

1000

BỘ NÃO

A THOUSAND BRAINS

Copyright © 2021 by Jeffrey C. Hawkins

1000 BỘ NÃO

Bản quyền tiếng Việt © Công ty CP Xuất bản và Dữ liệu ETS, 2024

Bản tiếng Việt được xuất bản theo thỏa thuận với tác giả.

Không phần nào trong xuất bản phẩm này được phép sao chép hay phát hành dưới bất kỳ hình thức hoặc phương tiện nào mà không có sự cho phép trước bằng văn bản của Công ty Cổ phần Xuất bản và Dữ liệu ETS. Chúng tôi luôn mong muốn nhận được những ý kiến đóng góp của quý vị độc giả để sách ngày càng hoàn thiện hơn.

Biên mục trên xuất bản phẩm của Thư viện Quốc gia Việt Nam

Hawkins, Jeff

1000 bộ não - Lý thuyết mới về trí tuệ con người / Jeff Hawkins ; Nguyễn Duy Anh dịch.

- H. : Thế giới ; Công ty Xuất bản và Dữ liệu ETS, 2024. - 372 tr. ; 21 cm

Tên sách tiếng Anh: A thousand brains

ISBN 978-604-345-664-6

1. Khoa học não bộ 2. Trí tuệ

612.82 - dc23

TGF0560p-CIP

Góp ý về sách, liên hệ về bản thảo và bản dịch: publication@etsdata.vn

Liên hệ hợp tác về Nội dung số: digital@etsdata.vn

Liên hệ hợp tác xuất bản & truyền thông trên sách: project@etsdata.vn

Liên hệ dịch vụ tư vấn, đại diện & giao dịch bản quyền: rights@etsdata.vn

Kinh doanh Miền Bắc: Hotline: 093 614 9090 | Email: sales@alphabooks.vn

Kinh doanh Miền Nam: Hotline: 091 942 5402 | Email: sales.hcm@alphabooks.vn

1000

BỘ NÃO

Lý thuyết mới về trí tuệ con người

Nguyễn Duy Anh dịch

Jeff Hawkins

ĐỘI NGŨ TRIỂN KHAI EH

Phụ trách xuất bản: Phùng Thị Ngọc Linh

Điều phối viên: Dương Thu Hương

Thiết kế bìa: Hoàng Khánh

Trình bày: Mỹ Mây

Thư ký xuất bản: Kim Khuyên

Công ty Cp Xuất bản & Dữ liệu ETS (ETS DATA JSC.) - Thành viên của Alpha Publishing Group



Dòng sách Einstein House

Tập trung vào các từ sách Khoa học, Bách khoa thư, Sách tham khảo, Giáo dục cho trẻ em và
Tủ sách cho các nhà quản lý giáo dục. | **Einstein House - Every child is a Genius**
Fanpage: @ets.khoahocluanguonsong | @ets.sach.cho.con

Gamma Dòng sách Gamma

Chuyên phát triển và phát hành các sản phẩm học ngôn ngữ có chất lượng cho cộng đồng
dạy và học ngôn ngữ trong và ngoài nước với ba mảng sản phẩm chính: Gamma Test - Prep,
Gamma Junior và Gamma Gen. | **Gamma - Bridging languages**
Fanpage: @GammaBooks | @GammaJunior | @TuhocTOEIC.IELTSmoingay

Thư viện thông minh Einstein House

Hệ sinh thái đọc tối ưu cho học sinh và thầy cô trong nhà trường, bao gồm không gian mở
đa chức năng, sự kiện trải nghiệm, hệ thống sản phẩm số và các khóa học.

Einstein House - Trái tim của trường học

*ETS DATA JSC. nhận hỗ trợ tư vấn miễn phí bản thảo và quy trình xuất bản cho các cá nhân,
tổ chức muốn viết và xuất bản sách; hỗ trợ xây dựng các khóa học ngoại ngữ, STEAM và kỹ năng
sống cho trẻ; đồng thời mong muốn hợp tác và đồng hành cùng các nhà trường để xây dựng
1.000 không gian thư viện sáng tạo tại các trường học ở Việt Nam.*

Chúng tôi tin rằng, mỗi đứa trẻ là một thiên tài!

Mục lục

Lời tựa của Richard Dawkins 7

PHẦN 1: HIỂU BIẾT MỚI VỀ NÃO BỘ 17

CHƯƠNG 1: Não cũ - Não mới..... 31

CHƯƠNG 2: Ý tưởng lớn của Vernon Mountcastle 44

CHƯƠNG 3: Một mô hình thế giới trong não bộ 54

CHƯƠNG 4: Bật mí bí mật của não bộ..... 68

CHƯƠNG 5: Bản đồ trong não bộ..... 91

CHƯƠNG 6: Khái niệm, ngôn ngữ và tư duy bậc cao 107

CHƯƠNG 7: Thuyết Ngàn não về trí tuệ..... 135

PHẦN 2: TRÍ TUỆ MÁY 165

CHƯƠNG 8: Tại sao không có trí tuệ trong
trí tuệ nhân tạo? 171

CHƯƠNG 9: Khi máy móc có ý thức 194

CHƯƠNG 10: Tương lai của trí tuệ máy 208

CHƯƠNG 11: Các mối nguy hiện sinh của trí tuệ máy..... 230

PHẦN 3: TRÍ TUỆ CỦA CON NGƯỜI	244
CHƯƠNG 12: Những niềm tin sai lệch.....	247
CHƯƠNG 13: Các mối nguy hiện sinh từ trí tuệ con người	264
CHƯƠNG 14: Não và máy hợp nhất.....	282
CHƯƠNG 15: Lập kế hoạch di sản cho nhân loại.....	296
CHƯƠNG 16: Gen và kiến thức.....	317
Những suy nghĩ cuối cùng	343
Tài liệu đọc thêm	350
Lời cảm ơn	357
Nguồn hình ảnh	360
Chỉ mục	361

Lời tựa của Richard Dawkins

Đừng đọc cuốn sách này trước khi đi ngủ. Nó chẳng đáng sợ và cũng không gây ác mộng đâu. Nhưng nó gây phấn khích tột độ và sẽ biến trí óc bạn thành một cơn lốc quay cuồng với những ý tưởng đầy thú vị, và thay vì đi ngủ, bạn sẽ muốn chạy đi kể ngay cho ai đó. Chính tôi, người viết lời tựa, là nạn nhân của cơn lốc này, và tôi hy vọng rằng các bạn cũng thấy vậy.

Charles Darwin là một nhà khoa học hiếm hoi có đủ phương tiện để nghiên cứu ngoài trường đại học mà không cần chính phủ tài trợ. Jeff Hawkins hẳn cũng chẳng thích bị gọi là một nhà khoa học tự chủ theo kiểu Thung lũng Silicon nhưng, chà, bạn thấy điểm tương đồng rồi đấy. Ý tưởng đầy sức mạnh của Darwin là quá cách mạng để gây được chú ý nếu chỉ được viết trong một bài báo ngắn, và gần như chẳng có ai quan tâm đến các bài báo chung của Darwin-Wallace năm 1858 cả. Như chính Darwin đã nói, ý tưởng đó cần được thể hiện bằng cả một pho sách. Và quả vậy, một năm sau, cuốn sách vĩ đại của ông đã làm rung chuyển mọi nền móng thời Victoria.

Lý thuyết Ngàn não (Thousand Brains Theory) của Jeff Hawkins cũng cần viết thành sách vở. Và khái niệm hệ quy chiếu của ông, rằng “bản thân suy nghĩ cũng là một dạng chuyển động” cũng vậy. Hai ý tưởng này đều đủ sâu sắc để lấp đầy một cuốn sách. Nhưng đó không phải là tất cả.

T. H. Huxley đã có câu nói nổi tiếng sau khi đọc *On the Origin of Species* (Nguồn gốc các loài): “Tôi thật ngu ngốc làm sao khi không nghĩ đến điều đó.” Ý tôi không phải là các nhà khoa học não bộ cũng nên nói thế khi đọc xong cuốn sách này; nó chứa nhiều ý tưởng hay, chứ không phải một ý tưởng khổng lồ như của Darwin.

Tôi ngờ rằng không chỉ T. H. Huxley mà cả ba người cháu trai tài giỏi của ông cũng sẽ thích nó: Andrew, vì đã khám phá ra cách hoạt động của xung thần kinh (Hodgkin và Huxley là Watson và Crick của hệ thần kinh¹); Aldous vì những chuyến đi nhìn xa trông rộng và đầy chất thơ đến những vùng xa xôi nhất của trí óc; và Julian, vì đã viết bài thơ dưới đây ca ngợi khả năng của bộ não trong việc xây dựng một mô hình của thực tại, một bản sao thu nhỏ của vũ trụ:

The world of things entered your infant mind

To populate that crystal cabinet.

1. Francis Crick và James Watson tìm ra cấu trúc phân tử chuỗi xoắn kép của ADN, còn Andrew Huxley và Alan Hodgkin khám phá ra mô hình toán học mô tả cách thức các điện thế hoạt động trong tế bào thần kinh được khởi tạo và lan truyền. (BTV)

*Within its walls the strangest partners met,
And things turned thoughts did propagate their kind.*

For, once within, corporeal fact could find

A spirit. Fact and you in mutual debt

*Built there your little microcosm—which yet
Had hugest tasks to its small self assigned.*

Dead men can live there, and converse with stars:

*Equator speaks with pole, and night with day;
Spirit dissolves the world's material bars—*

A million isolations burn away.

*The Universe can live and work and plan,
At last made God within the mind of man.*

(Tạm dịch:

Vạn vật trên đời đi vào trí óc ngây thơ

Và lấp đầy chiếc bình pha lê đó

Trong bình, những sự lạ trên đời gặp gỡ

Và thứ gì hóa thành suy nghĩ sẽ lan xa

Bởi bên trong, những sự thật hữu hình sẽ tìm ra linh khí

Khi chúng và ta, đã mắc nợ nhau nhiều

Xây nên ở đây các hình dung nhỏ bé

Nhưng mang trong mình sứ mệnh cao siêu

Nơi đây, các vì sao chuyện trò cùng người đã khuất
 Xích đạo thì thăm qua địa cực, và đêm-ngày cũng thủ
 thủ với nhau
 Vật chất thì tan ra bởi tinh thần
 Hàng triệu thứ đơn thân thì thi nhau bốc cháy
 Vũ trụ thì cứ thế vận hành và tính toán
 Và Chúa ra đời từ trí óc của nhân gian.)

Bộ não nằm trong bóng tối, nhận biết thế giới bên ngoài qua một cơn mưa đá chứa đầy xung thần kinh của Andrew Huxley. Xung thần kinh từ mắt không khác gì từ tai hay ngón chân cái. Điểm dừng là não bộ, nơi chúng được phân loại. Jeff Hawkins không phải nhà khoa học hay triết gia đầu tiên cho rằng thực tại mà ta cảm nhận là một thực tại được tạo dựng, một mô hình cập nhật thông tin từ các giác quan truyền tới. Nhưng tôi nghĩ Hawkins là người đầu tiên đưa ra ý tưởng hùng hồn rằng có tới hàng ngàn mô hình như vậy chứ không phải một, mỗi mô hình nằm trong một cột thông tin, và các cột này được sắp xếp gọn gàng để tạo thành vỏ não. Có khoảng 150.000 cột như vậy và chúng là ngôi sao của phần đầu tiên trong cuốn sách này, cùng với cái mà ông gọi là “hệ quy chiếu”. Luận điểm của Hawkins về cả hai đều mang tính gợi mở, và xem xét cách các nhà khoa học não bộ khác tiếp nhận chúng là một việc hay ho. Một trong những ý tưởng thú vị của ông là các “cột vỏ não” (cortical column) hoạt động bán tự động

khi mô hình hóa thế giới. Những gì “chúng ta” tiếp nhận được đến từ một kiểu đồng thuận dân chủ giữa chúng.

Dân chủ trong não bộ? Sự đồng thuận, và cả tranh chấp? Thật là một ý tưởng tuyệt vời. Đó là một chủ đề lớn trong cuốn sách. Loài người chúng ta thuộc lớp thú, là nạn nhân của một cuộc tranh chấp lặp đi lặp lại: tranh đấu giữa vùng não bò sát cũ kỹ (thứ vận hành cỗ máy sinh tồn trong vô thức) và vùng não thú mới đang ngồi ghế lái phía trên. Vùng não thú mới này – chính là vỏ não – biết suy nghĩ. Ý thức đến từ đây. Vùng não này nhận thức được quá khứ, hiện tại, tương lai và gửi cho vùng não cũ các chỉ dẫn để thi hành.

Được chọn lọc tự nhiên rèn giũa giữa hàng triệu năm, vùng não cũ khi thấy thiếu đường và cần nạp thêm để sống sẽ thốt lên: “Bánh kem. Muốn bánh kem. Muốn ăn bánh kem. Đưa đây nào.” Được sách vở trau dồi mới chỉ mấy chục năm, vùng não mới khi thấy cơ thể đã quá nhiều đường lại nói: “Không, không. Không phải bánh. Không được. Làm ơn đừng ăn cái bánh đó.” Vùng não cũ nói: “Đau, đau, đau khủng khiếp, hãy *lập tức* chấm dứt cơn đau này.” Vùng não mới lại nói: “Không, không, hãy cố chịu đựng tra tấn, đừng phản bội tổ quốc mà đầu hàng nó. Lòng trung thành với đất nước và đồng đội còn quý hơn tính mạng đấy!”

Xung đột giữa vùng não bò sát cũ và vùng não thú mới mang đến đáp án cho các câu hỏi hóc búa như “Tại sao cơn đau lại phải kinh khủng như vậy?” Rốt cuộc, cơn đau sinh ra

để làm gì? Cơn đau đại diện cho cái chết, là một lời cảnh báo dành cho não bộ: “Đừng làm điều đó nữa: đừng trêu rần, đừng nhạt than nóng, đừng nhảy từ trên cao xuống. Lần này chỉ đau thôi; lần sau là chết đấy!” Nhưng giới kỹ sư thiết kế giờ sẽ nói rằng thứ chúng ta cần là một “lá cờ không gây đau” cho não bộ. Khi lá cờ được phát lên, đừng lặp lại những gì bạn vừa làm. Nhưng thay vì lá cờ dễ chịu này, thứ ta có lại là cơn đau – thường là những cơn đau dữ dội, không thể chịu đựng nổi. Tại sao lại thế? Lá cờ kia có vấn đề gì hay sao?

