

Giới Thiệu Về Toán Mô Hình

Cần Trần Thành Trung¹, Lê Việt Hải²

¹*Duke University, USA*

²*Pennsylvania State University, USA*

.....
Toán mô hình giống với một công cụ hơn là một môn học, nắm được những ý tưởng cơ bản của công cụ này các em có thể sử dụng nó vào những lĩnh vực khác như: vật lý, hóa học, sinh học, hay các vấn đề và hiện tượng xã hội.
.....

Subject Index Math modeling, graph theory, logistic model.

1. Giới thiệu

If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is

John von Neumann

As far as the laws of mathematics refer to reality, they are not certain, and as far as they are certain, they do not refer to reality

Albert Einstein

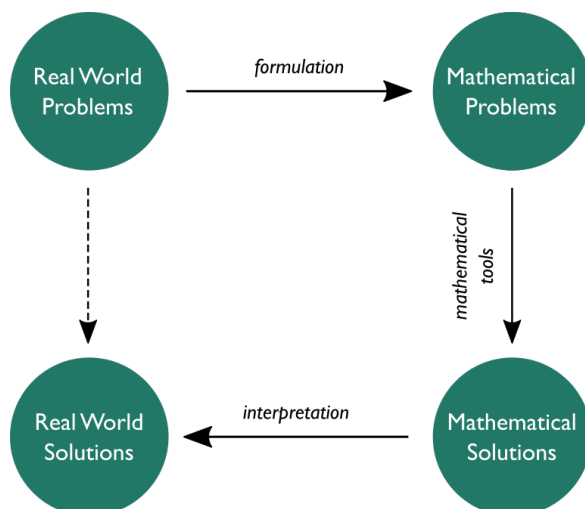
Toán là một ngôn ngữ gọn gàng với những quy tắc và phép tính được định nghĩa rõ ràng, chặt chẽ, có thể được sử dụng để mô tả các vấn đề thực tế một cách cụ thể, dễ hình dung hơn. Toán lý thuyết đã được nghiên cứu và xây dựng cả ngàn năm nay nên cực kỳ “giàu” và “đẹp,” cung những công cụ hữu ích để giải quyết các bài toán. Với sự hỗ trợ của máy tính, cái phép tính toán phức tạp hay một khối lượng lớn các phép tính có thể được xử lý dễ dàng.

Nhờ áp dụng công cụ Toán học vào giải quyết các vấn đề kinh tế, xã hội, nhiều nhà khoa học đã đạt những thành tựu đáng kể. Tiêu biểu là hai giải Nobel dựa trên hai bài viết, “*Equilibrium points in n-person games*” của John Nash năm 1994, và “*Stable allocations and the practice of market design*” của Lloyd Shapley năm 2012.

Qua bài viết này, chúng ta sẽ tìm hiểu khái niệm, đặc điểm, tính chất của Toán mô hình và những ví dụ cụ thể trong thực tế.

2. Khái niệm

Hiểu một cách tổng quan, Toán Mô Hình là sử dụng các biến, công thức, mệnh đề, công cụ toán học để mô tả một “hệ” và tìm câu trả lời cho những câu hỏi, bài toán liên quan đến “hệ” đó.



Hình 1: “Hệ” bao gồm các đối tượng trong thực tế và mối liên quan giữa các đối tượng này.

Các ví dụ về đối tượng (và các mối liên hệ) để mô hình có thể kể đến là

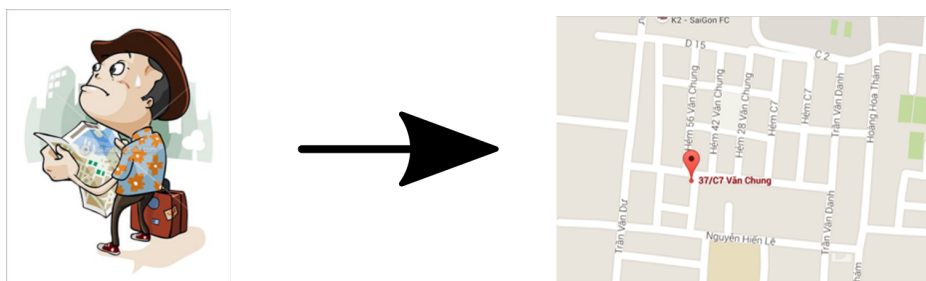
Ví dụ 2.1.

