بسم الله الرحمن الرحیم

هوش مصنوعی

درس ۱

مقدمه‌ای بر هوش مصنوعی

An Introduction to Artificial Intelligence

کاظم فولادی قلعه

دانشکده مهندسی، دانشکدگان فارابی

دانشگاه تهران

http://courses.fouladi.ir/ai

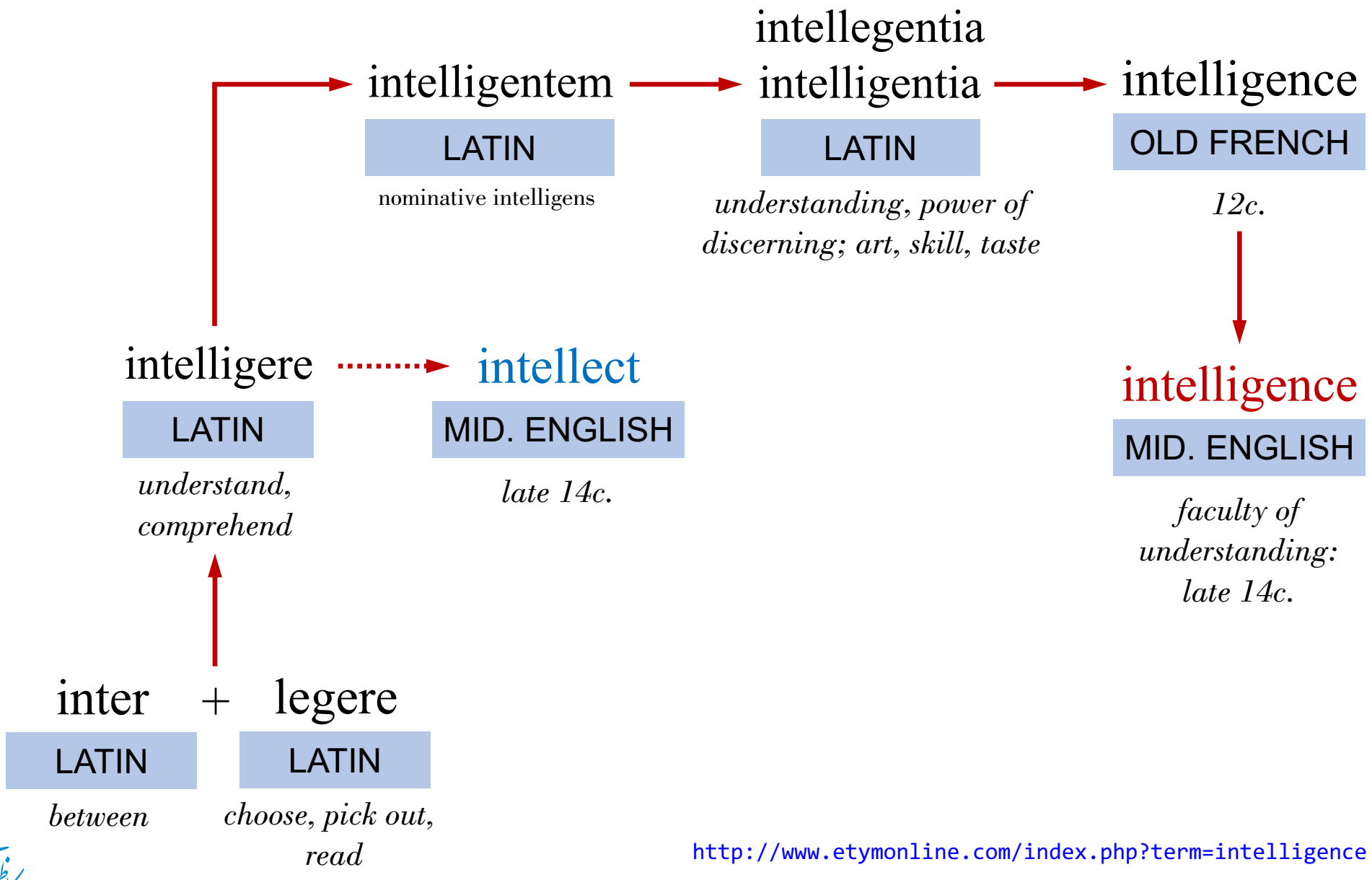
هوش مصنوعی

مقدمه‌ای بر هوش مصنوعی

◆

معرفی موضوع

اتیمولوژی: اینتلیجنس



http://www.etymonline.com/index.php?term=intelligence

ترمینولوژی: اینتلیجنس

Full Definition of *INTELLIGENCE*

1

a (1) : the ability to learn or understand or to deal with new or trying situations : reason; *also* : the skilled use of reason *(2)* : the