Câu trả lời có lẽ nằm ở bản chất gây tranh cãi của quá trình ra quyết định trong não bộ: cuộc đấu tranh giữa não cũ và não mới. Não mới sẽ dễ dàng vượt qua não cũ và lá cờ không gây đau sẽ không hoạt động. Việc tra tấn cũng vậy.

Não mới sẽ thoải mái bỏ qua lá cờ giả định của tôi và chịu ong chích, bong gân hay đóng đinh tra tấn bất kể mức độ thế nào, miễn là nó có lý do để “muốn” thế. Não cũ, vốn thực sự “quan tâm” đến việc sống sót để truyền lại gen, có thể “phản đối” trong vô ích. Có thể chọn lọc tự nhiên, vì mục đích sống sót, đã đảm bảo “chiến thắng” cho não cũ bằng cách tạo ra cơn đau mà não mới không thể chịu nổi. Lại ví dụ, nếu não cũ “nhận thức được” rằng mục đích kiểu Darwin của tính dục đang bị phản bội thì việc đeo bao cao su sẽ vô cùng đau đớn.

Hawkins đứng về phía đa số nhà khoa học và triết gia hiểu biết, những người sẽ không ủng hộ thuyết nhị nguyên: không có bóng ma trong cỗ máy, không có linh hồn ma quái

tách biệt khỏi phần cứng đến mức có thể sống sót sau khi phần cứng chết đi, không có nhà hát kiểu Descartes (thuật ngữ của Dan Dennett) với màn hình màu hiển thị một bộ phim về thế giới. Thay vào đó, Hawkins đề xuất nhiều mô hình về thế giới, các thế giới vi mô được tạo dựng, cung cấp thông tin và điều chỉnh bằng cơn mưa xung thần kinh rơi xuống từ các giác quan. Và Hawkins không loại trừ hoàn toàn khả năng chiến thắng tử thần trong tương lai bằng cách “tải” não bộ lên máy tính, nhưng cũng không cho rằng đó là một điều hay.

Những mô hình quan trọng hơn của bộ não là các mô hình của chính cơ thể. Các mô hình này có nhiệm vụ giải thích việc chuyển động cơ thể thay đổi góc nhìn của chúng ta về thế giới bên ngoài “nhà tù thể xác” ra sao. Và điều này có liên quan đến nội dung chính của phần giữa cuốn sách: trí tuệ máy. Jeff Hawkins và cả tôi rất tôn trọng những người bạn thông thái của mình, những người sợ rằng cách tiếp cận máy móc siêu trí tuệ sẽ thay thế chúng ta, chinh phục hoặc thậm chí tận diệt chúng ta. Nhưng Hawkins không sợ chúng, một phần bởi kỹ năng cờ vua hay cờ vây siêu phàm không giúp chúng đương đầu với sự phức tạp của thế giới thực. Những đứa trẻ không biết chơi cờ vua cũng “biết nước tràn ra, bóng lăn và chó sủa ra sao. Chúng biết dùng bút chì, bút đánh dấu, giấy và keo dán. Chúng cũng biết gỡ sách như thế nào và biết giấy thì dễ rách.” Và chúng cũng nhận thức được hình ảnh của bản thân (self-image), một hình ảnh về cơ thể đưa chúng vào thế giới thực và giúp chúng di chuyển dễ dàng trong đó.

Hawkins không đánh giá thấp sức mạnh của trí tuệ nhân tạo và người máy trong tương lai. Ngược lại là đằng khác. Nhưng ông cho rằng hầu hết nghiên cứu ngày nay đang đi sai hướng. Theo quan điểm của ông, đúng ra chúng ta cần hiểu các phương thức hoạt động của bộ não và mượn chúng để thúc đẩy cho các phương thức này.

Và không có lý do gì để (thật lòng, xin đừng) mượn các phương thức của não cũ, những tham-sân-si cùng những cảm xúc và nỗi sợ của nó, vì chúng có thể lái ta vào các con đường mà não mới xem là nguy hại. Ít nhất là có hại khi nhìn nhận theo quan điểm của những giá trị mà tôi và Hawkins (gần như chắc chắn là cả bạn) đều đánh giá cao. Hawkins hiểu các giá trị Khai sáng của chúng ta phải – và chắc chắn – rất khác giá trị khởi thủy của các gen vị kỷ. Theo quan điểm của ông (mà tôi ngờ là sẽ gây tranh cãi), không có não cũ thì sẽ không có lý nào để dự đoán AI sẽ nuôi dưỡng những cảm xúc tiêu cực với chúng ta. Cũng vì vậy, và có thể cũng gây tranh cãi, ông không cho rằng tắt nguồn một AI có ý thức là hành vi giết chóc: Không có não cũ, AI không thể thấy buồn hay sợ, và cũng không muốn sinh tồn.

Trong chương “Gen vs Kiến thức”, chúng ta nói về khác biệt rõ ràng giữa mục tiêu của não cũ (phục vụ các gen vị kỷ) và não mới (kiến thức). Vỏ não người – duy nhất trong mọi loài động vật và chưa từng có tiền lệ trong mọi niên đại địa chất – có khả năng chống lại mệnh lệnh của các gen vị kỷ, và đây là hào quang của riêng nó. Chúng ta có thể tận hưởng tình dục mà không cần sinh sản. Chúng ta có thể tận hiến cho triết

học, toán học, thơ ca, thiên văn, âm nhạc, địa chất hay hơi ấm của tình yêu, bất chấp lời thúc giục di truyền của não cũ rằng những điều này là lãng phí thời gian mà đáng ra “nên” được dùng để chiến đấu và theo đuổi nhiều bạn tình: “Tôi thấy ta có lựa chọn sâu sắc cần đưa ra: ủng hộ não cũ hoặc não mới. Chúng ta có muốn tương lai của mình bị định hướng bởi những thứ đã tạo nên mình là chọn lọc tự nhiên, cạnh tranh và động lực của các gen vị kỷ? Hay muốn tương lai được dẫn dắt bởi trí tuệ và khao khát tìm hiểu thế giới?”

Tôi đã bắt đầu bằng cách trích dẫn nhận xét vô cùng khiêm tốn của T. H. Huxley khi đọc xong *Nguồn gốc các loài* của Darwin. Tôi sẽ kết thúc bằng một trong rất nhiều ý tưởng hấp dẫn của Jeff Hawkins – mà ông trình bày gọn gàng trong vài trang giấy – và điều này làm tôi giống với Huxley. Hawkins cảm thấy cần một tấm bia mang tầm vũ trụ để cả thiên hà biết rằng chúng ta từng ở đây và có khả năng nói ra sự thật, vì ông cho rằng mọi nền văn minh đều sớm lụi tàn. Trên thang thời gian thiên văn, từ khi một nền văn minh phát minh ra truyền tin điện tử đến khi tuyệt chủng chỉ ngắn ngủi như đốm sáng đom đóm. Chẳng mấy khi hai đốm sáng như vậy chớp lên cùng lúc. Thứ chúng ta cần – tấm bia mà tôi vừa nhắc đến – là thông điệp “Chúng tôi từng ở đây” chứ không phải “Chúng tôi ở đây”. Tấm bia này cũng phải trường tồn cùng vũ trụ: không chỉ có thể thấy từ hàng tỷ tỷ ki-lô-mét mà còn phải tồn tại hàng tỷ năm để truyền tải thông điệp khi các đốm sáng trí tuệ khác lóe lên, khi mà toàn nhân loại đã về với

cát bụi. Truyền đi các số nguyên tố hoặc số pi sẽ chẳng thành công, cũng như sóng radio hay tia laser vậy. Sóng radio hay laser chắc chắn là dấu hiệu của trí tuệ sinh học – lý do SETI (chương trình tìm kiếm trí tuệ ngoài Trái Đất) và khoa học viễn tưởng “ghim hàng” những thứ này – nhưng chúng tồn tại quá ngắn. Vậy tín hiệu nào sẽ tồn tại đủ lâu và dễ phát hiện từ khoảng cách lớn đến vậy từ mọi hướng? Hawkins khơi dậy tinh thần Huxley trong tôi chính ở điểm này.

Điều này hiện nằm ngoài tầm tay chúng ta nhưng trong tương lai, trước khi đốm sáng của nhân loại vụt tắt, chúng ta có thể đưa một loạt vệ tinh vào quỹ đạo Mặt Trời để “chặn ánh sáng Mặt Trời theo một quy luật trái tự nhiên. Các vệ tinh cản sáng này sẽ quay quanh Mặt Trời hàng triệu năm, rất lâu sau khi chúng ta tan biến, và có thể phát hiện từ xa.” Ngay cả khi khoảng cách của các vệ tinh cản sáng này không phải một dãy số nguyên tố theo nghĩa đen thì thông điệp vẫn rất rõ ràng: “Sự sống thông minh từng ở đây.”

Một điều khác khiến tôi vui thích – tôi viết lời tựa ngắn này cho Jeff Hawkins như một lời cảm ơn dành cho cuốn sách tuyệt vời của ông – là việc một thông điệp vũ trụ mã hóa dưới dạng một mẫu hình từ khoảng trống giữa các gai Mặt Trời (với ví dụ trên là các “phản gai” vì các vệ tinh của ông cản sáng) sẽ sử dụng cùng một loại mã giống của các neuron.

Cuốn sách này viết về việc não bộ hoạt động như thế nào, theo cách khiến ta chẳng thể thôi phẩn chấn.

PHẦN 1

Hiểu biết mới về não bộ

Các tế bào trong đầu bạn đang đọc chữ. Nghĩ mà xem, điều đó đáng chú ý biết bao. Các tế bào rất đơn giản. Một tế bào đơn lẻ không thể đọc, suy nghĩ hay làm bất cứ điều gì. Tuy nhiên, nếu kết hợp các tế bào để tạo thành não bộ, chúng không chỉ đọc mà còn viết sách. Chúng có thể thiết kế nhà cửa, phát minh ra công nghệ và giải mã các bí ẩn của vũ trụ. Làm sao một bộ não với toàn các tế bào đơn giản lại tạo ra được trí tuệ là một câu hỏi vô cùng lý thú và vẫn còn là bí ẩn.

Hiểu cách não bộ hoạt động được coi là thách thức lớn với nhân loại. Nhiệm vụ này đã tạo ra hàng chục sáng kiến quốc gia và quốc tế, như *Dự án bộ não người của châu Âu* và *Sáng kiến bộ não quốc tế*. Vô vàn nhà thần kinh học với vô số chuyên ngành ở gần như mọi quốc gia đang cố gắng tìm hiểu về bộ não. Tuy các nhà thần kinh học nghiên cứu não của các loài động vật và đặt ra nhiều câu hỏi khác nhau, nhưng mục tiêu cuối cùng vẫn là tìm hiểu xem trí tuệ con người được bộ não tạo ra như thế nào.

Có thể bạn sẽ ngạc nhiên khi tôi nói rằng bộ não người vẫn còn là bí ẩn. Hằng năm, nhiều khám phá mới liên quan đến não bộ được công bố, sách mới về não bộ được xuất bản và các nhà nghiên cứu trong các lĩnh vực liên quan như trí tuệ nhân tạo vẫn thường tuyên bố rằng những sáng tạo của họ đã tiệm cận, chẳng hạn, não chuột hay não mèo. Từ đó, ta có thể dễ dàng kết luận rằng các nhà khoa học có một ý tưởng tương đối hay ho về cách thức hoạt động của bộ não. Nhưng nếu được hỏi, hầu hết các nhà khoa học thần kinh sẽ thừa nhận chúng ta vẫn còn rất mù mờ. Ta đã biết được lượng lớn kiến thức và dữ kiện về bộ não, nhưng lại hiểu rất ít về hoạt động tổng thể của nó.

Năm 1979, Francis Crick, người nổi tiếng với công trình về ADN, đã viết một tiểu luận về tình trạng của khoa học não bộ với tiêu đề *Thinking About the Brain* (Suy nghĩ về bộ não). Ông mô tả lượng lớn dữ kiện mà các nhà khoa học đã thu thập được về bộ não, nhưng lại kết luận rằng “tuy liên tục tích lũy những kiến thức chi tiết nhưng cách thức hoạt động của bộ não người vẫn còn là một bí ẩn sâu sắc”. Ông tiếp lời: “Rõ ràng là còn thiếu một khung ý tưởng rộng để diễn giải những kết quả này.”

Crick thấy rằng các nhà khoa học đã thu thập dữ liệu về não bộ trong nhiều thập kỷ. Họ có vô vàn dữ kiện, nhưng không ai tìm ra cách tập hợp những dữ kiện đó thành một thứ có ý nghĩa. Bộ não giống như một bức hình ghép khổng lồ với hàng ngàn mảnh. Những mảnh ghép này đang bày ra trước

mắt, nhưng ta không thể tìm ra ý nghĩa của chúng. Không ai biết đáp án trông ra sao. Theo Crick, bộ não bí ẩn không phải vì chúng ta chưa thu thập đủ dữ liệu, mà vì chúng ta không biết cách sắp xếp những mảnh ghép đã có. Trong 40 năm kể từ khi Crick viết bài luận trên, đã có nhiều khám phá quan trọng về bộ não. Tôi sẽ nói về chúng sau nhưng nhìn chung, quan sát của ông là đúng. Các tế bào trong não tạo ra trí tuệ ra sao vẫn hoàn toàn bí ẩn. Với nhiều mảnh ghép được thu thập mỗi năm, đôi khi chúng ta cảm thấy như mình đang ngày càng xa rời mục tiêu tìm hiểu não bộ chứ không phải gần hơn.

Tôi đã đọc bài luận của Crick khi còn trẻ và được nó truyền cảm hứng. Tôi cảm thấy bí ẩn về não bộ có thể được giải quyết trong khoảng thời gian mình còn sống, và tôi đã theo đuổi mục tiêu này từ đó. Trong 15 năm qua, tôi đã điều hành một nhóm ở Thung lũng Silicon nghiên cứu về tân vỏ não (neocortex). Tân vỏ não chiếm khoảng 70% thể tích não người, chịu trách nhiệm về mọi thứ liên quan đến trí tuệ, từ các giác quan (thị giác, cảm giác và thính giác), ngôn ngữ dưới mọi hình thức, đến các hình thức tư duy trừu tượng như toán và triết học. Mục tiêu của chúng tôi là tìm hiểu cách hoạt động của tân vỏ não với mức chi tiết đủ để giải thích đặc điểm sinh học của não bộ và dùng các nguyên tắc ấy để chế tạo những cỗ máy thông minh.