- Người dân trong thành phố (tính tuổi thọ trung bình),
- Các học sinh và mentors của PiMA (sắp xếp lịch phỏng vấn phù hợp),
- Các tài khoản trên Facebook (tìm ra xu hướng sử dụng của giới trẻ),
- Các bãi biển ở Việt Nam (xây dựng tiêu chí để xếp hạng).

Có một câu hỏi được đặt ra: *Tại sao không giải quyết thẳng ngay trên “hệ” đó mà cần phải thông qua mô hình toán của nó?*

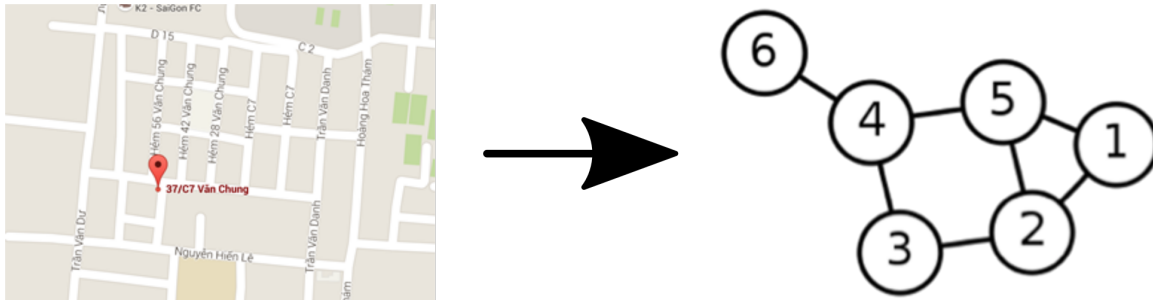
Ví dụ 2.2. Tính tuổi thọ trung bình chỉ cần thống kê được số người và tuổi từng người. Mỗi người được đại diện bằng một biến x_i (tuổi). Như vậy tuổi trung bình là $\bar{x} = \frac{\sum x_i}{n}$, với n là tổng số người.

Ví dụ 2.3. Tìm đường đi bằng Google Map. Nếu 2 người khác nhau ở cùng một vị trí thì sẽ được xem là cùng một điểm trên Google Maps. Các đường đi sẽ được biểu diễn thành các đường trắng.



Hình 2: Mô hình hóa việc tìm đường.

Nhưng đó vẫn chưa phải là mô hình toán cuối cùng. Đơn giản hơn nữa, ngôn ngữ mà cả máy tính và con người đều có thể hiểu và làm việc được: mô hình đồ thị. Như vậy, mỗi địa điểm sẽ được biểu diễn thành một đỉnh của đồ thị, các đường đi sẽ thành các cạnh của đồ thị.



Hình 3: Tối giản hóa bằng ngôn ngữ đồ thị.

Chúng ta sẽ được tìm hiểu về đồ thị và cả thuật toán liên quan Google sử dụng để tìm đường đi trên Google Maps.

Câu hỏi: Chúng ta đã từng mô hình hóa người bằng 1 điểm chưa? Tại sao không giải quyết thẳng ngay trên “hệ” đó mà cần phải thông qua mô hình toán của nó?

Các “hệ” trong thực tế rất phức tạp và có những yếu tố không liên quan hoặc không quan trọng đến những câu hỏi cần trả lời. Mô hình toán làm cho các đối tượng trong “hệ” trở nên đơn giản hơn, nhưng vẫn đủ thông tin và đủ phức tạp để giải quyết các câu hỏi không tầm thường. Ví dụ là ở bài toán tìm đường ở trên khi chuyển qua ngôn ngữ đồ thị ta không quan tâm đến hình dạng của đường mà chỉ gán mỗi đường bởi một đường thẳng có giá trị riêng thể hiện độ dài.

Ví dụ 2.4. (Google Page Rank) Thuật toán xếp hạng các trang web của Google để tìm kiếm nhanh.