ability to apply knowledge to manipulate one's environment or to think abstractly as measured by objective criteria (as tests)

b Christian Science : the basic eternal quality of divine Mind

c : mental acuteness : shrewdness

2

a : an intelligent entity; *especially* : angel

b : intelligent minds or mind <cosmic *intelligence*>

3

: the act of understanding : comprehension

4

a : information, news

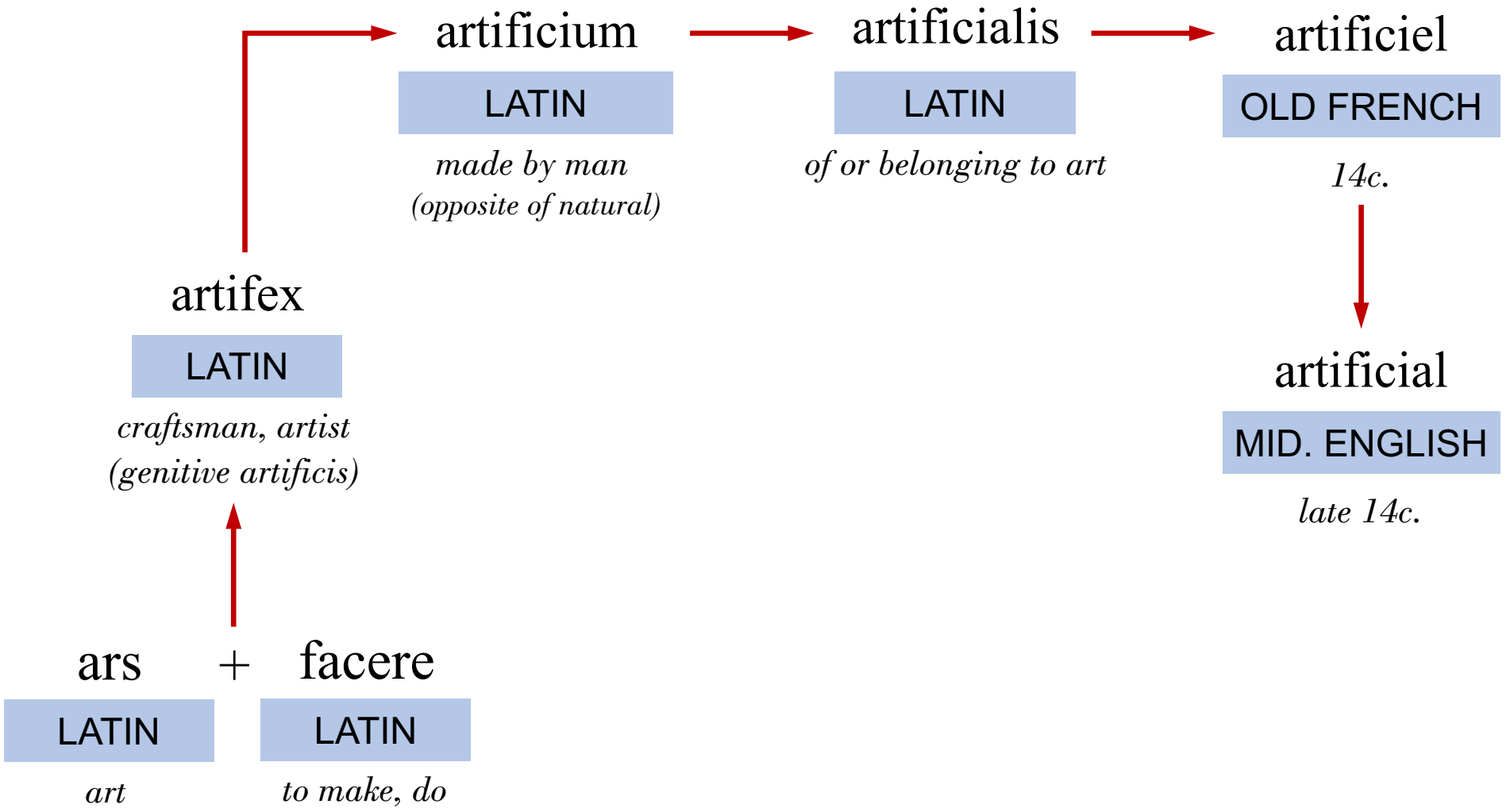
b : information concerning an enemy or possible enemy or an area; *also* : an agency engaged in obtaining such information