Đầu năm 2016, tiến độ nghiên cứu tăng tốc đáng kể. Chúng tôi đã có một bước đột phá về hiểu biết. Chúng tôi nhận ra rằng mình cùng các nhà khoa học khác đã bỏ sót một

thành phần quan trọng. Với góc nhìn mới này, chúng tôi đã thấy các mảnh ghép lắp được vào nhau ra sao. Nói cách khác, tôi tin rằng chúng tôi đã khám phá ra bộ khung mà Crick đã nhắc đến, không chỉ giải thích những điều cơ bản về cách hoạt động của tân vỏ não mà còn tạo ra một cách nghĩ mới về trí tuệ. Chúng ta vẫn chưa có một lý thuyết hoàn chỉnh về bộ não và phải còn lâu mới có. Các lĩnh vực khoa học thường bắt đầu với một khung lý thuyết, và chỉ khi đã có khung thì các chi tiết mới được hoàn thiện. Có lẽ ví dụ nổi tiếng nhất là thuyết tiến hóa của Darwin. Ông đã đề xuất một cách nghĩ mới mẻ, táo bạo về nguồn gốc các loài, nhưng các chi tiết (như cách hoạt động của gen và ADN) thì phải nhiều năm sau mới được biết đến.

Để có trí tuệ, bộ não phải học rất nhiều điều về thế giới. Tôi không chỉ nói về những gì chúng ta học ở trường mà cả những thứ cơ bản, chẳng hạn như việc các đối tượng hằng ngày có hình dạng, âm thanh như thế nào hay có cảm giác ra sao. Chúng ta phải học mọi thứ, từ cách đóng mở cửa đến cách phản hồi của các ứng dụng khi chạm vào màn hình điện thoại. Chúng ta cần tìm hiểu xem mọi thứ nằm ở đâu, từ nơi bạn cất giữ tài sản cá nhân trong nhà đến vị trí của bưu điện hay thư viện trong thành phố. Và tất nhiên, chúng ta cũng học các khái niệm cao cấp như “lòng trắc ẩn” hay “chính phủ”. Trên hết, mỗi chúng ta đều học nghĩa của hàng vạn từ, đều sở hữu lượng kiến thức khổng lồ về thế giới. Một số kỹ năng cơ bản được ghi sẵn trong gen như cách ăn uống hoặc cách vượt

qua cơn đau, nhưng hầu hết kiến thức ta biết về thế giới là do học tập.

Các nhà khoa học nói rằng bộ não học được một mô hình về thế giới. Từ “mô hình” ngụ ý rằng những gì ta biết không chỉ được lưu dưới dạng dữ kiện rời rạc mà còn được tổ chức để phản ánh cấu trúc của thế giới và mọi thứ trong đó. Để biết xe đạp là gì, chúng ta không nhớ danh sách dữ kiện về xe đạp, mà não tạo ra một mô hình xe đạp với các bộ phận, cách lắp ráp chúng vào nhau, cách các bộ phận chuyển động và hoạt động cùng nhau. Để nhận diện một thứ, trước tiên cần tìm hiểu xem nó trông như thế nào và tạo ra cảm giác gì, còn để đạt được mục tiêu, cần tìm hiểu xem mọi thứ trên thế giới phản hồi ra sao khi ta tương tác với chúng. Trí tuệ gắn bó mật thiết với mô hình thế giới của bộ não; để hiểu được cách não bộ tạo ra trí tuệ, chúng ta cần biết não bộ (vốn chỉ được tạo thành từ các tế bào đơn giản) học một mô hình về thế giới và mọi thứ trong đó ra sao.

Khám phá năm 2016 của chúng tôi giải thích cách bộ não học mô hình này. Chúng tôi suy luận rằng tân vỏ não lưu trữ mọi kiến thức của con người bằng các hệ quy chiếu. Tôi sẽ lý giải các hệ này đầy đủ hơn sau, còn giờ hãy coi một bản đồ giấy như một phép loại suy. Bản đồ là một loại mô hình: bản đồ thị trấn là mô hình của thị trấn, còn các đường lưới như kinh tuyến và vĩ tuyến là một loại hệ quy chiếu mang lại cấu trúc cho bản đồ. Hệ quy chiếu cho bạn biết vị trí tương đối của các đối tượng, và cách đạt được một mục tiêu cụ thể như đi từ điểm

này sang điểm kia. Chúng tôi nhận ra mô hình thế giới của bộ não được xây dựng bằng các hệ quy chiếu giống như bản đồ, nhưng không phải một mà hàng trăm ngàn hệ quy chiếu. Thậm chí, giờ chúng ta hiểu rằng hầu hết tế bào trong vỏ não được dành riêng để tạo ra và vận dụng các hệ quy chiếu, rồi các hệ này được não bộ dùng để suy nghĩ và lập kế hoạch.

Với góc nhìn mới này, đáp án cho một số câu hỏi lớn nhất của khoa học thần kinh bắt đầu xuất hiện. Những câu hỏi như là: Thông tin đầu vào từ các giác quan của chúng ta hợp nhất thành một trải nghiệm đơn nhất bằng cách nào? Điều gì xảy ra khi chúng ta suy nghĩ? Làm thế nào hai người lại có niềm tin khác nhau từ những quan sát giống nhau? Và tại sao chúng ta có ý thức về bản ngã?

Cuốn sách này kể về những khám phá nói trên và ý nghĩa của chúng với tương lai nhân loại. Hầu hết nội dung đã được công bố trên các tạp chí khoa học và đường link tham khảo được cung cấp ở cuối sách. Nhưng các bài báo khoa học chưa hẳn là phù hợp để giải thích các lý thuyết quy mô lớn, đặc biệt là theo cách mà các độc giả thường thức có thể hiểu.

Cuốn sách có ba phần. Phần đầu tiên mô tả lý thuyết về hệ quy chiếu của chúng tôi, và chúng tôi gọi đó là Lý thuyết Ngàn não. Lý thuyết này một phần dựa trên suy luận logic, nên tôi sẽ trình bày các bước mà chúng tôi đã thực hiện để đi đến kết luận. Tôi cũng giới thiệu vài thông tin lịch sử cơ bản để giúp bạn thấy lý thuyết này liên quan ra sao đến lịch sử tư

duy về bộ não. Đọc xong phần đầu tiên, hy vọng bạn phần nào hiểu được điều gì xảy ra trong đầu khi chúng ta suy nghĩ và hành động trong thế giới thực, và ý nghĩa của trí tuệ là gì.

Phần thứ hai của sách nói về trí tuệ máy. Máy móc thông minh sẽ biến đổi thế kỷ 21 như những gì máy tính đã làm với thế kỷ 20. Lý thuyết Ngàn não giải thích tại sao AI ngày nay chưa thông minh và chúng ta cần làm gì để tạo ra những cỗ máy thực sự có trí tuệ. Tôi sẽ mô tả những cỗ máy thông minh trong tương lai và cách chúng ta có thể sử dụng chúng. Tôi giải thích tại sao một số máy móc có thể có ý thức và các biện pháp mà chúng ta có thể làm. Cuối cùng, nhiều người sợ rằng máy móc thông minh là một mối nguy hiểm sinh, rằng chúng ta sắp tạo ra một công nghệ hủy diệt loài người. Tôi không đồng ý. Những khám phá của chúng tôi chỉ ra tại sao trí tuệ máy lại vô hại. Tuy vậy, cách con người sử dụng công nghệ mạnh mẽ này lại là hiểm họa.

Trong phần thứ ba, tôi xem xét tình trạng con người từ góc độ não bộ và trí tuệ. Mô hình thế giới của bộ não chứa mô hình về chính con người. Điều này dẫn đến một sự thật kỳ lạ rằng những gì bạn và tôi cảm nhận trong từng khoảnh khắc chỉ là bản mô phỏng của thế giới chứ không phải thế giới thực. Một hệ quả của Lý thuyết Ngàn não là niềm tin của chúng ta về thế giới có thể không đúng. Tôi sẽ giải thích điều này xảy ra như thế nào, tại sao những niềm tin sai lệch có thể khó loại bỏ và việc chúng kết hợp với những cảm xúc nguyên thủy đe dọa sự sống sót của chúng ta trong dài hạn ra sao.

Các chương cuối thảo luận về điều mà theo tôi là quyết định hệ trọng nhất mà giống loài chúng ta sẽ đối mặt. Có hai hướng suy nghĩ về con người. Thứ nhất, chúng ta là sinh vật thành phẩm của tiến hóa và chọn lọc tự nhiên. Từ quan điểm này, con người được định dạng bằng gen và mục đích tồn tại là để tái tạo chúng. Nhưng chúng ta hiện đang dần thoát khỏi quá khứ thuần sinh học của mình. Chúng ta đã trở thành một loài thông minh. Chúng ta là loài đầu tiên trên Trái Đất biết kích thước và tuổi của vũ trụ. Chúng ta là loài đầu tiên biết về quá trình biến đổi của Trái Đất cũng như quá trình tiến hóa của chính giống loài mình. Chúng ta là loài đầu tiên phát triển các công cụ cho phép khám phá và tìm hiểu những bí mật của vũ trụ. Từ quan điểm này, con người được nhận diện bằng trí tuệ và kiến thức chứ không phải bằng gen. Quyết định mà ta phải đối mặt khi nghĩ về tương lai là nên tiếp tục để quá khứ sinh học chi phối hay đi theo trí tuệ của mình, thứ chỉ vừa xuất hiện?

Có thể chúng ta không làm được cả hai. Chúng ta đang tạo ra những công nghệ mạnh mẽ về cơ bản có thể thay đổi cả hành tinh này, điều biến sinh học và chẳng bao lâu nữa là tạo ra những cỗ máy thông minh hơn chính mình. Nhưng chúng ta vẫn có những hành vi nguyên thủy đã giúp mình tồn tại đến nay. Sự kết hợp này thực sự là một mối nguy hiểm sinh cần giải quyết. Nếu sẵn sàng coi trí tuệ và kiến thức là những yếu tố định nghĩa con người chứ không phải gen, thì có lẽ ta có thể tạo ra một tương lai dài hơn với mục đích cao cả hơn.

Hành trình dẫn đến Lý thuyết Ngăn não rất dài và phức tạp. Tôi học kỹ thuật điện ở đại học và chỉ vừa bắt đầu công việc đầu tiên tại Intel khi đọc bài luận của Francis Crick. Nó tác động mạnh đến mức tôi quyết định chuyển nghề và cống hiến cho việc nghiên cứu não bộ. Sau khi nỗ lực xin một vị trí nghiên cứu về não bộ tại Intel không thành, tôi đăng ký làm nghiên cứu sinh tại phòng thí nghiệm AI của MIT. (Tôi cảm thấy cách tốt nhất để tạo ra những cỗ máy thông minh là nghiên cứu về bộ não trước.) Khi được các giảng viên MIT phỏng vấn, đề xuất tạo ra những cỗ máy thông minh dựa trên lý thuyết về não bộ của tôi đã bị từ chối. Họ nói rằng bộ não chỉ là một máy tính lộn xộn và nghiên cứu nó chẳng ích gì. Chán nhưng không nản, tôi đăng ký chương trình tiến sĩ khoa học thần kinh tại Đại học California, Berkeley. Tôi bắt đầu học từ tháng 1 năm 1986.

Khi đến Berkeley, tôi đã liên hệ với trưởng nhóm sau đại học về sinh học thần kinh là Tiến sĩ Frank Werblin để xin lời khuyên. Ông yêu cầu tôi viết một bài mô tả nghiên cứu mà tôi muốn thực hiện cho luận án tiến sĩ của mình. Trong bài viết, tôi giải thích rằng mình muốn nghiên cứu lý thuyết về tân vỏ não. Tôi biết mình muốn tiếp cận vấn đề bằng cách nghiên cứu xem tân vỏ não đưa ra dự đoán như thế nào. Thấy Werblin đã cho một số giảng viên đọc bài viết của tôi và nó được đón nhận nồng nhiệt. Ông nói rằng tham vọng của tôi thật đáng ngưỡng mộ, cách tiếp cận của tôi rất hợp lý và vấn đề tôi muốn nghiên cứu là một trong những vấn đề quan trọng nhất của

khoa học, nhưng – tôi không lường trước được điều này – ông không biết tôi sẽ theo đuổi ước mơ của mình ra sao vào thời điểm đó. Là cử nhân khoa học thần kinh, tôi sẽ phải làm việc cho một giáo sư, làm công việc tương tự như những gì giáo sư đã làm. Và không ai ở Berkeley hoặc bất cứ nơi nào ông biết đang làm điều đủ gần với những gì tôi muốn thực hiện.

Việc cố gắng phát triển một lý thuyết tổng thể về chức năng não bộ được coi là quá tham vọng, nên cũng quá rủi ro. Nếu một sinh viên làm việc này trong năm năm mà không có tiến triển thì có thể không tốt nghiệp được. Các giáo sư cũng gặp rủi ro tương tự; họ có thể không được làm giảng viên cơ hữu. Các cơ quan tài trợ cũng nghĩ rằng nó quá rủi ro, và các đề cương nghiên cứu tập trung vào lý thuyết cũng thường xuyên bị từ chối.

Tôi có thể chọn làm việc trong một phòng thí nghiệm, nhưng sau vài lần phỏng vấn, tôi biết việc này không phù hợp với mình: tôi sẽ dành phần lớn thời gian để huấn luyện động vật, chế tạo thiết bị thí nghiệm và thu thập dữ liệu. Bất kỳ lý thuyết nào tôi phát triển sẽ đều bị giới hạn ở phần nào được nghiên cứu trong phòng thí nghiệm đó.

Trong hai năm tiếp theo, tôi ngồi cả ngày trong thư viện đại học, đọc hết bài báo này đến bài báo khác về khoa học thần kinh. Tôi đọc hàng trăm bài, gồm tất cả những bài báo quan trọng nhất được xuất bản trong hơn 50 năm trước. Tôi cũng đọc xem các nhà tâm lý học, ngôn ngữ học, toán học và

triết học nghĩ gì về bộ não và trí tuệ. Tôi đã được tiếp thu một nền giáo dục hạng nhất nhưng hơi khác thường. Sau hai năm tự học, tôi phải thay đổi. Tôi đã nghĩ ra một kế hoạch. Tôi sẽ quay lại ngành công nghiệp điện toán trong bốn năm rồi sẽ đánh giá lại các cơ hội của mình trong giới học thuật. Nghĩ vậy, tôi bèn quay về nghiên cứu phát triển máy tính ở Thung lũng Silicon.