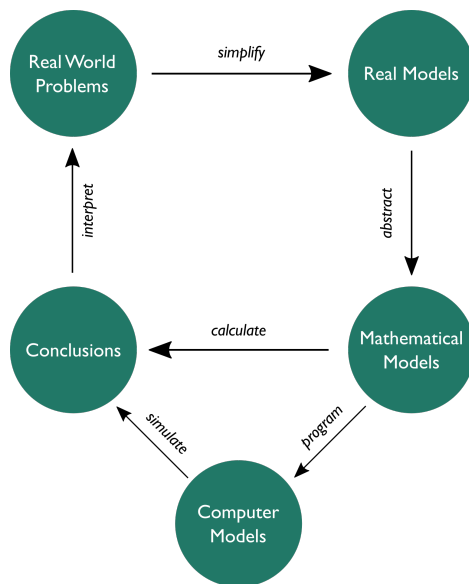
Ví dụ 2.5. Lên kế hoạch vận chuyển sữa hộp từ các kho đến các đại lý một cách tối ưu nhất.

Câu hỏi: “Hệ” ở đây là gì? Vấn đề cần giải quyết? Yếu tố quan trọng của đối tượng mình có thể sử dụng được để trả lời câu hỏi?

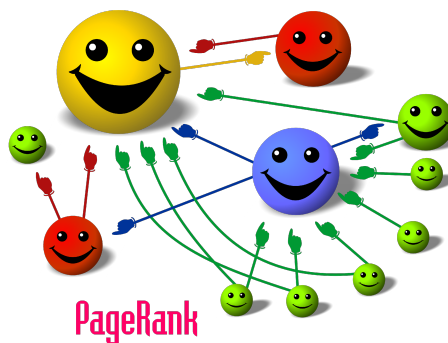
Fun Fact: “PageRank” là kết hợp giữa “Page” (người sáng lập Google, không phải trang giấy) + “Rank” (xếp hạng).

Ví dụ của Google PageRank

- o Thuật toán đầu tiên của Google: dựa trên đường links giữa các trang web này để tính xem trang web nào là quan trọng nhất.
- o Có hai giả thiết quan trọng:



Hình 4: Sơ đồ tóm tắt Toán mô hình.



Hình 5: Google PageRank.

- 1) Trang web sẽ càng quan trọng nếu có càng nhiều trang web dẫn đến nó và các trang web ấy cũng được nhiều trang web dẫn đến.
- 2) Một người đang ở một web bất kì sẽ phải ngẫu nhiên đến một web khác có liên kết từ web này với xác suất như nhau.
 - o Biểu diễn mối liên hệ giữa các web bằng đồ thị có hướng (sẽ được đề cập sau).
 - o Giải quyết bài toán này dựa trên xác suất: mô hình Markov Chain (sẽ được đề cập sau)

Liên hệ: Sử dụng mô hình trên cho vấn đề gì khác được nữa không? Người ta còn thường sử dụng mô hình trên để xếp hạng độ quan trọng của các điểm trong một mạng lưới. Có những mô hình, bài toán đã được xây dựng sẵn và chúng ta có thể áp dụng nó cho những trường hợp tương tự mà mình gặp phải.

Qua các ví dụ trên, ta đưa ra định nghĩa sau

Định nghĩa 2.6. Một mô hình toán học là một bộ ba (S, Q, M) trong đó S là một "hệ", Q là một/các câu hỏi liên quan đến S , và M là tập hợp các đối tượng và mệnh đề dùng để mô tả S và trả lời Q .

Hoạt động: Hãy xác định S, Q, M trong các ví dụ trên.

3. Các ví dụ cụ thể

Ví dụ 3.1. Thiết kế lon thiếc đựng thực phẩm và Bất Đẳng Thức



Hình 6: Các lon thiếc.

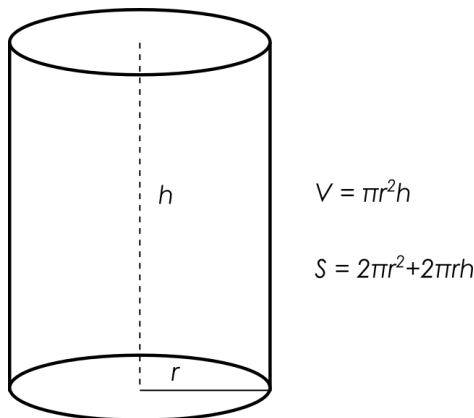
Các yếu tố quan trọng cần lưu ý

- Thể tích của lon: sẽ cố định, cho người tiêu dùng tham khảo và định giá.
- Hình dáng: hình trụ do vẫn đứng chắc đợc và không có góc nhọn có thể gây thương tích. Cũng là dạng thiết kế phổ biến nhất hiện nay.
- Do lon mỏng nên tỉ lệ với diện tích bề mặt.