5

: the ability to perform computer functions

http://www.merriam-webster.com/dictionary/intelligence

اتیمولوژی: آرتیفیشال



http://www.etymonline.com/index.php?term=artificial

ترمینولوژی: آرتیفیشال

Full Definition of *ARTIFICIAL*

1
: humanly contrived often on a natural model : man-made <an *artificial* limb> <*artificial* diamonds>

2
a : having existence in legal, economic, or political theory
b : caused or produced by a human and especially social or political agency <an *artificial* price advantage> <*artificial* barriers of discrimination — R. C. Weaver>

3
obsolete : artful, cunning

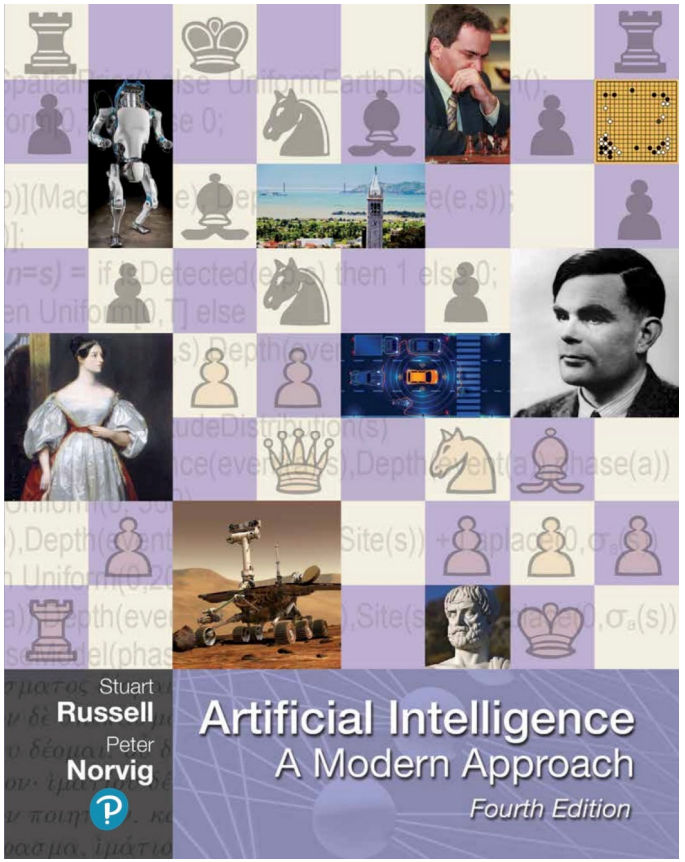
4
a : lacking in natural or spontaneous quality <an *artificial* smile> <an *artificial* excitement>
b : imitation, sham <*artificial* flavor>