Tôi bắt đầu thành công trong vai trò doanh nhân. Từ năm 1988-1992, tôi đã tạo ra một trong những máy tính bảng đầu tiên: GridPad. Năm 1992, tôi thành lập Palm Computing, bắt đầu thiết kế một số máy tính cầm tay và điện thoại thông minh đầu tiên như PalmPilot và Treo trong mười năm. Bất cứ ai làm việc với tôi ở Palm đều biết tâm trí tôi dành cho khoa học thần kinh, rằng công việc trong lĩnh vực điện toán di động chỉ là tạm thời. Thật hào hứng biết bao khi thiết kế được một số loại máy tính cầm tay và điện thoại thông minh đầu tiên, và tôi biết hàng tỷ người sẽ dựa vào những thiết bị này. Nhưng tôi cũng cảm thấy hiểu được bộ não còn quan trọng hơn nữa. Tôi tin rằng so với máy tính, lý thuyết về não bộ sẽ có tác động tích cực lớn hơn đến tương lai nhân loại. Do đó, tôi cần quay lại nghiên cứu não bộ.

Vì chẳng có lúc nào là tiện, nên tôi đã chọn đại một ngày để rời khỏi chính những doanh nghiệp mà mình chung tay xây dựng. Với sự hỗ trợ và thúc đẩy của vài người bạn cũng là nhà thần kinh học (đáng chú ý là Bob Knight tại UC Berkeley, Bruno Olshausen tại UC Davis và Steve Zornetzer

tại NASA Ames Research), tôi đã thành lập viện khoa học thần kinh Redwood (RNI) năm 2002. RNI chỉ tập trung vào lý thuyết về tân vỏ não và có mười nhà khoa học làm việc toàn thời gian. Tất cả đều quan tâm đến các lý thuyết quy mô lớn về não bộ, và RNI là một trong những nơi duy nhất trên thế giới không chỉ chấp nhận mà còn mong đợi sự tập trung này. Trong suốt ba năm tôi điều hành RNI, chúng tôi đã có hơn 100 học giả đến thăm, có người ở lại vài ngày hoặc vài tuần. Chúng tôi có các bài giảng hằng tuần mà công chúng cũng có thể tham dự, và những giờ giảng này thường biến thành thời gian tranh luận.

Bất kỳ ai từng làm việc tại RNI, kể cả tôi, đều nghĩ rằng điều đó thật tuyệt. Tôi được làm quen và dành thời gian với nhiều nhà thần kinh học hàng đầu thế giới. Điều đó mang đến cho tôi hiểu biết trong nhiều lĩnh vực về khoa học thần kinh, thứ mà một vị trí học thuật thông thường khó lòng có được. Vấn đề là tôi muốn biết đáp án cho một tập hợp câu hỏi cụ thể và không thấy nhóm của mình tiến tới sự đồng thuận về những câu hỏi đó. Các nhà khoa học hài lòng khi ai làm việc này. Vì vậy, sau ba năm điều hành viện nghiên cứu, tôi quyết định rằng cách tốt nhất để đạt mục tiêu là lãnh đạo nhóm nghiên cứu của riêng mình.

Ở mọi mặt khác, RNI đều hoạt động tốt, vì vậy chúng tôi quyết định chuyển nó đến UC Berkeley. Vâng, mười chín năm sau, chính nơi đã nói rằng tôi không thể nghiên cứu lý thuyết não bộ đã quyết định rằng một trung tâm lý thuyết não

bộ chính xác là thứ họ cần. RNI ngày nay tiếp tục là Trung tâm Khoa học Thần kinh Lý thuyết Redwood.

Khi RNI chuyển đến UC Berkeley, tôi và vài đồng nghiệp đã khởi động Numenta, một công ty nghiên cứu độc lập. Mục tiêu chính của chúng tôi là phát triển một lý thuyết về cách hoạt động của tân vỏ não, còn mục tiêu thứ yếu là áp dụng những gì học được về bộ não vào học máy và trí tuệ máy. Numenta cũng giống các phòng nghiên cứu thường thấy ở đại học nhưng linh hoạt hơn. Nó cho phép tôi chỉ đạo một nhóm, đảm bảo rằng mọi người tập trung vào cùng một nhiệm vụ và thử thực hiện những ý tưởng mới thường xuyên.

Numenta đã hơn 15 năm tuổi, nhưng ở một số mặt, chúng tôi vẫn như một công ty mới thành lập. Đi tìm cách hoạt động của tân vỏ não là một con đường trắc trở. Để có tiến bộ, chúng tôi cần sự linh hoạt và tập trung của một môi trường khởi nghiệp, cũng cần rất nhiều kiên nhẫn – thứ mà một công ty khởi nghiệp thường không có. Khám phá quan trọng đầu tiên – cách các neuron đưa ra dự đoán – xuất hiện năm 2010, năm năm sau khi khởi động. Đến năm 2016, chúng tôi phát hiện ra rằng trong tân vỏ não có các hệ quy chiếu giống như bản đồ.

Năm 2019, chúng tôi bắt đầu thực hiện nhiệm vụ thứ hai: áp dụng các nguyên tắc của bộ não vào học máy. Đó cũng là năm mà tôi bắt đầu viết cuốn sách này để chia sẻ những gì chúng tôi đã biết.

Tôi kinh ngạc khi thấy rằng thứ duy nhất trong vũ trụ biết vũ trụ tồn tại là khối tế bào nặng chỉ hơn 1 kg trôi nổi trong đầu chúng ta. Nó làm tôi nhớ đến một câu hỏi xưa cũ: Nếu một cái cây đổ trong rừng và không có ai ở đó để nghe thì nó có phát ra âm thanh không? Tương tự, chúng ta có thể đặt câu hỏi: Nếu vũ trụ trải qua quá trình sinh diệt mà không có bộ não nào nhận biết được, thì vũ trụ có thực sự tồn tại hay không? Ai sẽ biết điều đó? Vài tỷ tế bào lơ lửng trong hộp sọ của bạn không chỉ biết vũ trụ có tồn tại, mà còn biết rằng nó rất cổ xưa và rộng lớn. Những tế bào này đã học được một mô hình về thế giới, điều mà ta biết là không tồn tại ở nơi nào khác. Tôi dành cả đời để tìm hiểu xem não bộ làm vậy như thế nào, và rất phấn khích với những gì chúng tôi tìm hiểu được. Hy vọng bạn cũng vậy. Cùng bắt đầu nào.

CHƯƠNG 1

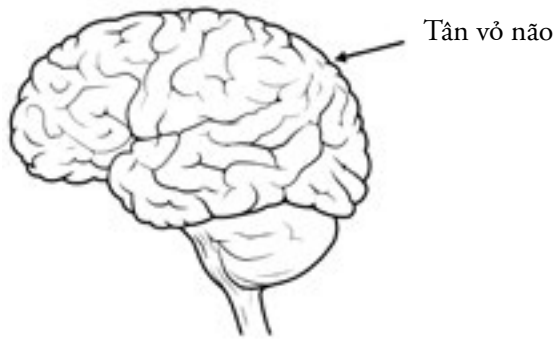
Não cũ - Não mới

Để hiểu cách bộ não tạo ra trí tuệ, bạn cần biết vài điều cơ bản.

Ngay sau khi Charles Darwin công bố thuyết tiến hóa, các nhà sinh học đã nhận ra rằng bản thân bộ não người cũng tiến hóa theo thời gian và chỉ cần nhìn vào nó là thấy được một lịch sử tiến hóa rõ ràng. Không như các loài thường biến mất khi loài mới xuất hiện, bộ não tiến hóa nhờ bổ sung bộ phận mới trên nền các bộ phận cũ. Ví dụ, một số hệ thần kinh lâu đời và đơn giản nhất là tập hợp neuron chạy dọc lưng các con giun nhỏ. Những neuron này cho phép giun thực hiện các chuyển động đơn giản và chúng là tiền thân của tủy sống, chịu trách nhiệm tương tự với nhiều chuyển động cơ bản của chúng ta. Tiếp theo, một khối neuron xuất hiện ở một đầu của cơ thể, kiểm soát các chức năng như tiêu hóa và thở. Khối này là tiền thân của thân não, nơi kiểm soát quá trình tiêu hóa và hô hấp ở người. Thân não mở rộng nhưng không thay thế những gì

đã có. Theo thời gian, não bộ phát triển khả năng thực hiện những hành vi phức tạp bằng cách tiến hóa các bộ phận mới dựa trên các bộ phận đã có. Phương pháp tăng trưởng qua bổ sung này đúng với não của hầu hết các loài động vật phức tạp. Thật dễ hiểu tại sao các phần não cũ vẫn còn đó. Dù thông minh và tinh vi đến đâu thì hô hấp, ăn uống, tình dục và phản xạ vẫn rất quan trọng với sự sống còn của chúng ta.

Phần mới nhất trong bộ não của chúng ta là tân vỏ não (neocortex nghĩa là “lớp mới bên ngoài”). Mọi loài thú đều có tân vỏ não, và chỉ lớp thú mới có. Ở người, tân vỏ não đặc biệt lớn, chiếm khoảng 70% thể tích não. Nếu lấy tân vỏ não ra khỏi đầu và là phẳng, nó sẽ có kích thước bằng một chiếc khăn ăn lớn và dày gấp đôi (khoảng 2,5 mm). Tân vỏ não bao quanh những phần cũ hơn và khi nhìn vào não người, đó là hầu hết những gì bạn thấy (với những nếp gấp đặc trưng), ngoài ra là những phần nhỏ của não cũ và tủy sống nhô ra phía dưới.



Bộ não người

Tân vỏ não là cơ quan của trí tuệ. Hầu như mọi năng lực mà chúng ta gán cho trí tuệ – như thị giác, ngôn ngữ, âm nhạc, toán học, khoa học và kỹ thuật – đều do tân vỏ não tạo ra. Khi chúng ta nghĩ về điều gì đó, việc suy nghĩ này hầu hết do tân vỏ não thực hiện. Tân vỏ não của bạn đang đọc hoặc nghe cuốn sách này, còn tân vỏ não của tôi thì đang viết nó. Nếu muốn hiểu về trí tuệ, ta cần biết tân vỏ não làm gì và làm ra sao.

Loài vật không có tân vỏ não vẫn có thể sống một cuộc sống phức tạp. Bộ não của cá sấu gần tương đương não người, nhưng không có tân vỏ não đích thực. Cá sấu có hành vi phức tạp, quan tâm đến con cái và có khả năng định vị. Hầu hết sẽ nói rằng cá sấu cũng thông minh ở một mức độ nào đó nhưng còn xa mới bì được với con người.

Tân vỏ não và các phần cũ hơn của não kết nối bằng các sợi thần kinh, nên ta không thể coi chúng như những cơ quan hoàn toàn riêng biệt. Chúng như những người bạn cùng phòng với năng lực và tính cách khác nhau nhưng phải hợp tác để hoàn thành công việc. Tân vỏ não nằm ở một vị trí vô cùng bất công khi không được trực tiếp kiểm soát hành vi. Không giống các phần khác của não, không tế bào nào trong tân vỏ não kết nối trực tiếp với các cơ nên tự nó không thể khiến bất kỳ cơ nào cử động. Khi muốn làm điều gì đó, tân vỏ não sẽ gửi tín hiệu để yêu cầu não cũ thực hiện. Ví dụ, thở là một chức năng của thân não, không cần suy nghĩ hay dữ liệu đầu vào từ tân vỏ não. Tân vỏ não có thể tạm thời kiểm soát hơi thở, như

khi bạn quyết định nín thở. Nhưng nếu phát hiện cơ thể cần nhiều oxy hơn, thân não sẽ bỏ qua tân vỏ não và lấy lại quyền kiểm soát. Tương tự, tân vỏ não có thể nghĩ: “Đừng ăn miếng bánh này, không tốt cho sức khỏe.” Nhưng nếu những phần cũ hơn và nguyên thủy hơn trong não nói “Trông nó rất thơm ngon, hãy ăn đi” thì khó mà cưỡng lại. Trận chiến não cũ-não mới là một chủ đề cơ bản của cuốn sách này và đóng vai trò quan trọng trong việc luận bàn về những rủi ro hiện sinh mà nhân loại phải đối mặt.

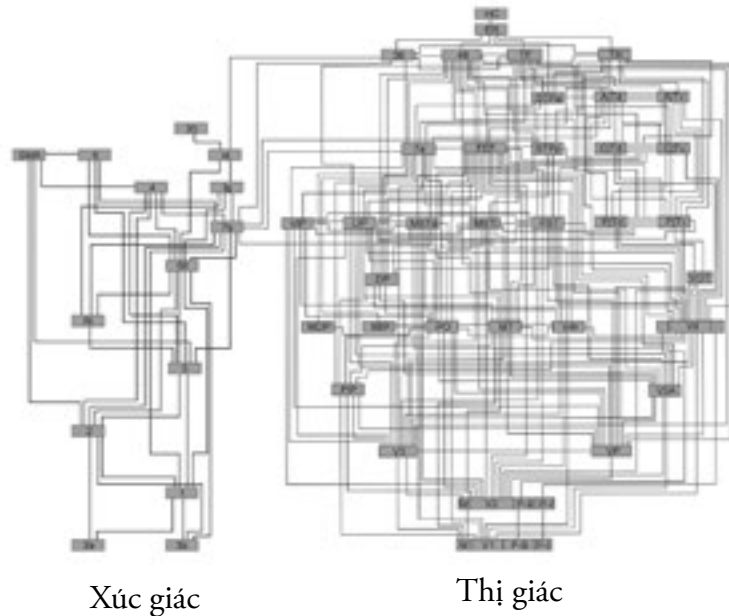
Não cũ có hàng chục cơ quan riêng biệt với chức năng riêng biệt. Chúng khác nhau về trực quan, đồng thời hình dạng, kích thước và kết nối của chúng cũng phản ánh chức năng. Ví dụ, một số cơ quan có kích thước bằng hạt đậu trong hạch hạnh nhân (amygdala), một phần cũ hơn của não, chịu trách nhiệm cho các hình thức của sự gây hấn như gây hấn có chủ đích và gây hấn do bốc đồng.

Tân vỏ não khác biệt đến kinh ngạc. Mặc dù chiếm gần 3/4 thể tích não và chịu trách nhiệm cho vô số chức năng nhận thức, nhưng về bề ngoài của vùng não này không có sự phân chia rõ ràng. Nó cần có các nếp gấp, uốn mới vừa trong hộp sọ, tương tự khi bạn cố nhét khăn ăn vào một ly rượu lớn. Nếu không có các nếp gấp, tân vỏ não sẽ trông hệt như một mảng tế bào phẳng và không được phân chia rõ ràng.