Như vậy vấn đề còn lại là thiết kế chiều cao và đáy như thế nào để diện tích bề mặt ít nhất nhằm tiết kiệm nguyên vật liệu.

Ta đưa ra một phương án như sau. Gọi h, r lần lượt là chiều cao và bán kính đáy, V là thể tích không đổi:

- Tìm mối liên hệ giữa các biến trên?
- Hàm cần tìm tối ưu là gì?



Hình 7: Tham số hóa lon thiếc.

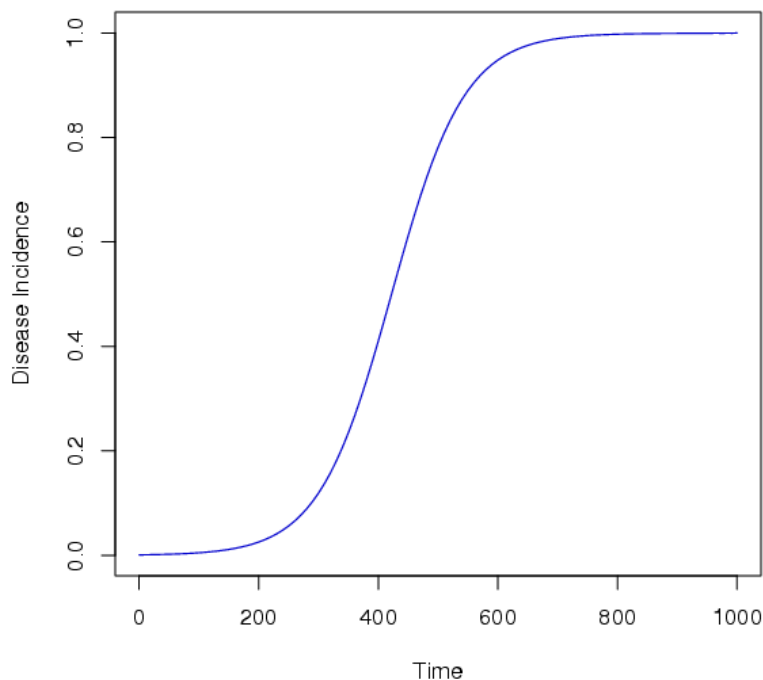
Nhận xét:

- o Hàm số và bất đẳng thức là một công cụ hiệu quả trong các bài toán tối ưu hóa. Chúng ta có một cách nhìn khác ở môn bất đẳng thức theo khía cạnh hình học này.
- o Các bước cơ bản là:
 - (1) Xác định các biên.
 - (2) Xác định hàm cần tối ưu.
 - (3) Xây dựng các điều kiện ràng buộc, mối liên quan giữa các biến để tìm giá trị tối ưu và dấu bằng.

Câu hỏi: Hãy liên hệ ý tưởng trên với bài toán số 4 trong đề tuyển sinh của PIMA.

Ngoài ra công cụ hàm số còn hiệu quả trong những bài toán cần tính toán hay đo đạc sự thay đổi, như ví dụ sau.

Ví dụ 3.2. *Tính toán số lượng cá thể trong một quần thể theo thời gian và đạo hàm (Logistics Model)*



Hình 8: Đồ thị tiêu biểu của một hàm logistic.

Giả thiết P là số lượng cá thể phụ thuộc vào thời gian: $P = P(t)$ trở thành một hàm liên tục ở thời điểm t . *Tỉ lệ gia tăng dân số* ở thời điểm t là $r(t)$. Đạo hàm $P'(t)$ đo tốc độ tăng của hàm số. Như vậy, theo các định nghĩa trên, ta có

$$\frac{P'(t)}{P(t)} = r(t) \quad \text{hay} \quad P'(t) = r(t)P(t).$$

Trong điều kiện lý tưởng, tài nguyên khoáng sản là không giới hạn thì $r(t) = r_0$ là một hằng số, nhờ vào lý thuyết phương trình sai phân, giải phương trình trên ta được

$$P(t) = P_0 e^{r_0 t} \text{ với } P_0 \text{ là số lượng cá thể ban đầu lúc } t = 0.$$

Có thể thấy là dân số sẽ gia tăng rất nhanh theo hàm mũ, khá hợp lý trong trường hợp có nguồn tài nguyên vô hạn. Tất nhiên trong thực tế, tài nguyên chỉ là hữu hạn, do đó tỉ lệ tăng dân số sẽ tỉ lệ nghịch với mật độ (các cá thể sẽ tranh giành với nhau, có thể dẫn đến là đào thải lẫn nhau)

Mật độ được tính bằng công thức $\frac{P(t)}{K}$ trong đó K là số lượng cá thể tối đa để tài nguyên có thể đáp ứng được. Người ta đã kiểm tra được giả thiết này.