5
: based on differential morphological characters not necessarily indicative of natural relationships <an *artificial* key for plant identification>

http://www.merriam-webster.com/dictionary/artificial

کتاب درس

هوش مصنوعی: رویکردی نوین

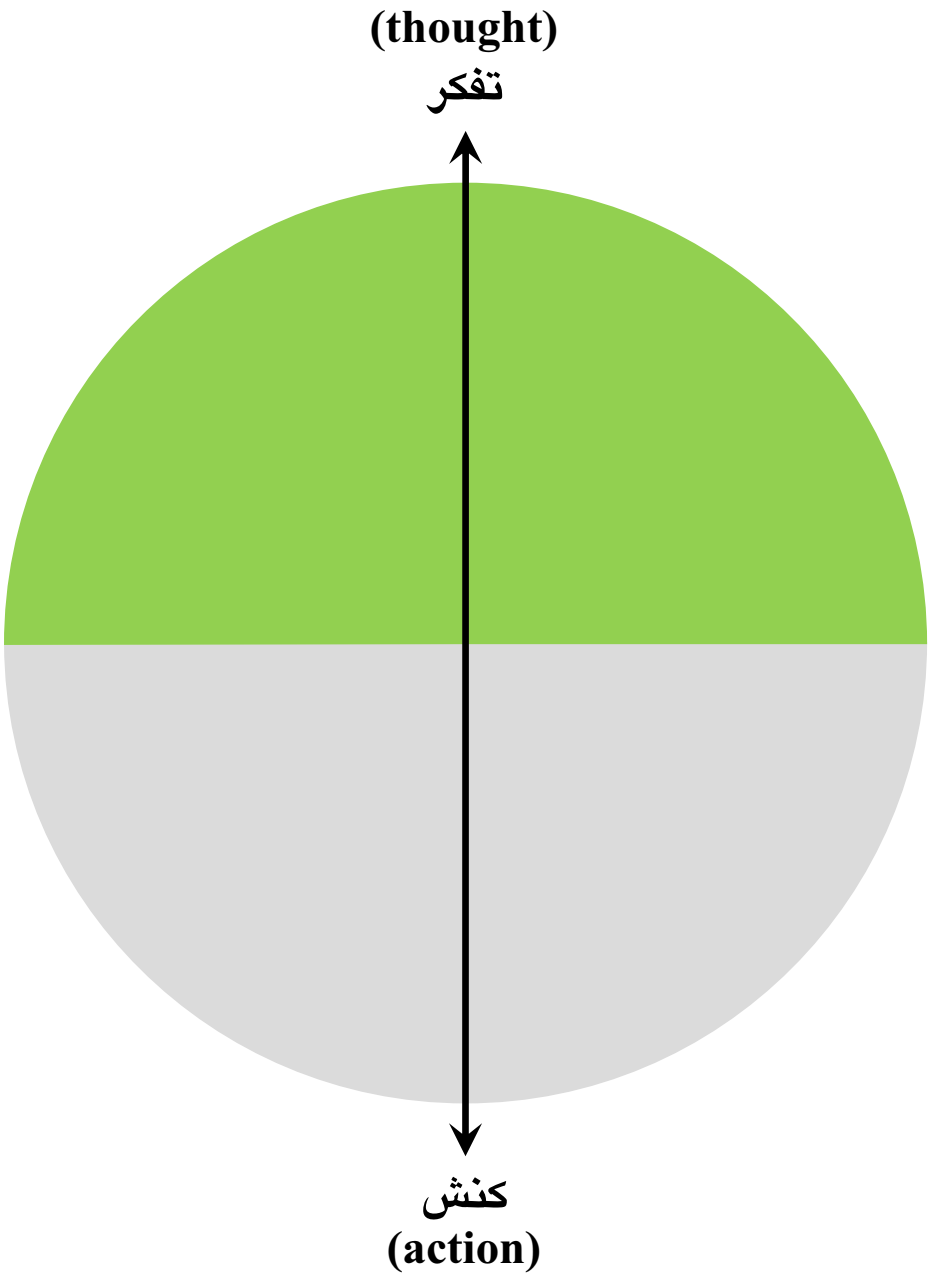


Stuart Russell and Peter Norvig,
Artificial Intelligence: A Modern Approach,
4th Edition, Prentice Hall, 2020.