Tuy nhiên, tân vỏ não vẫn được chia thành nhiều phân khu, hoặc vùng, với chức năng khác nhau. Một số vùng chịu

trách nhiệm về thị giác, một số về thính giác, còn số khác là cảm giác. Cũng có các vùng chịu trách nhiệm về ngôn ngữ và lập kế hoạch. Những khiếm khuyết khi tân vỏ não bị tổn thương phụ thuộc vào việc vùng nào bị tổn hại. Tổn thương sau đầu dẫn đến mù lòa còn tổn thương bên trái có thể dẫn đến mất ngôn ngữ.

Các vùng của tân vỏ não được kết nối bằng các bó sợi thần kinh di chuyển bên dưới tân vỏ não, được gọi là chất trắng. Bằng cách cẩn thận lần theo các sợi thần kinh này, các nhà khoa học có thể xác định số lượng các vùng cũng như cách kết nối giữa chúng. Rất khó để nghiên cứu bộ não người nên loài động vật lớp thú phức tạp đầu tiên được phân tích là khỉ macaca. Năm 1991, hai nhà khoa học Daniel Felleman và David Van Essen đã tổng hợp dữ liệu từ hàng chục nghiên cứu riêng biệt để tạo ra một bản minh họa nổi tiếng cho tân vỏ não loài khỉ này. Dưới đây là một hình ảnh mà họ tạo ra (bản đồ tân vỏ não ở người sẽ khác về chi tiết nhưng tương tự về cấu trúc tổng thể).



Các kết nối trong tân vỏ não

Các hình chữ nhật nhỏ ở hình trên biểu thị các vùng khác nhau của tân vỏ não, còn các đường thẳng mô tả dòng thông tin chảy từ vùng này sang vùng khác qua chất trắng.

Một cách giải thích phổ biến cho hình ảnh này là tân vỏ não có tính thứ bậc giống một lưu đồ. Dữ liệu đầu vào từ các giác quan được nạp vào ở phía dưới (trong sơ đồ này, đầu vào từ da ở bên trái còn đầu vào từ mắt ở bên phải). Thông tin đầu vào được xử lý bằng một loạt bước, mỗi bước trích xuất đặc điểm ngày càng phức tạp hơn. Ví dụ: vùng đầu tiên nhận thông tin đầu vào từ mắt có thể phát hiện các mẫu đơn giản

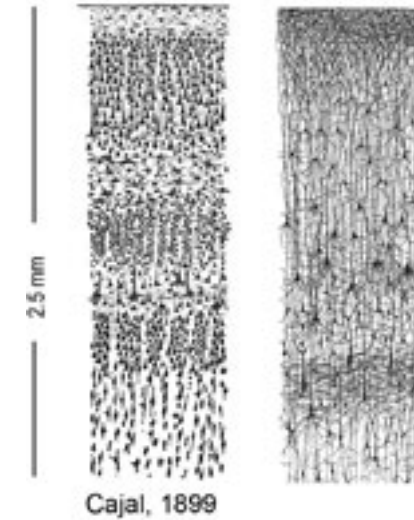
như đường kẻ hoặc cạnh. Khi thông tin này được gửi đến vùng tiếp theo, vùng này có thể phát hiện các đặc điểm phức tạp hơn như góc hoặc hình dạng. Quá trình này tiếp diễn cho đến khi có vùng phát hiện được các đối tượng hoàn chỉnh.

Có rất nhiều bằng chứng hỗ trợ cho cách diễn giải này. Chẳng hạn, khi xem xét các tế bào trong vùng dưới cùng của hệ thống phân cấp, các nhà khoa học thấy chúng phản hồi tốt nhất với các đặc điểm đơn giản, trong khi các tế bào trong vùng kế tiếp lại phản hồi với các đặc điểm phức tạp hơn. Đôi khi họ cũng thấy các tế bào ở vùng cao hơn phản hồi với các đối tượng hoàn chỉnh. Tuy nhiên, có nhiều bằng chứng cho thấy tân vỏ não không giống lưu đồ. Như bạn thấy trong sơ đồ trên, các vùng không chồng lên nhau như trong lưu đồ. Mỗi cấp độ trên sơ đồ có nhiều vùng, và hầu hết các vùng đều kết nối cùng lúc với nhiều cấp độ trên hệ thống phân cấp. Trên thực tế, phần lớn kết nối giữa các vùng đều không giống sơ đồ phân cấp. Ngoài ra, ở mỗi vùng chỉ có một số tế bào hoạt động như các bộ phận phát hiện đặc điểm; các nhà khoa học chưa xác định được đa số các tế bào trong mỗi vùng đang làm gì.

Vậy là chúng ta có một câu đố phức tạp. Tân vỏ não, cơ quan của trí tuệ, được chia thành hàng chục vùng với nhiều nhiệm vụ khác nhau nhưng bề ngoài lại giống hệt nhau. Các khu vực kết nối trong một mô hỗn độn phức tạp đôi khi giống một lưu đồ, nhưng hầu hết thì không giống. Hiện vẫn chưa rõ tại sao cơ quan này lại trông như thế.

Điều hiển nhiên tiếp theo cần làm là quan sát “mạch điện” chi tiết trong 2,5 mm bề dày của tân vỏ não. Có thể tưởng tượng rằng ngay cả khi các vùng trong tân vỏ não trông không khác gì nhau, thì các mạch thần kinh chi tiết tạo ra thị giác, xúc giác và ngôn ngữ bên trong vẫn khác nhau. Nhưng thực ra không phải vậy.

Người đầu tiên nghiên cứu chi tiết mạng lưới bên trong tân vỏ não là Santiago Ramón y Cajal. Cuối những năm 1800, người ta đã tìm ra các kỹ thuật nhuộm màu, cho phép nhìn thấy các neuron riêng lẻ trong não bằng kính hiển vi. Cajal đã sử dụng những vệt màu này để tạo ra hình ảnh của mọi bộ phận trong não. Ông đã tạo ra hàng ngàn hình ảnh và lần đầu tiên cho thấy bộ não trông ra sao ở cấp tế bào. Mọi hình ảnh đẹp đẽ và phức tạp về bộ não của Cajal đều được vẽ tay. Cuối cùng ông đã giành giải Nobel cho công trình đó. Dưới đây là hai hình vẽ tân vỏ não của Cajal, hình bên trái chỉ hiển thị thân tế bào của các neuron còn bên phải thể hiện kết nối giữa các tế bào. Những hình vẽ này cho thấy một lát cắt ngang qua bề dày 2,5 mm của tân vỏ não.



Các neuron trên một mặt cắt của tân vỏ não

Các vệt màu được sử dụng để tạo ra những hình ảnh này chỉ tô màu cho số ít tế bào. May là vậy, vì nếu nhuộm màu hết thì ta sẽ chỉ thấy một màu đen kịt. Hãy nhớ rằng số lượng neuron thực tế lớn hơn những gì bạn thấy ở đây rất nhiều.

Điều đầu tiên mà Cajal và những người khác thấy là các neuron trong tân vỏ não dường như được xếp theo lớp. Các lớp song song với bề mặt tân vỏ não (nằm ngang trong hình) do khác biệt về kích thước giữa các neuron và mật độ mà chúng được sắp xếp. Hãy tưởng tượng bạn đổ đậu Hà Lan, đậu lăng và đậu nành vào một ống thủy tinh với độ dày khoảng 2 cm mỗi loại. Nhìn ống từ bên cạnh, bạn sẽ thấy ba lớp. Bạn cũng

có thể thấy các lớp như vậy ở hình trên. Số lượng các lớp phụ thuộc vào việc ai là người đếm và họ dùng tiêu chí nào để phân lớp. Cajal thấy sáu lớp. Cách giải thích đơn giản là mỗi lớp neuron đang làm một việc khác nhau.

Ngày nay, chúng ta biết rằng có hàng chục loại neuron trong tân vỏ não chứ không chỉ sáu, nhưng các nhà khoa học vẫn sử dụng hệ thuật ngữ sáu lớp. Ví dụ, một loại tế bào có thể được tìm thấy ở Lớp 3 còn loại khác lại ở Lớp 5. Lớp 1 ở bề mặt ngoài cùng của tân vỏ não, nơi gần hộp sọ nhất (ở trên cùng trong bản vẽ của Cajal). Lớp 6 gần trung tâm não nhất, xa hộp sọ nhất. Cần lưu ý rằng các lớp chỉ là hướng dẫn sơ bộ về nơi có thể tìm thấy một loại neuron cụ thể. Điều quan trọng hơn là neuron kết nối với bộ phận nào và bộ phận đó hoạt động ra sao. Sẽ có cả tá loại neuron nếu bạn phân biệt chúng dựa trên đặc điểm kết nối.

Quan sát thứ hai từ những hình ảnh này là hầu hết liên kết neuron đều chạy dọc giữa các lớp. Các neuron có các phần phụ trông giống một cái cây gọi là sợi trục và sợi nhánh, cho phép chúng gửi thông tin cho nhau. Cajal thấy rằng hầu hết sợi trục chạy giữa các lớp, vuông góc với bề mặt của tân vỏ não (từ trên xuống hoặc từ dưới lên, trong hình trang 39). Các neuron trong một số lớp tạo ra kết nối theo chiều ngang với khoảng cách lớn, nhưng hầu hết là theo chiều dọc. Điều này nghĩa là thông tin gửi đến một vùng của tân vỏ não chủ yếu lên xuống giữa các lớp trước khi đến nơi khác.

Trong 120 năm kể từ lần đầu tiên Cajal chụp ảnh não bộ, hàng trăm nhà khoa học đã nghiên cứu tân vỏ não để khám phá càng nhiều chi tiết càng tốt về các neuron và kết nối giữa chúng. Có hàng ngàn bài báo khoa học về chủ đề này và tôi không thể tóm tắt hết. Thay vào đó, tôi muốn nhấn mạnh ba quan sát tổng quát.

1. Các mạch cục bộ trong tân vỏ não rất phức tạp

Có khoảng 100.000 neuron, 500 triệu liên kết giữa các neuron (khop thần kinh) và vài kilomet sợi trục, sợi nhánh trong 1 mm^2 của tân vỏ não (khoảng $2,5 \text{ mm}^3$). Hãy tưởng tượng việc giăng vài kilomet dây điện dọc một con đường rồi cố nén thành 2 mm^3 , gần bằng kích thước một hạt gạo. Có hàng chục loại neuron dưới mỗi milimet vuông, mỗi loại tạo ra các kết nối nguyên mẫu với các loại neuron khác. Các nhà khoa học thường mô tả các vùng của tân vỏ não thực hiện những chức năng đơn giản (như phát hiện các đặc điểm), nhưng chỉ cần số ít neuron để làm việc này. Các mạch thần kinh chính xác và cực kỳ phức tạp ở mọi nơi trong tân vỏ não cho thấy mọi khu vực đang làm việc gì đó phức tạp hơn nhiều chứ không phải chỉ phát hiện đặc điểm.

2. Tân vỏ não trông rất đồng đều

Mạch phức tạp của tân vỏ não giống nhau đáng kể ở các vùng thị giác, ngôn ngữ và xúc giác; thậm chí mạch này cũng giống nhau giữa các loài như chuột, mèo và người. Nhưng thực chất

vẫn có khác biệt: một số vùng của tân vỏ não có rất nhiều loại tế bào nào đó và tế bào khác thì ít hơn, đồng thời một số vùng có một loại tế bào không thấy ở nơi khác. Có lẽ bất kỳ hoạt động nào của các vùng như vậy trong tân vỏ não đều được hưởng lợi từ những khác biệt này nhưng nhìn chung, khác biệt giữa các vùng là tương đối nhỏ so với điểm tương đồng.

3. Mọi bộ phận của tân vỏ não đều tạo ra chuyển động

Trong một thời gian dài, người ta tin rằng thông tin đi vào tân vỏ não qua “vùng cảm giác”, lên xuống theo thứ bậc của các vùng và cuối cùng xuống “vùng vận động”. Các tế bào trong vùng vận động nối với neuron trong tủy sống giúp cử động các cơ và tứ chi. Giờ ta đã biết mô tả này không đúng. Ở mọi khu vực đã kiểm tra, các nhà khoa học thấy các tế bào kết nối đến một số vùng liên quan đến chuyển động của não cũ. Ví dụ, các vùng thị giác nhận thông tin đầu vào từ mắt sẽ gửi tín hiệu xuống phần não cũ chịu trách nhiệm di chuyển mắt. Tương tự, các vùng thính giác nhận thông tin đầu vào từ tai sẽ gửi tín hiệu đến phần não cũ có chức năng di chuyển đầu. Việc di chuyển đầu thay đổi những gì ta nghe thấy, cũng giống như di chuyển mắt thay đổi những gì ta nhìn thấy. Bằng chứng chỉ ra rằng hệ thống mạch phức tạp xuất hiện mọi nơi trong tân vỏ não thực hiện nhiệm vụ vận động-cảm giác. Không có vùng vận động thuần túy và cũng không có vùng cảm giác thuần túy.

Tóm lại, tân vỏ não là cơ quan của trí tuệ. Đó là một mảng mô thần kinh có kích thước bằng chiếc khăn ăn được

chia thành hàng chục vùng. Có các vùng chịu trách nhiệm về thị giác, thính giác, xúc giác và ngôn ngữ; cũng có những khu vực tương đối khó gán nhãn, chịu trách nhiệm cho việc lập kế hoạch và tư duy cấp cao. Các vùng được kết nối bằng các bó sợi thần kinh. Một số kết nối giữa các vùng có tính thứ bậc, gợi ý rằng thông tin qua lại giữa các khu vực này cũng có trật tự như trong một lưu đồ. Nhưng một số kết nối khác giữa các khu vực dường như lại ít trật tự hơn, cho thấy thông tin được truyền đi khắp nơi cùng lúc. Mọi vùng trông đều giống nhau về chi tiết, bất kể chức năng của chúng là gì.

Chúng ta sẽ gặp người đầu tiên hiểu được những quan sát này trong chương tiếp theo.

Đây là thời điểm thích hợp để nói về phong cách viết của cuốn sách này. Tôi đang viết cho các tín đồ tò mò về trí tuệ. Mục tiêu của tôi hầu như chỉ xoay quanh việc truyền đạt mọi thứ bạn cần biết để hiểu lý thuyết mới, với giả định rằng hầu hết độc giả sẽ có kiến thức hạn chế về khoa học thần kinh. Tuy nhiên, nếu có kiến thức cơ bản về khoa học thần kinh, bạn sẽ biết tôi đang bỏ qua các chi tiết và đơn giản hóa các chủ đề phức tạp ở đâu; tôi mong các bạn hiểu cho. Cuối sách có một danh mục tài liệu đọc thêm để người quan tâm có thể tìm hiểu chi tiết.