Như vậy $r(t)$ tỉ lệ thuận với $(1 - P(t)/K)$ và gần bằng r_0 khi $(1 - P(t)/K)$ gần bằng 1. Do đó, $r(t) = r_0 \left(1 - \frac{P(t)}{K}\right)$, trong đó r_0 là tỉ lệ gia tăng dân số trong điều kiện lý tưởng. Như vậy,

$$P'(t) = r_0 P(t) \left(1 - \frac{P(t)}{K}\right).$$

Phương trình này được gọi là phương trình logistic, hay còn có tên là phương trình Verhulst, một nhà toán học người Bỉ. Năm 1840, ông dùng phương pháp này cũng số liệu những năm gần đó để dự đoán dân số của Mĩ vào năm 1940 và chỉ sai số 1%.

Nhận xét:

- o Cách dùng đạo hàm và thêm các giả thiết lý tưởng hóa để xây dựng mối quan hệ rất phổ biến trong vật lý.
- o Để giải được các phương trình vi phân cần đến kiến thức toán cao cấp.

Câu hỏi:

- o Hãy thử áp dụng mô hình trên để giải bài toán số 1 đề tuyển sinh PiMA.

4. Các thành phần quan trọng

4.1. Phát biểu của bài toán

Mô tả ngắn gọn về S và những kết quả (output) mà mô hình sẽ cho ra được để trả lời cho Q .

Ví dụ 4.1. Một cách phát biểu của các bài toán ở trên:

- o Tìm ra đường đi ngắn nhất từ hai vị trí trên bản đồ và phương tiện đi lại, cũng như ước tính thời gian cần để di chuyển.
- o Xếp hạng các trang web dựa trên liên kết giữa chúng.
- o Tính ra số lượng dân số trong tương lai dựa trên số dân trong một số năm cũ.

Có thể kết hợp thêm mô tả ngắn gọn về input hay các mối liên hệ quan trọng khác.

4.2. Các giả thiết, giả sử

Những điều kiện/giả thiết không đúng hoàn toàn với thực tế nhưng có lý để làm tiền đề cho việc xây dựng mô hình toán.

Chúng thường được dùng khi cần bỏ qua những chi tiết không quá quan trọng, những yếu tố khó kiểm soát được hoặc không có dữ liệu để kiểm chứng.

Ngoài ra, những giả thiết ấy còn đơn giản hóa được các vấn đề nhưng vẫn đủ phức tạp và không quá xa rời thực tế.

Lưu ý: giả thiết/giả sử không thể chứng minh được.

Ví dụ 4.2.

- Google Maps giả sử đường xá luôn ở trạng thái lưu thông bình thường, tốc độ trung bình của mọi người là như nhau.
- Xác định các giả thiết, giả sử trong một ví dụ ở trên.

4.3. *Mô hình toán, bao gồm các biến (input) và công thức liên quan thu được sau khi đã có các giả thiết/giả sử*

Ví dụ 4.3. Xác định mô hình toán trong các ví dụ trên.

4.4. *Giải quyết mô hình toán để tìm nghiệm/ đáp số (output)*

Tùy theo mô hình để giải quyết ra đúng đáp số hoặc đôi khi là gần đúng đáp số.

4.5. *Nghiên cứu nghiệm/đáp số (thường là so sánh với thực tiễn) và đánh giá mô hình để rút ra kết luận hoặc cải thiện mô hình*

Ví dụ 4.4. Các kinh nghiệm sau được suy ra từ kết quả của các mô hình trên:

- Thuật toán PageRank được cải thiện sau một thời gian, sau khi đã có thêm dữ liệu về số người truy cập và các thông tin khác trong thực tế. Nhưng đó vẫn chưa là thuật toán cuối cùng để ra kết quả khi search google do vẫn còn những yếu tố khác như lịch sử tìm kiếm (Search History) của mỗi người.
- Mô hình tính toán dân số của Verhulst sai lệch đi nhiều sau 1940, nguyên nhân có thể là do chiến tranh thế giới lần thứ 2 và mô hình giả sử dân số là một hàm liên tục ổn định theo thời gian. Như vậy, ta cần cải thiện để mô hình điều chỉnh phù hợp với những biến cố/sự kiện.