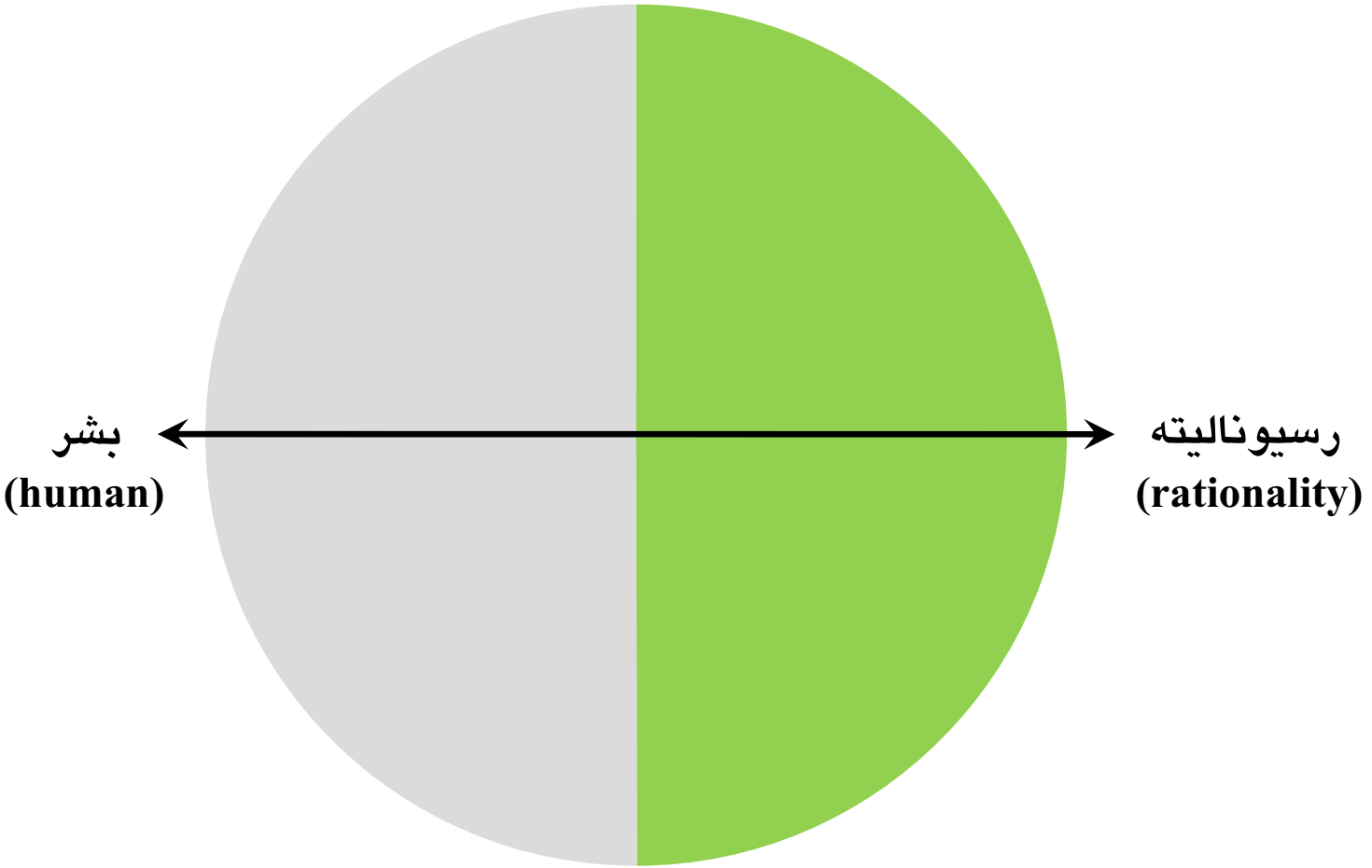
۱

هوش مصنوعی چیست؟