CHƯƠNG 2

Ý tưởng lớn của Vernon Mountcastle

The Mindful Brain (Bộ não biết suy tư) là một cuốn sách mỏng chỉ 100 trang. Xuất bản năm 1978, sách gồm hai tiểu luận về bộ não của hai nhà khoa học lỗi lạc. Trong đó, bài luận của nhà thần kinh học Vernon Mountcastle tại Đại học Johns Hopkins vẫn là chuyên khảo quan trọng và mang tính biểu tượng bậc nhất từng được viết về não. Mountcastle đã đề xuất một cách suy nghĩ về bộ não rất tao nhã – một đặc điểm luôn có ở những lý thuyết lớn – nhưng cũng gây kinh ngạc đến mức vẫn luôn khiến cộng đồng khoa học thần kinh phân cực.

Lần đầu tiên tôi đọc *The Mindful Brain* là năm 1982. Bài luận của Mountcastle lập tức tác động mạnh đến tôi, và bạn sẽ thấy đề xuất của ông ảnh hưởng rất nhiều đến lý thuyết mà tôi trình bày trong cuốn sách này.

Bài viết của Mountcastle chính xác và uyên bác nhưng cũng khó đọc. Nó có tiêu đề không mấy hấp dẫn là “An Organizing Principle for Cerebral Function: The Unit Module and the Distributed System” (Một nguyên tắc tổ chức của chức năng não bộ: module đơn vị và hệ thống phân phối). Những dòng mở đầu của bài luận này khá khó hiểu; tôi trích dẫn chúng dưới đây để bạn hiểu được cảm giác khi đọc nó:

Chẳng gợn chút nghi ngờ về sức ảnh hưởng toàn diện của cuộc cách mạng Darwin giữa thế kỷ 19 với các khái niệm về cấu trúc và chức năng của hệ thần kinh. Ý tưởng của Spencer, Jackson và Sherrington cùng đồng đảo những người ủng hộ bắt nguồn từ lý thuyết theo hướng tiến hóa rằng bộ não phát triển theo kiểu phát sinh chủng loại (phylogeny) nhờ liên tục bổ sung các bộ phận khác lên trên. Theo lý thuyết này, mỗi sự bổ sung hoặc mở rộng mới đều đi kèm với việc phát sinh hành vi phức tạp hơn, đồng thời chinh đốn các bộ phận nằm dưới và nguyên thủy hơn mà chúng điều khiển.

Những gì Mountcastle nói trong ba câu đầu tiên này là bộ não phát triển kích thước theo thời gian bằng cách bổ sung các phần não mới lên trên các phần não cũ. Các phần cũ hơn kiểm soát các hành vi nguyên thủy hơn, còn các phần mới hơn tạo ra các hành vi tinh vi hơn. Hy vọng các bạn thấy điều này quen thuộc, như tôi đã thảo luận về ý tưởng này trong chương trước.

Tuy nhiên, Mountcastle tiếp tục nói rằng dù phần lớn bộ não tăng kích thước nhờ thêm các bộ phận mới lên trên các bộ phận cũ, nhưng đó không phải là cách tân vỏ não phát triển và chiếm đến 70% não bộ. Tân vỏ não “lớn lên” bằng cách tạo ra nhiều bản sao của cùng một thứ: một mạch cơ bản. Hãy tưởng tượng việc xem một video về quá trình phát triển của não. Ban đầu, bộ não rất nhỏ. Một “mảnh” mới xuất hiện ở một đầu, sau đó một mảnh khác hiện ra ngay phía trên, rồi một mảnh khác nữa lại đè lên nó. Vào thời điểm nào đó hàng triệu năm trước, xuất hiện một mảnh mới mà ngày nay chúng ta gọi là tân vỏ não. Tân vỏ não khởi đầu cũng nhỏ rồi lớn lên, không phải bằng cách tạo ra thứ gì mới mà là sao chép lặp đi lặp lại một mạch cơ bản. Khi phát triển, tân vỏ não trải rộng hơn nhưng không dày thêm. Mountcastle lập luận rằng dù tân vỏ não của con người lớn hơn chuột hay chó nhiều, nhưng chúng đều hình thành từ cùng một yếu tố – chúng ta chỉ có nhiều bản sao của yếu tố đó hơn mà thôi.

Bài luận của Mountcastle làm tôi nhớ đến cuốn *On the Origin of Species* (Nguồn gốc các loài) của Charles Darwin. Darwin sợ rằng thuyết tiến hóa của ông sẽ gây náo động. Vì vậy, trong sách, ông đề cập đến rất nhiều tài liệu tương đối nhằm chận về sự biến đổi trong giới động vật trước khi mô tả lý thuyết của mình ở phần cuối. Nhưng ngay cả vậy, ông cũng chưa từng nói rõ rằng tiến hóa xảy ra cả ở người. Khi đọc bài luận của Mountcastle, tôi cũng có ấn tượng tương tự. Tôi cảm thấy Mountcastle biết đề xuất của mình sẽ thu hút ý kiến trái

chiều, nên đã rất thận trọng. Đây là trích dẫn thứ hai trong phần sau bài tiểu luận của Mountcastle:

Nói ngắn gọn, về bản chất vỏ não vận động không có gì là vận động, cũng như vỏ não cảm giác không có gì là cảm giác. Do đó, việc làm sáng tỏ phương thức hoạt động của mạch module cục bộ ở bất kỳ đâu trong tân vỏ não đều có ý nghĩa khá quát hóa lớn.

Trong hai câu này, Mountcastle tóm tắt ý tưởng chính trong bài luận của mình. Ông cho rằng mọi bộ phận của tân vỏ não đều hoạt động theo cùng một nguyên tắc. Tất cả những thứ mà ta coi là trí tuệ – từ thị giác, xúc giác đến ngôn ngữ và tư duy cấp cao – cơ bản đều giống nhau.

Nhắc lại là tân vỏ não được chia thành hàng chục vùng, mỗi vùng thực hiện một chức năng khác nhau. Nếu quan sát tân vỏ não từ bên ngoài, bạn không thể thấy các vùng vì ở đó không có ranh giới, giống như ảnh vệ tinh không cho thấy biên giới chính trị giữa các quốc gia. Nếu quan sát mặt cắt của tân vỏ não, bạn sẽ thấy một cấu trúc phức tạp và chi tiết. Tuy nhiên, các chi tiết trông sẽ giống nhau bất kể lát cắt đi qua vùng nào. Lát cắt qua vỏ não thị giác cũng hết như lát cắt qua vỏ não xúc giác, và cả lát cắt vỏ não ngôn ngữ cũng vậy.

Mountcastle đề xuất rằng các khu vực trông giống nhau là vì chúng làm những việc giống nhau. Điều khiến chúng khác biệt không phải là chức năng nội tại mà là những gì

chúng kết nối. Khi kết nối một vùng vỏ não với mắt, bạn sẽ có thể nhìn; nếu kết nối chính vùng vỏ não đó với tai, bạn sẽ nghe được; nếu kết nối các vùng với nhau, bạn sẽ có các loại tư duy cấp cao như ngôn ngữ. Sau đó, Mountcastle chỉ ra rằng nếu có thể khám phá chức năng cơ bản của bất kỳ bộ phận nào trong tân vỏ não, thì chúng ta sẽ hiểu cách thức hoạt động của toàn bộ cơ quan này.

Ý tưởng của Mountcastle đáng kinh ngạc và sâu sắc chẳng kém gì thuyết tiến hóa của Darwin. Darwin đã đề xuất một cơ chế – hay “thuật toán” như nhiều bạn đọc ưa dùng – lý giải tính đa dạng đáng kinh ngạc của sự sống. Các sinh vật nhìn bề ngoài có thể là vô số loài động thực vật khác nhau, nhưng thực ra chỉ là những dạng biểu lộ trong thực tế của cùng một thuật toán tiến hóa cơ bản. Đến lượt mình, Mountcastle cũng đề xuất rằng mọi thứ mà ta gán cho trí tuệ tuy có vẻ khác nhau, nhưng thực ra là biểu hiện của cùng một thuật toán vỏ não cơ bản. Tôi hy vọng bạn có thể nhận ra tính bất ngờ và cách mạng trong đề xuất của Mountcastle. Darwin đề xuất rằng sự sống đa dạng nhờ một thuật toán cơ bản, Mountcastle đề xuất rằng trí tuệ cũng vậy.

Cũng giống nhiều vấn đề mang tính lịch sử, người ta vẫn đang tranh luận xem Mountcastle có phải người đầu tiên đề xuất ý tưởng này hay không. Theo kinh nghiệm của tôi, mọi ý tưởng đều có ít nhất một tiền thân nào đó. Nhưng cũng theo những gì tôi biết, Mountcastle là người đầu tiên lập luận rõ ràng và cẩn thận về một thuật toán chung của vỏ não.

Các đề xuất của Mountcastle và Darwin khác nhau ở một điểm rất thú vị. Darwin biết thuật toán *là gì*: tiến hóa dựa trên biến đổi ngẫu nhiên và chọn lọc tự nhiên. Nhưng ông không biết thuật toán nằm ở đâu trong cơ thể. Mãi đến khi ADN được phát hiện nhiều năm sau đó, người ta mới trả lời được câu hỏi này. Ngược lại, Mountcastle không biết thuật toán vỏ não *là gì* và cũng không biết đến các nguyên tắc của trí tuệ, nhưng ông biết thuật toán cơ bản nằm ở đâu trong não bộ.

Vậy Mountcastle nói gì về vị trí của thuật toán vỏ não? Ông cho rằng đơn vị cơ bản của tân vỏ não (đơn vị của trí tuệ) là cột vỏ não. Nhìn vào bề mặt của tân vỏ não, một cột vỏ não chiếm khoảng 1 mm² và toàn bộ bề dày 2,5 mm để tạo thành thể tích 2,5 mm³. Theo định nghĩa này, có khoảng 150.000 cột vỏ não xếp cạnh nhau trong tân vỏ não ở người. Bạn có thể tưởng tượng mỗi cột vỏ não như một sợi spaghetti mỏng, tân vỏ não khi đó hết như 150.000 sợi spaghetti ngắn xếp chồng lên nhau theo chiều dọc.

Chiều rộng của các cột vỏ não ở mỗi loài và mỗi vùng một khác. Ví dụ, ở chuột nhắt và chuột cống, mỗi râu ứng với một cột vỏ não; những cột này có đường kính khoảng 0,5 mm. Ở mèo, các cột thị giác có đường kính khoảng 1 mm. Chúng tôi không có nhiều dữ liệu về kích thước của các cột trong não người. Để đơn giản hóa, tôi sẽ tiếp tục giả định các cột này có diện tích 1 mm², nghĩa là mỗi chúng ta có 150.000 cột. Mặc dù con số thực tế có thể khác nhưng không quá quan trọng với những gì ta muốn làm ở đây.

Cột vỏ não không thấy được dưới kính hiển vi. Ở số ít ngoại lệ, không có ranh giới rõ ràng giữa chúng. Các nhà khoa học biết chúng tồn tại bởi mọi tế bào trong cùng một cột sẽ phản ứng với cùng một phần võng mạc hoặc cùng một mảng da, còn các tế bào trong cột khác sẽ phản ứng với phần võng mạc hoặc mảng da khác. Việc phân nhóm phản ứng như vậy xác định ranh giới giữa các cột và được tìm thấy trên toàn bộ tân vỏ não. Mountcastle chỉ ra rằng mỗi cột lại được chia thành vài trăm cột nhỏ. Nếu một cột vỏ não như một sợi spaghetti mỏng thì các cột nhỏ giống như những sợi mảnh hơn, như sợi tóc, xếp cạnh nhau tạo thành sợi spaghetti. Mỗi cột nhỏ chứa hơn 100 neuron bao trùm tất cả các lớp. Không giống như cột vỏ não, ta có thể phân định ranh giới giữa các cột nhỏ này và thấy chúng dưới kính hiển vi.

Mountcastle không đề xuất và cũng không biết về chức năng của cột vỏ não cùng các cột nhỏ. Ông chỉ đề xuất rằng các cột vỏ não đang làm cùng một việc và các cột nhỏ là một thành phần phụ quan trọng.

Hãy cùng tổng kết lại một chút. Tân vỏ não là một mảng mô có kích thước bằng một chiếc khăn ăn lớn. Nó được chia thành hàng chục khu vực với các chức năng khác nhau. Mỗi khu vực như vậy được chia thành hàng ngàn cột, mỗi cột gồm vài trăm cột nhỏ giống sợi tóc, và mỗi cột nhỏ này lại chứa hơn 100 tế bào. Mountcastle đề xuất rằng các cột và cột nhỏ của tân vỏ não có cùng một chức năng: thực thi một thuật toán cơ bản chịu trách nhiệm cho mọi khía cạnh của nhận thức và trí tuệ.

Mountcastle đưa ra đề xuất về thuật toán phổ quát dựa trên một số nhóm bằng chứng. Đầu tiên, như tôi đã đề cập, các mạch chi tiết được tìm thấy ở mọi nơi trong tân vỏ não là khá giống nhau. Khi thấy hai con chip silic có thiết kế mạch gần như giống hệt nhau, sẽ là có lý khi cho rằng các chức năng của chúng cũng gần như giống hệt nhau. Lập luận tương tự cũng đúng với các mạch chi tiết trong tân vỏ não. Thứ hai, quá trình mở rộng chính của tân vỏ não người hiện đại so với các tổ tiên trong họ người mới chỉ vài triệu năm, khá nhanh trên khung thời gian tiến hóa. Nó có thể không đủ để quá trình tiến hóa phát hiện ra nhiều năng lực phức tạp mới, nhưng vẫn đủ để tạo thêm các bản sao của cùng một thứ. Thứ ba, chức năng của các vùng trong tân vỏ não là không cố định. Ví dụ, ở những người khiếm thị bẩm sinh, vùng thị giác của tân vỏ não không nhận được thông tin hữu ích từ mắt. Vì thế, vùng này có thể đảm nhận vai trò mới liên quan đến thính giác hoặc xúc giác. Cuối cùng là lập luận về tính linh hoạt cực độ. Con người có thể làm nhiều việc dù không có áp lực tiến hóa. Ví dụ, não ta không phát triển để lập trình máy tính hay làm kem – cả hai đều là những phát minh gần đây. Khả năng làm những việc này cho ta biết não bộ hoạt động nhờ một phương pháp học tập có mục đích khái quát. Theo tôi, lập luận cuối cùng này là thuyết phục nhất. Để học được gần như mọi thứ, não bộ phải làm việc dựa trên một nguyên tắc phổ quát.