Đây là một phần vô cùng quan trọng và rất có ích trong việc hiểu thêm về mô hình, về đối tượng. Phần này còn có thể cho thêm ý tưởng để cải thiện mô hình cũ hay xây dựng các mô hình mới khác, hay thậm chí là hiểu rõ hơn về vấn đề trong thực tế, vì sao nó không đúng như mô hình dự đoán.

5. Hai tính chất quan trọng của M

5.1. Input/ Output

Hiểu nôm na, mọi mô hình toán đều hoạt động theo nguyên lý nhập/xuất: nhận vào những tham số, thông tin (input) và cho ra đáp số, câu trả lời (output).

Ví dụ 5.1 Xác định input/output trong các ví dụ trên.

input sẽ là các tham số, các thông tin, mệnh đề độc lập với nhau còn output sẽ phụ thuộc vào input.

5.2. *Open-ended* (tạm hiểu là có thể xây dựng nhiều M cho cùng một S và Q):

Ý nghĩa quan trọng nhất của toán mô hình không dừng lại ở việc tìm ra được đáp số hay giải quyết một vấn đề mà là qua đó giúp chúng ta hiểu thêm về các đối tượng và mối liên hệ giữa chúng trong “hệ”.

Trong nhiều vấn đề thực tế, không có một lời giải, mô hình duy nhất cho một bài toán mà tùy thuộc vào cách tiếp cận, cách thiết lập các giả thiết hay giả định, và cách định nghĩa những yếu tố “tốt nhất” hay “hiệu quả nhất”.

Ví dụ 5.2. Chi phí vs thời gian: Trang tìm vé máy bay nổi tiếng *Skiplagged* sẽ tìm ra đường bay “tốt” nhất theo nghĩa là rẻ nhất, nhưng cũng có thể đường bay “tốt” nhất là đường bay tiết kiệm được nhiều thời gian nhất trong một khoảng thời gian nhất định.

Ta có nhận xét:

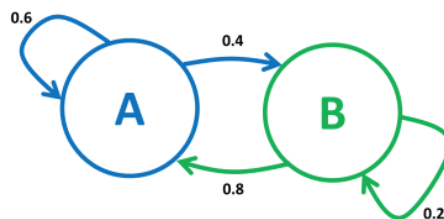
- o Rất khó để tìm ra một giải pháp hay lời giải hoàn hảo về mọi mặt, mỗi cách giải hay mô hình có những được và mất riêng (trade-off), ta cần chọn ra những yếu tố mình muốn tối ưu hóa, ghi lại những yếu tố mà mô hình chưa đánh giá và tìm cách cải thiện nó.
- o Đôi khi những mô hình toán tạo nên các bài toán mở và chưa có những công cụ trong lý thuyết để mô tả và giải dễ dàng.
- o Nếu không giải quyết được toàn bộ bài toán, hãy giải quyết một phần của nó và lưu ý những phần còn lại.

5.3. *Interdisciplinary* (tạm hiểu là có thể dùng một M cho nhiều S)

Một mô hình toán M có thể được sử dụng nhiều vấn đề khác nhau (S , Q thay đổi) nhưng gần tương tự về bản chất hay mô tả.

Ví dụ 5.4. Tìm vé máy bay giá rẻ *Skiplagged*. Dùng lại thuật toán tìm đường đi ngắn nhất của Google Maps, thay độ dài của quãng đường bằng số tiền. *Fun fact:* *Skiplagged* đã từng bị kiện do làm ảnh hưởng đến doanh thu của các hãng máy bay.

Ví dụ 5.5. Mô hình Pagerank của Google có bản chất là mô hình chuỗi Markov: được ứng dụng rộng rãi trong việc tính toán sự phân bố sau một thời gian trong một hệ gồm nhiều trạng thái khép kín như sự phân bố dân số, sự phân bố thị trường v.v.



Hình 9: Mô hình cơ bản của một chuỗi Markov.