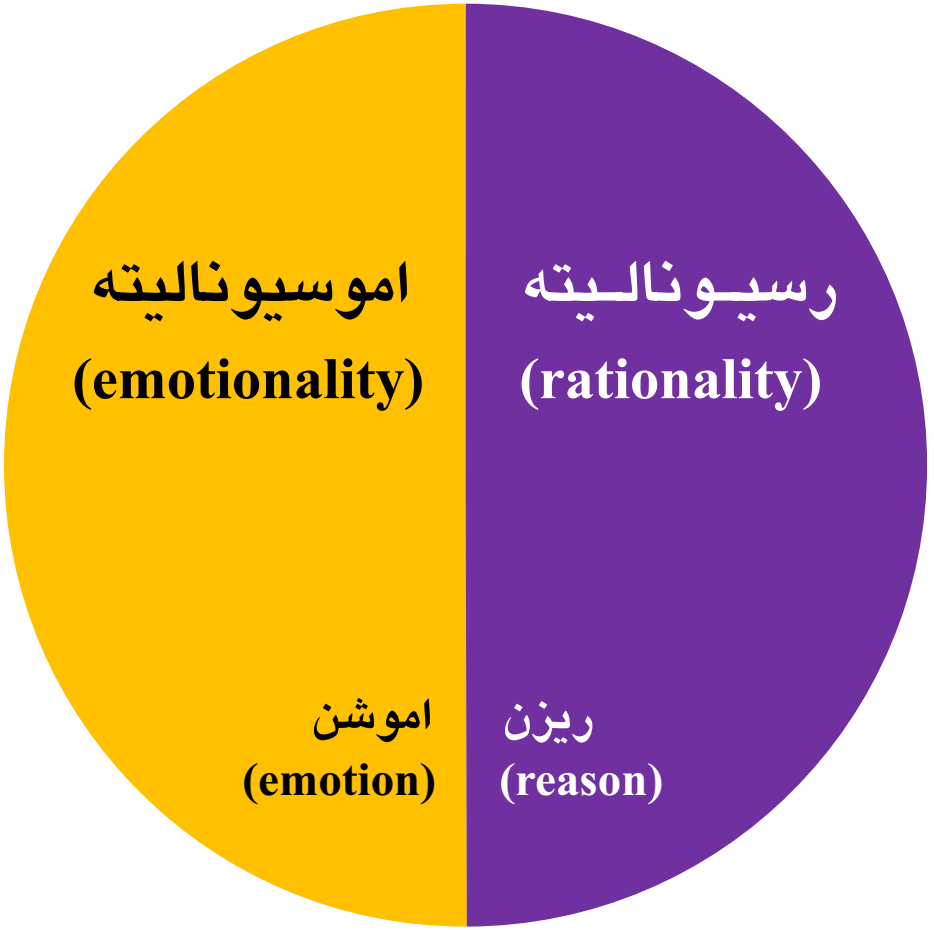
هوش: در تفکر یا کنش



