phân có các thông số sau:

  Độ dài từ mã: $n = 2^m - 1$

  Số bit kiểm tra: $n - k \leq mt$

  Khoảng cách Hamming: $d_{\min} \geq 2t + 1$

Định lý

- Định lý 13.8
  - Cho a là một phần tử của trường $GF(2^m)$ có đa thức tối thiểu là một đa thức căn bản bậc m. Thì mã có ma trận sau làm ma trận kiểm tra là một mã vòng có khoảng cách Hamming $\geq 2t + 1$, trong đó mỗi phần tử trong ma trận bên dưới được thay thế bằng vectơ m thành phần tương ứng của nó.

$$H = \begin{bmatrix} 1 & a & a^2 & \cdots & a^{n-2} & a^{n-1} \\ 1 & a^3 & a^6 & \cdots & a^{3(n-2)} & a^{3(n-1)} \\ 1 & a^5 & a^{10} & \cdots & a^{5(n-2)} & a^{5(n-1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a^{2t-1} & a^{2(2t-1)} & \cdots & a^{(2t-1)((n-2)} & a^{(2t-1)((n-1)} \end{bmatrix}$$

Định lý (tt)

- Hơn nữa đa thức sinh $g(x)$ của bộ mã là đa thức bội số chung nhỏ nhất của các đa thức tối thiểu của các phần tử a, a^3, a^5, …, a^{2t-1}.

- Bổ đề 13.2
  - Ma trận A sau có định thức bằng $\prod_{i>j}(y_i - y_j)$ với $i, j \in \{1, 2, \ldots, r\}$. Định thức này được gọi là định thức Vandermonde.

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ y_1 & y_2 & \cdots & y_r \\ y_1^2 & y_2^2 & \cdots & y_r^2 \\ \vdots & \vdots & \vdots & \vdots \\ y_1^{r-1} & y_2^{r-1} & \cdots & y_r^{r-1} \end{bmatrix}$$

Ví dụ

- Cho $m = 4$, $t = 2$ chúng ta sẽ xây dựng một mã vòng có chiều dài từ mã là $2^4 - 1 = 15$ và có khoảng cách Hamming $d \geq 5$. Việc xây dựng sẽ dựa vào trường GF(2^4).
- Gọi a là phần tử có đa thức tối thiểu là đa thức căn bản bậc 4 sau $$f_1(x) = 1 + x + x^4$$
- Đây chính là trường $GF(2^4)$ trong ví dụ ở slide 250.
- a có chu kỳ $n = 2^m - 1 = 15$. Chúng ta có ma trận kiểm tra của bộ mã như sau.

$$H=\begin{bmatrix} 1 & a & a^2 & a^3 & a^4 & a^5 & a^6 & a^7 & a^8 & a^9 & a^{10} & a^{11} & a^{12} & a^{12} & a^{14} \\ 1 & a^3 & a^6 & a^9 & a^{12} & a^{15} & a^{18} & a^{21} & a^{24} & a^{27} & a^{30} & a^{33} & a^{36} & a^{39} & a^{42} \end{bmatrix}$$

- Thay mỗi phần tử a^i bằng vectơ 4 thành phần tương ứng

Ví dụ (tt)

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Ví dụ (tt)

- Đa thức sinh g(x) là bội số của hai đa thức tối thiểu tương ứng với phần tử a và a^3.
- Theo ví dụ ở slide 250, đa thức tối thiểu của a^3 là

$$f_3(x) = 1 + x + x^2 + x^3 + x^4.$$

- Từ đây suy ra

$$\begin{aligned} g(x) &= f_1(x) * f_3(x) \\ &= (1 + x + x^4) * (1 + x + x^2 + x^3 + x^4) \\ &= 1 + x^4 + x^6 + x^7 + x^8 \end{aligned}$$

- Chú ý
  - Trong trường hợp đa thức tối thiểu của a không phải là đa thức căn bản, chúng ta sẽ tìm được mã vòng có chiều dài $n \neq 2^m + 1$, với n là chu kỳ của a.