Có nhiều bằng chứng khác ủng hộ đề xuất của Mountcastle nhưng dù vậy, ý tưởng của ông vẫn gây tranh cãi

từ khi công bố đến tận ngày nay. Tôi tin rằng có hai lý do cho việc này. Một là Mountcastle không biết chức năng của cột vỏ não. Tuy đã đưa ra một tuyên bố đáng ngạc nhiên dựa trên rất nhiều bằng chứng gián tiếp, nhưng ông không đề xuất cách mà một cột vỏ não thực sự làm tất cả những điều mà ta gán cho trí tuệ. Lý do khác là nhiều người thấy đề xuất của ông kéo theo những điều khó tin. Ví dụ, sẽ khó để bạn chấp nhận rằng thị giác và ngôn ngữ về cơ bản là giống nhau. Ta không có cảm giác rằng chúng tương đồng. Trước những điều không chắc chắn này, một số nhà khoa học bác bỏ đề xuất của Mountcastle bằng cách chỉ ra rằng có sự khác biệt giữa các vùng trong tân vỏ não. Sự khác biệt là tương đối nhỏ so với những điểm tương đồng, nhưng nếu tập trung vào chúng, bạn có thể lập luận rằng các vùng của tân vỏ não không giống hệt nhau.

Đề xuất của Mountcastle giống như chén thánh của khoa học thần kinh. Dù có nghiên cứu loài vật nào hay phần nào của não thì toàn bộ giới thần kinh học đều muốn hiểu về cách vận hành của toàn thể não bộ, hay hiểu cách thức vận hành của tân vỏ não. Và để làm vậy, phải hiểu được chức năng của cột vỏ não. Tóm lại, nhiệm vụ tìm hiểu bộ não và trí tuệ của chúng ta là tìm ra chức năng cũng như cách hoạt động của các cột vỏ não. Chúng không phải bí ẩn duy nhất liên quan đến não hay tân vỏ não. Nhưng hiểu biết về cột vỏ não đến nay vẫn là miếng ghép lớn nhất và quan trọng nhất trong tổng thể vấn đề.

Năm 2005, tôi được mời thuyết trình về nghiên cứu của chúng tôi tại Đại học Johns Hopkins. Tôi nói về nhiệm vụ tìm hiểu tân vỏ não, cách chúng tôi tiếp cận vấn đề và những tiến bộ đạt được. Sau những buổi nói chuyện kiểu này, diễn giả thường gặp gỡ từng giảng viên. Trong chuyến đi này, chuyến thăm cuối cùng của tôi là với Vernon Mountcastle cùng trưởng khoa của ông. Tôi cảm thấy vinh dự khi được gặp người đã mang đến cho cuộc đời mình nhiều cảm hứng và cái nhìn sâu sắc. Sau khi dự khán, Mountcastle nói rằng tôi nên đến làm việc tại Johns Hopkins và ông sẽ thu xếp cho tôi một ghế. Đề nghị này thật bất ngờ và kỳ lạ. Tôi không thể nghiêm túc cân nhắc nó vì gia đình và các mối kinh doanh của mình ở California, nhưng nó làm tôi nhớ lại năm 1986, khi đề xuất nghiên cứu tân vỏ não của tôi bị UC Berkeley từ chối. Nếu hồi đó ông đề nghị như thế, hẳn tôi đã nhảy cẫng lên rồi.

Trước khi tạm biệt ông, tôi xin Mountcastle ký tên vào cuốn *The Mindful Brain* mà tôi đã đọc rất kỹ. Khi rời đi, tôi buồn vui lẫn lộn. Tôi rất vui khi được gặp Mountcastle và nhẹ nhõm vì được ông đánh giá cao. Tôi buồn khi biết rằng có thể mình sẽ không bao giờ gặp lại ông nữa. Ngay cả khi thành công, có thể tôi cũng không chia sẻ được với ông những gì mình tìm hiểu được và nhờ ông cho ý kiến. Khi chờ taxi, quyết tâm hoàn thành sứ mệnh của ông trào dâng trong tôi.

CHƯƠNG 3

Một mô hình thế giới trong não bộ

Với bạn, hoạt động của não bộ có vẻ hiển nhiên. Não nhận đầu vào từ các “cảm biến”, xử lý chúng rồi bắt đầu hành động. Cuối cùng, cách một con vật phản ứng với những gì nó cảm nhận được sẽ quyết định việc nó thành công hay thất bại. Việc ta hành động tức thời ngay khi nhận thông tin đầu vào là đúng với một số phần của não. Ví dụ, vô tình chạm vào một bề mặt nóng sẽ khiến ta rút tay lại. Mạch đầu vào-đầu ra chịu trách nhiệm cho việc này nằm trong tủy sống. Nhưng còn tân vỏ não thì sao? Liệu ta có thể nói rằng nhiệm vụ của tân vỏ não là lấy thông tin đầu vào từ các cảm biến rồi hành động tức thời hay không? Nói ngắn gọn là không.

Bạn đang đọc cuốn sách này và tân vỏ não không gây ra bất kỳ hành động tức thời nào ngoài lật trang hoặc chạm vào màn hình. Hàng ngàn từ đang truyền vào tân vỏ não và bạn

gần như không hành động theo chúng. Có thể sau này bạn sẽ có hành động khác vì đã đọc cuốn sách. Có lẽ bạn sẽ có những cuộc trò chuyện trong tương lai về lý thuyết não bộ và tương lai nhân loại, thứ mà bạn sẽ không có nếu không đọc những trang sách này. Có lẽ việc suy nghĩ và lựa chọn từ ngữ của bạn trong tương lai sẽ bị ảnh hưởng một cách tinh tế qua những gì tôi chia sẻ ở đây. Có lẽ bạn sẽ nghiên cứu chế tạo những cỗ máy thông minh dựa trên các nguyên tắc của não bộ và được truyền cảm hứng từ những gì tôi viết. Nhưng ngay bây giờ, bạn chỉ đang đọc mà thôi. Nếu nhất định phải mô tả tân vỏ não như một hệ thống đầu vào-đầu ra, thì điều tốt nhất có thể nói là tân vỏ não nhận được rất nhiều thông tin đầu vào, học hỏi và sau đó – có thể vài giờ hoặc hàng năm – hành động khác đi dựa trên chúng.

Kể từ khi bắt đầu quan tâm đến cách não bộ vận hành, tôi nhận ra rằng việc coi tân vỏ não như một hệ thống đầu vào-dẫn-đến-đầu-ra không mang lại hiệu quả. May thay, khi còn học cao học ở Berkeley, tôi đã có một cái nhìn sâu sắc dẫn bản thân đến một con đường khác thành công hơn. Khi đó tôi đang làm việc ở nhà, mặt bàn và khắp phòng đầy những thứ tạp nham. Tôi để ý rằng nếu bất kỳ thứ gì thay đổi dù chỉ một chút, tôi cũng sẽ nhận ra. Cốc đựng bút chì của tôi luôn ở bên phải bàn; nếu ngày nào đó thấy nó nằm bên trái, tôi sẽ nhận ra và tự hỏi làm sao nó chạy sang đó được. Tôi cũng sẽ nhận ra nếu cái dập ghim thay đổi độ dài bằng cách quan sát hoặc chạm vào nó, và cũng sẽ chú ý khi âm thanh phát ra

từ nó khác đi. Nếu đồng hồ trên tường thay đổi vị trí hoặc kiểu dáng, tôi sẽ biết. Nếu con trỏ trên màn hình máy tính di chuyển sang trái khi tôi di chuột sang phải, tôi sẽ lập tức nhận ra điều không ổn. Điều khiến tôi ấn tượng là mình sẽ nhận ra những thay đổi này ngay cả khi không chú ý vào các đồ vật. Khi nhìn quanh phòng, tôi không hỏi: “Cái dập ghim vẫn dài như cũ chứ?” và không nghĩ: “Phải kiểm tra xem kim giờ có ngắn hơn kim phút hay không.” Những điều khác thường cứ thế xuất hiện trong đầu, và tôi sẽ phải chú ý đến chúng. Thực sự, trong môi trường có hàng ngàn thay đổi mà não tôi sẽ lập tức nhận diện được.

Tôi chỉ nghĩ ra được một cách giải thích. Bộ não của tôi, cụ thể là vùng tân vỏ não, đang đưa ra nhiều dự đoán cùng lúc về những gì nó sắp nhìn thấy, nghe thấy và cảm thấy. Mỗi khi tôi di chuyển mắt, tân vỏ não sẽ dự đoán về những gì nó sắp thấy. Mỗi khi tôi nhặt một thứ gì đó, tân vỏ não sẽ dự đoán về cảm giác trên từng ngón tay. Và mọi hành động tôi thực hiện đều dẫn đến dự đoán về những gì tôi sẽ nghe thấy. Bộ não dự đoán những kích thích nhỏ nhất, như kết cấu của tay cầm trên cốc cà phê, và cả những ý tưởng lớn mang tính khái niệm như tháng nào sẽ hiện ra trên lịch. Các dự đoán này xuất hiện ở mọi phương thức cảm quan (sensory modality), từ các đặc điểm tri giác cấp thấp tới các khái niệm cấp cao, và từ đó tôi biết mọi phần của tân vỏ não – hay mọi cột vỏ não – đều đang đưa ra dự đoán. Dự đoán là một chức năng phổ quát của tân vỏ não.

Vào thời điểm đó, một số nhà khoa học thần kinh mô tả bộ não như một cỗ máy dự đoán. Tập trung vào cách tân vỏ não đưa ra nhiều dự đoán song song sẽ là một hướng đi mới lạ để nghiên cứu cách hoạt động của nó. Tôi biết tân vỏ não không chỉ đưa ra dự đoán, nhưng dựa vào việc này, ta có thể tìm ra một phương cách có tính hệ thống để giải đáp bí ẩn của các cột vỏ não. Tôi có thể đặt những câu hỏi cụ thể về cách neuron đưa ra dự đoán trong những điều kiện khác nhau, và câu trả lời có thể hé lộ chức năng của các cột vỏ não cùng cách chúng thực hiện các chức năng đó.

Để có thể dự đoán, não bộ phải học được rằng điều gì là bình thường – nghĩa là điều gì nên được mong đợi dựa trên những kinh nghiệm trong quá khứ. Cuốn *On Intelligence* (Bàn về trí tuệ) của tôi khám phá ý tưởng về học tập và dự đoán này. Trong đó, tôi dùng cụm từ “khung dự đoán trí nhớ” để mô tả ý tưởng tổng thể và viết về những gì mà việc nghĩ về não bộ theo hướng này có thể gợi ý. Tôi lập luận rằng thông qua việc nghiên cứu cách tân vỏ não đưa ra dự đoán, chúng ta có thể làm sáng tỏ cách thức hoạt động của nó.

Tôi giờ không còn dùng cụm từ “khung dự đoán trí nhớ” nữa. Thay vào đó, tôi mô tả ý tưởng tương tự bằng cách nói rằng tân vỏ não học được một mô hình về thế giới và dự đoán dựa trên mô hình đã học. Tôi thích từ “mô hình” hơn vì nó mô tả chính xác hơn loại thông tin mà tân vỏ não học được. Ví dụ, não tôi có mô hình của cái dập ghim gồm hình dạng của cái dập ghim, cảm giác về nó và âm thanh nó phát ra khi

sử dụng. Mô hình thế giới của bộ não gồm vị trí của các đối tượng và cách chúng thay đổi khi tương tác với ta. Ví dụ, mô hình cái dập ghim của tôi cũng chứa cách di chuyển của phần trên so với phần dưới và cách những chiếc ghim được đẩy ra khi bấm. Bạn đã học nó vào một thời điểm nào đó trong đời và kiến thức này đã được lưu vào tân vỏ não rồi.

Bộ não tạo ra một mô hình dự đoán, điều này chỉ có nghĩa là nó sẽ luôn dự đoán về thông tin đầu vào. Dự đoán không phải thứ bộ não thi thoảng mới làm, mà là một đặc tính nội tại không ngừng vận động, đóng vai trò thiết yếu trong việc học. Khi những dự đoán của não được xác minh thì mô hình thế giới của bộ não là chính xác. Khi có một dự đoán sai, bạn sẽ tập trung vào lỗi sai đó và cập nhật mô hình.

Chúng ta không nhận thức được phần lớn các dự đoán này trừ phi não nhận được thông tin đầu vào không khớp. Khi tình cờ với lấy cốc cà phê, tôi không biết não đang dự đoán cảm giác trên từng ngón tay cùng khối lượng, nhiệt độ của cái cốc và âm thanh mà cái cốc sẽ tạo ra khi tôi đặt nó lên bàn. Nhưng nếu cốc đột nhiên nặng hơn, lạnh hơn hoặc kêu cọt két, tôi sẽ nhận thấy sự thay đổi. Chúng ta có thể chắc chắn rằng những dự đoán này đang xảy ra bởi một thay đổi nhỏ trong thông tin đầu vào bất kỳ cũng sẽ được chú ý. Nhưng khi một dự đoán là chính xác (hầu hết là vậy), chúng ta sẽ không biết não từng có nó.

Khi bạn được sinh ra, tân vỏ não hầu như không biết gì. Nó không biết bất kỳ từ nào, không biết một ngôi nhà

trông như thế nào, không biết cách sử dụng máy tính và cũng không biết cánh cửa là gì, xoay trên bản lề ra sao. Nó phải học vô số thứ. Cấu trúc tổng thể của tân vỏ não không phải là ngẫu nhiên. Kích thước của nó, số vùng nó có và cách các vùng này kết nối phần lớn là do gen quyết định. Ví dụ, gen xác định phần tân vỏ não nào nối với mắt, phần nào nối với tai và cách các phần đó kết nối với nhau. Do đó, chúng ta có thể nói rằng tân vỏ não được cấu trúc ngay từ khi mới sinh để nhìn, nghe và thậm chí học ngôn ngữ. Nhưng cũng đúng là tân vỏ não không biết nó sẽ nhìn thấy gì, nghe thấy gì và có thể học những ngôn ngữ cụ thể nào. Chúng ta có thể coi là tân vỏ não ra đời với một số giả định tích hợp sẵn về thế giới nhưng không biết gì cụ thể. Qua kinh nghiệm, nó học được một mô hình phong phú và phức tạp về thế giới.