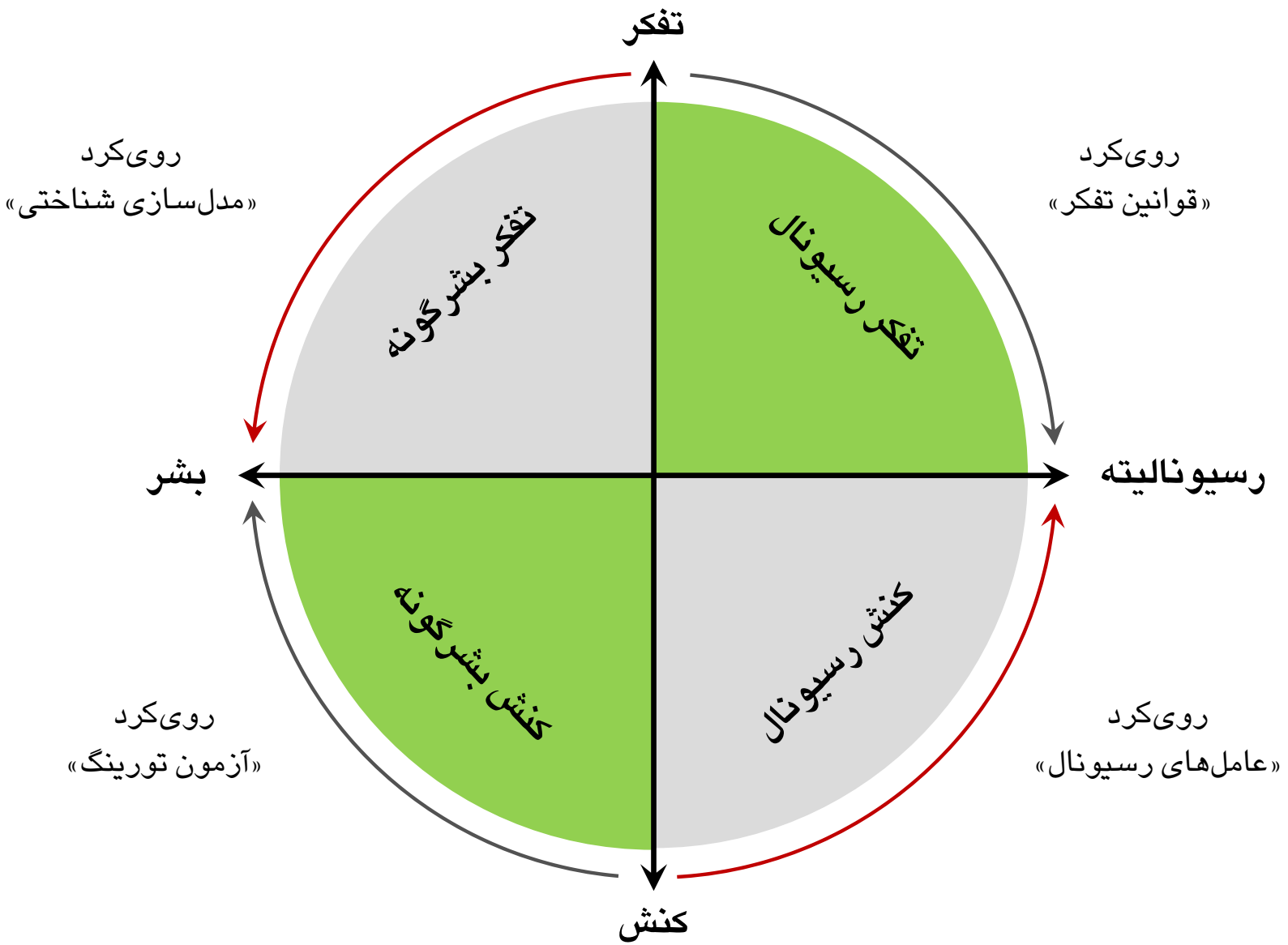
هوش: ایده‌آل هوشمندی (رسیونالیته) یا بشر