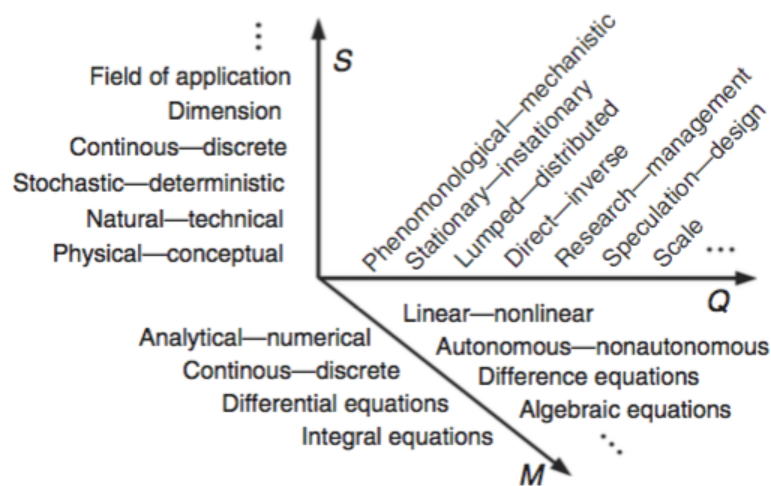
Nhận xét:

- o Trong nhiều trường hợp, người ta áp dụng cùng một M (hoặc gần tương tự M) cho những S khác nhau nhưng có cùng Q .
- o Cần phải thay đổi cách phát biểu bài toán, các giả thiết hoặc input/output cho phù hợp.

6. Phân loại

Dựa trên S , Q , M :

- o Hệ (S)
 - Các vấn đề tự nhiên (Vật lý, Hóa học, Sinh học. . .)
 - Các vấn đề trong xã hội (Kinh tế, Dân số, . . .)
 - Đối tượng vật lý cụ thể vs Đối tượng trừu tượng
 - Có yếu tố ngẫu nhiên vs không ngẫu nhiên
 - Vấn đề liên tục (sự thay đổi theo thời gian, . . .)
 - Vấn đề rời rạc (yếu tố thời gian ko quan trọng. . .)
- o Câu hỏi (Q)
 - Lên kế hoạch
 - Thiết kế
 - Tính toán, dự đoán
 - Tìm mối liên hệ
- o Mô hình (M)
 - Tuyến tính/Không tuyến tính
 - Liên tục/Rời rạc
 - Phương trình đại số/Phương trình vi phân
 - Xác suất/Quy hoạch tuyến tính/Đồ thị



Hình 10: Phân loại mô hình.

Hoạt động trên lớp:

- Hãy thử phân loại một ví dụ ở trên.
- Thảo luận và lập mô hình Toán cho 2 bài toán sau:

1. Mở tiệm Net để cạnh tranh (chế từ đề SIMC 2010)

Phường 9 quận 5 có rất nhiều học sinh và Internet Liên là chuỗi tiệm Net duy nhất. Nhận thấy thị trường tiềm năng, Linh muốn mở 3 tiệm Net để cạnh tranh với chuỗi tiệm Net đã có sẵn của Liên trong khu vực này. Hãy thiết kế một mô hình Toán giúp Linh chọn ra những địa điểm hợp lý nhất cho chuỗi cửa hàng của mình.

2. Sắp xếp lịch tàu vào cảng (chế từ đề SIMC 2008)

Hằng ngày, Cảng Sài Gòn phải đón rất nhiều tàu vào để chất/dỡ hàng. Tuy nhiên, do kích thước của cảng cũng như thời gian tàu dừng ở cảng không như nhau ban quản lý rất cần một lịch cho tàu ra vào để tối ưu hóa năng suất của Cảng. Hãy xây dựng một mô hình Toán học để sắp xếp lịch tàu ra vào hợp lý nhất.

Nhìn chung, một mô hình toán học có thể bao gồm nhiều hơn một loại mô hình nhất định, và là một tổ hợp của một số loại mô hình ở trên. Trong chương trình lần này, chúng ta sẽ được học các công cụ, kiến thức toán học cao cấp và lập trình để xây dựng mô hình toán học cho một vấn đề thực tế quan trọng của Việt Nam.

7. Tham khảo

www.indiana.edu/~hmathmod/modelmodel.html

K. Bliss, K. Fowler, J. Galluzzo, *Math Modeling: Getting Started and Getting Solutions*, link để download:

m3challenge.siam.org/sites/default/files/uploads/siam-guidebook-final-download.pdf

Kai Velten, *Mathematical Modeling and Simulation: Introduction for Scientists and Engineers*, link để download:

mcdtu.files.wordpress.com/2017/03/velten-mathematical-modeling-and-simulation.pdf