Tân vỏ não học vô khối thứ. Tôi đang ngồi trong một căn phòng với hàng trăm đồ vật. Tôi sẽ chọn ngẫu nhiên một thứ: máy in. Tôi đã học được một mô hình về máy in gồm việc nó có một khay giấy và cách khay này đi ra đi vào. Tôi biết cách đổi khổ giấy và bóc một ram giấy mới rồi đặt vào khay. Tôi biết các bước xử lý kẹt giấy. Tôi biết dây nguồn có phích hình chữ D và chỉ có thể cắm theo một hướng. Tôi biết máy in phát ra âm thanh gì, và cũng biết nó sẽ phát ra âm thanh khác nếu in hai mặt chứ không phải một. Phòng tôi cũng có một tủ tài liệu nhỏ hai ngăn. Tôi có thể nhớ lại cả chục thứ mình biết về nó, mỗi ngăn có những gì và sắp xếp ra sao. Tôi biết có một ổ khóa, chỗ cất chìa khóa, cách cắm và xoay chìa để khóa tủ.

Tôi biết cảm giác khi chạm vào chìa và ổ khóa cũng như âm thanh mà chúng tạo ra khi dùng. Trên chiếc chìa khóa có một khuyên đeo và tôi biết dùng móng tay cạy cái khuyên để thêm bột chìa.

Hãy tưởng tượng bạn đang đi từ phòng này sang phòng khác trong nhà mình. Trong mỗi phòng, bạn có thể nghĩ ra hàng trăm thứ và với mỗi thứ, bạn có thể theo dõi một loạt những điều đã học. Bạn cũng có thể làm tương tự với thị trấn mình đang sống: nhớ lại những tòa nhà, công viên, khu để xe đạp và từng cái cây ở từng địa điểm. Với mỗi đối tượng, bạn có thể nhớ lại những trải nghiệm liên quan và cách tương tác với đối tượng đó. Bạn biết rất nhiều điều và những liên kết hữu quan của kiến thức xem ra là vô kể.

Chúng ta cũng học được nhiều khái niệm cấp cao. Theo ước tính, mỗi người biết khoảng 40.000 từ. Chúng ta có khả năng học tiếng nói, chữ viết, ngôn ngữ ký hiệu, ký hiệu toán học và ký hiệu âm nhạc. Chúng ta tìm hiểu cách hoạt động của các biểu mẫu điện tử, chức năng của máy điều nhiệt và cả ý nghĩa của khái niệm sự đồng cảm hay dân chủ, mặc dù cách hiểu về những điều này có thể khác nhau. Độc lập với những thứ khác mà tân vỏ não có thể làm, chúng ta chắc chắn rằng tân vỏ não học một mô hình vô cùng phức tạp về thế giới. Mô hình này là cơ sở để chúng ta dự đoán, tri nhận và hành động.

Học tập thông qua chuyển động

Thông tin đầu vào cho não luôn thay đổi. Có hai lý do cho việc này. Thứ nhất, thế giới quanh ta luôn thay đổi. Ví dụ, khi nghe nhạc, đầu vào từ tai thay đổi nhanh chóng, phản ánh những dịch chuyển trong bài nhạc. Tương tự, một cái cây đứng đưa trong gió sẽ dẫn đến những thay đổi về thị giác và cả thính giác. Trong hai ví dụ này, thông tin đầu vào của não thay đổi theo từng thời điểm, không phải vì bạn đang chuyển động mà vì mọi thứ xung quanh đang chuyển động và tự thay đổi.

Lý do thứ hai là vì chúng ta di chuyển. Mỗi khi ta bước một bước, di chuyển một chi, đảo mắt, nghiêng đầu hoặc tạo ra âm thanh, đầu vào từ các giác quan sẽ thay đổi. Ví dụ, mắt thực hiện các chuyển động nhanh (hay saccade) khoảng ba lần/giây. Với mỗi chuyển động saccade, mắt tập trung vào một điểm mới trong môi trường và thông tin từ mắt đến não thay đổi hoàn toàn. Thay đổi này sẽ không xảy ra nếu ta không đảo mắt.

Bộ não học mô hình về thế giới bằng cách quan sát cách thông tin đầu vào thay đổi theo thời gian, ngoài ra không còn cách nào khác. Khác với máy tính, chúng ta không thể tải tệp lên não. Cách duy nhất để não học bất cứ điều gì là thông qua thay đổi trong đầu vào. Nếu đầu vào không thay đổi, não sẽ không học được gì.

Một số thứ như giai điệu có thể học được mà không cần cử động cơ thể. Chúng ta có thể ngồi yên, nhắm mắt và học một giai điệu mới chỉ bằng cách lắng nghe âm thanh thay

đổi theo thời gian. Nhưng hầu hết việc học đòi hỏi tích cực di chuyển và khám phá. Hãy tưởng tượng bạn bước vào một ngôi nhà mình chưa từng vào. Nếu không di chuyển, sẽ không có thay đổi nào trong thông tin đầu vào và bạn không thể học được bất cứ gì về ngôi nhà. Để học được mô hình về ngôi nhà, bạn phải quan sát khắp nơi và đi từ phòng này sang phòng khác. Bạn cần mở cửa, nhìn vào ngăn kéo và nhặt nhạnh đồ vật. Ngôi nhà và mọi thứ bên trong chủ yếu là tĩnh; chúng không tự di chuyển. Để tìm hiểu mô hình của một ngôi nhà, chính bạn phải di chuyển.

Hãy xét một đồ vật đơn giản như chuột máy tính. Để biết cảm giác khi chạm vào chuột, bạn phải lướt ngón tay trên nó. Để biết hình thù của chuột, bạn phải nhìn nó từ các góc độ và quan sát kỹ các vị trí khác nhau. Để tìm hiểu hoạt động của chuột, bạn phải nhấp chuột, mở nắp pin hoặc di chuột để xem, cảm nhận và nghe những gì đang diễn ra.

Hoạt động này được gọi là học tập cảm giác-vận động. Nói cách khác, bộ não học mô hình về thế giới qua cách thông tin cảm giác đầu vào thay đổi khi ta di chuyển. Chúng ta có thể học một bài hát mà không cần di chuyển bởi khác với việc đi từ phòng này sang phòng khác, thứ tự nốt nhạc trong một bài hát là cố định. Nhưng hầu hết thế giới không như vậy; đa phần, chúng ta phải di chuyển để khám phá cấu trúc của đồ vật, địa điểm và hành động. Trong học tập cảm giác-vận động, khác với giai điệu, thứ tự của các cảm giác không cố định. Những gì tôi thấy khi bước vào một căn phòng phụ thuộc vào

đầu tôi quay hướng nào. Cảm giác trên ngón tay khi cầm cốc cà phê phụ thuộc vào việc tôi di chuyển ngón tay lên xuống hay sang trái, sang phải.

Với mỗi chuyển động, tân vỏ não sẽ dự đoán cảm giác tiếp theo. Nếu rà ngón tay từ dưới lên với cốc cà phê, tôi sẽ cảm nhận được miệng cốc, còn nếu di chuyển ngón tay sang trái, sang phải, tôi sẽ cảm nhận được quai cốc. Nếu quay đầu sang trái khi vào bếp, tôi sẽ thấy tủ lạnh, còn nếu quay đầu sang phải, tôi sẽ thấy khu nấu ăn. Nếu liếc mắt sang mâm bếp phía trước bên trái, tôi sẽ thấy bộ phận đánh lửa bị hỏng và cần sửa. Nếu bất kỳ thông tin đầu vào nào không khớp với dự đoán của bộ não – hẳn vợ tôi đã sửa bộ phận đánh lửa – thì tôi sẽ chú ý đến khu vực dự đoán sai. Điều này cảnh báo tân vỏ não rằng mô hình của nó về phần đó của thế giới cần được cập nhật.

Câu hỏi về cách thức hoạt động của tân vỏ não giờ có thể được đặt chính xác hơn như sau: *Làm thế nào để tân vỏ não, bao gồm hàng ngàn cột vỏ não gắn như giống hạt nhau, học được một mô hình dự đoán về thế giới thông qua chuyển động?*

Đây là câu hỏi mà tôi và nhóm của mình cần trả lời. Chúng tôi tin rằng nếu trả lời được nó, ta có thể áp dụng kỹ thuật đảo ngược (reverse engineer) với tân vỏ não. Chúng ta sẽ hiểu chức năng mà tân vỏ não đang thực hiện và cả cách nó thực hiện chức năng đó. Và cuối cùng, chúng ta sẽ có thể chế tạo những cỗ máy hoạt động theo cách tương tự.

Hai nguyên lý của khoa học thần kinh

Trước khi bắt đầu trả lời câu hỏi trên, có một số ý tưởng cơ bản khác cần biết. Đầu tiên, như mọi bộ phận khác của cơ thể, não được cấu tạo từ tế bào. Các tế bào não được gọi là neuron và giống mọi loại tế bào khác về nhiều mặt. Ví dụ, một neuron có màng tế bào xác định ranh giới và một nhân chứa ADN. Tuy nhiên, neuron có một số đặc tính độc đáo không tồn tại trong các tế bào khác của cơ thể.

Thứ nhất, neuron trông như cây cối. Phần mở rộng của màng neuron trông giống cành nhánh, gọi là sợi trục và sợi nhánh. Sợi nhánh được nhóm lại gần neuron, có chức năng thu thập thông tin đầu vào; sợi trục thì xử lý đầu ra. Nó tạo rất nhiều kết nối với các neuron lân cận, nhưng thường di chuyển rất xa, như từ bên này sang bên kia của não hoặc từ tân vỏ não xuống tận tủy sống.

Điểm khác biệt thứ hai là neuron phát ra các “gai” (spike), hay còn gọi là điện thế hoạt động (action potential). Điện thế hoạt động là một tín hiệu điện khởi phát gần thân neuron và đi dọc sợi trục đến cuối mỗi nhánh.

Điểm độc đáo thứ ba là sợi trục của một neuron tạo kết nối với các sợi nhánh của các neuron khác; các điểm kết nối được gọi là synapse (hay khớp thần kinh). Khi một gai di chuyển dọc theo sợi trục đến synapse, nó sẽ giải phóng một chất vào sợi nhánh của neuron tiếp nhận. Tùy vào loại hóa chất được giải phóng, neuron tiếp nhận có thể cũng sẽ giải phóng gai hoặc không.

Xem xét cách các neuron hoạt động, chúng ta có thể nêu ra hai nguyên lý cơ bản, đóng vai trò quan trọng trong hiểu biết của chúng ta về bộ não và trí tuệ.

Nguyên lý 1: Suy nghĩ, ý tưởng và tri giác là hoạt động của các neuron

Bất kỳ thời điểm nào ở tân vỏ não đều có các neuron tích cực giải phóng gai và cả các neuron không làm thế. Thông thường, số lượng neuron hoạt động cùng lúc là nhỏ, có thể là 2%. Suy nghĩ và nhận thức của bạn được xác định bởi các neuron đang giải phóng gai. Ví dụ, khi phẫu thuật não, đôi khi các bác sĩ cần kích hoạt neuron trong não bệnh nhân đang tỉnh táo. Họ cắm một đầu dò cực nhỏ vào tân vỏ não và kích hoạt vài neuron bằng điện. Khi họ làm vậy, bệnh nhân có thể nghe, nhìn hoặc nghĩ một điều gì đó. Khi bác sĩ ngừng kích thích, bất cứ điều gì bệnh nhân đang trải qua đều dừng lại. Nếu bác sĩ kích hoạt các neuron khác, bệnh nhân sẽ có suy nghĩ hoặc nhận thức khác.

Suy nghĩ và trải nghiệm luôn là kết quả của một tập hợp neuron hoạt động cùng lúc. Các neuron riêng lẻ có thể tham gia vào nhiều suy nghĩ hoặc trải nghiệm khác nhau. Mọi suy nghĩ bạn có đều là hoạt động của các neuron. Mọi thứ bạn thấy, nghe hoặc cảm nhận cũng vậy. Trạng thái tinh thần của chúng ta và hoạt động của các neuron là một.

Nguyên lý 2: Mọi thứ chúng ta biết đều được lưu trong các kết nối neuron

Bộ não ghi nhớ rất nhiều thứ. Bạn có những ký ức vĩnh viễn như nơi bạn lớn lên, và cũng có những ký ức tạm thời như những gì ăn tối hôm trước. Ngoài ra, bạn có kiến thức cơ bản như cách mở cửa hoặc cách đánh vần từ “từ điển”; tất cả được lưu trữ bằng cách sử dụng các synapse – kết nối giữa các neuron.

Đây là ý tưởng cơ bản về cách học của não: Mỗi neuron có hàng ngàn synapse kết nối nó với hàng ngàn neuron khác. Việc hai neuron cùng lúc giải phóng gai sẽ tăng cường kết nối giữa chúng. Khi ta học một điều, các kết nối được củng cố; còn khi quên đi một điều, các kết nối sẽ yếu đi. Ý tưởng cơ bản này do Donald Hebb đề xuất vào những năm 1940, ngày nay được gọi là học tập kiểu Hebb (Hebbian learning).

Trong nhiều năm, người ta tin rằng kết nối giữa các neuron trong não người trưởng thành là cố định và học tập liên quan đến việc tăng hoặc giảm sức bền của các synapse. Đây vẫn là cách mà việc học diễn ra trong hầu hết các mạng thần kinh nhân tạo.

Tuy nhiên, trong vài thập kỷ qua, các nhà khoa học đã phát hiện ra rằng có các synapse mới hình thành và các synapse cũ mất đi ở nhiều phần của não (gồm cả tân vỏ não). Mỗi ngày, trên một neuron đơn lẻ có rất nhiều synapse biến mất và các synapse mới xuất hiện để thay thế chúng. Do đó,

phần lớn việc học xảy ra bằng cách hình thành kết nối mới giữa các neuron mà trước đây không kết nối với nhau. Chúng ta sẽ quên kiến thức khi các kết nối cũ hoặc không sử dụng hoàn toàn biến mất.

Các kết nối trong não lưu trữ mô hình thế giới mà ta đã học được qua kinh nghiệm. Mỗi ngày, chúng ta trải nghiệm những điều mới và bổ sung kiến thức mới vào mô hình bằng cách hình thành các synapse mới. Các neuron hoạt động tại một thời điểm bất kỳ đại diện cho những suy nghĩ và nhận thức của chúng ta vào thời điểm đó.

Ta đã xem xét một số đơn vị cơ bản của tân vỏ não – vài mảnh ghép trong bức tranh toàn cảnh. Trong chương tiếp theo, chúng ta sẽ ghép chúng lại để làm hé lộ cách hoạt động của toàn bộ tân vỏ não.