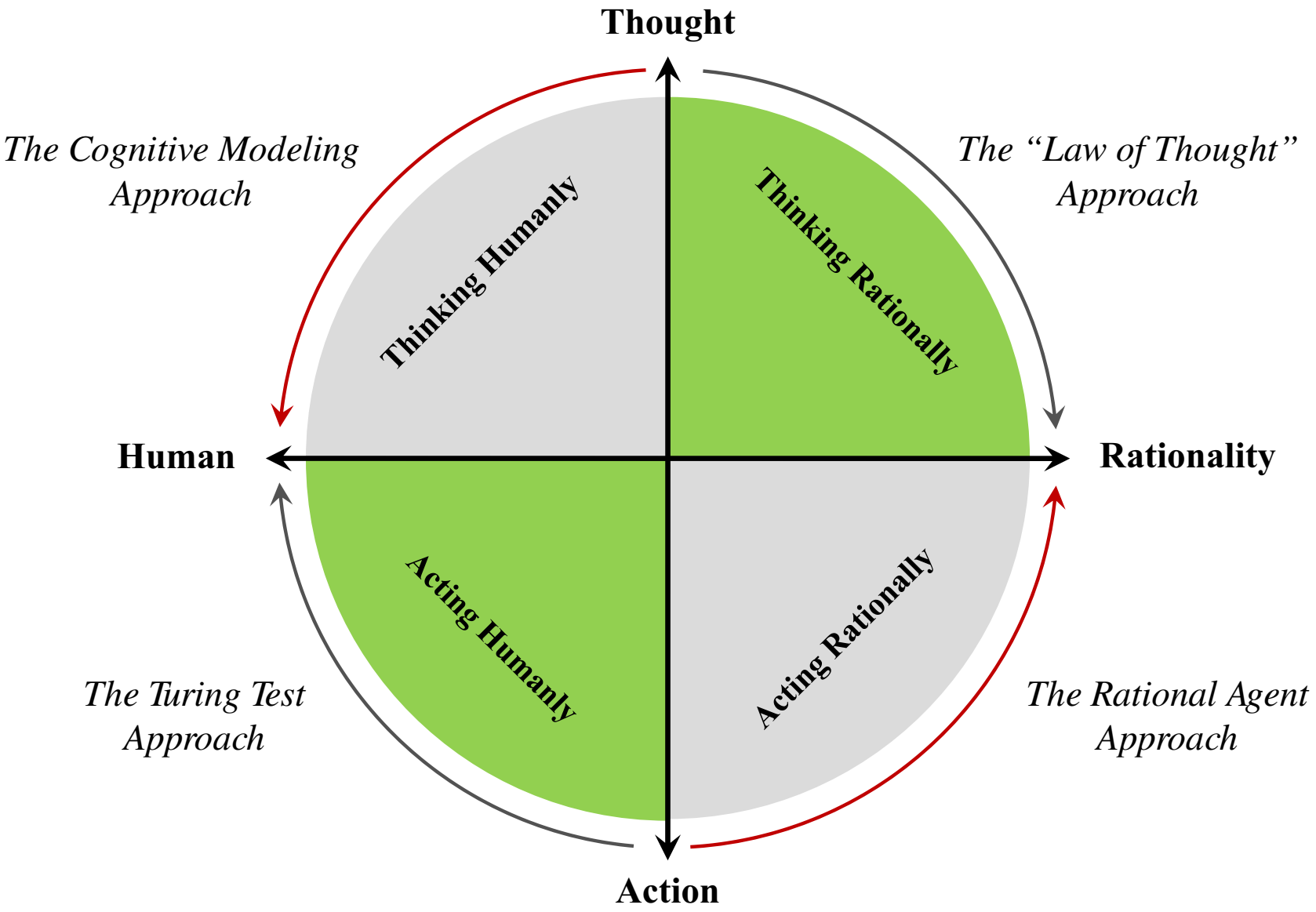
ساحات بشر در سایکولوژی



روی‌کردهای تعریف هوش مصنوعی



روی‌کردهای تعریف هوش مصنوعی

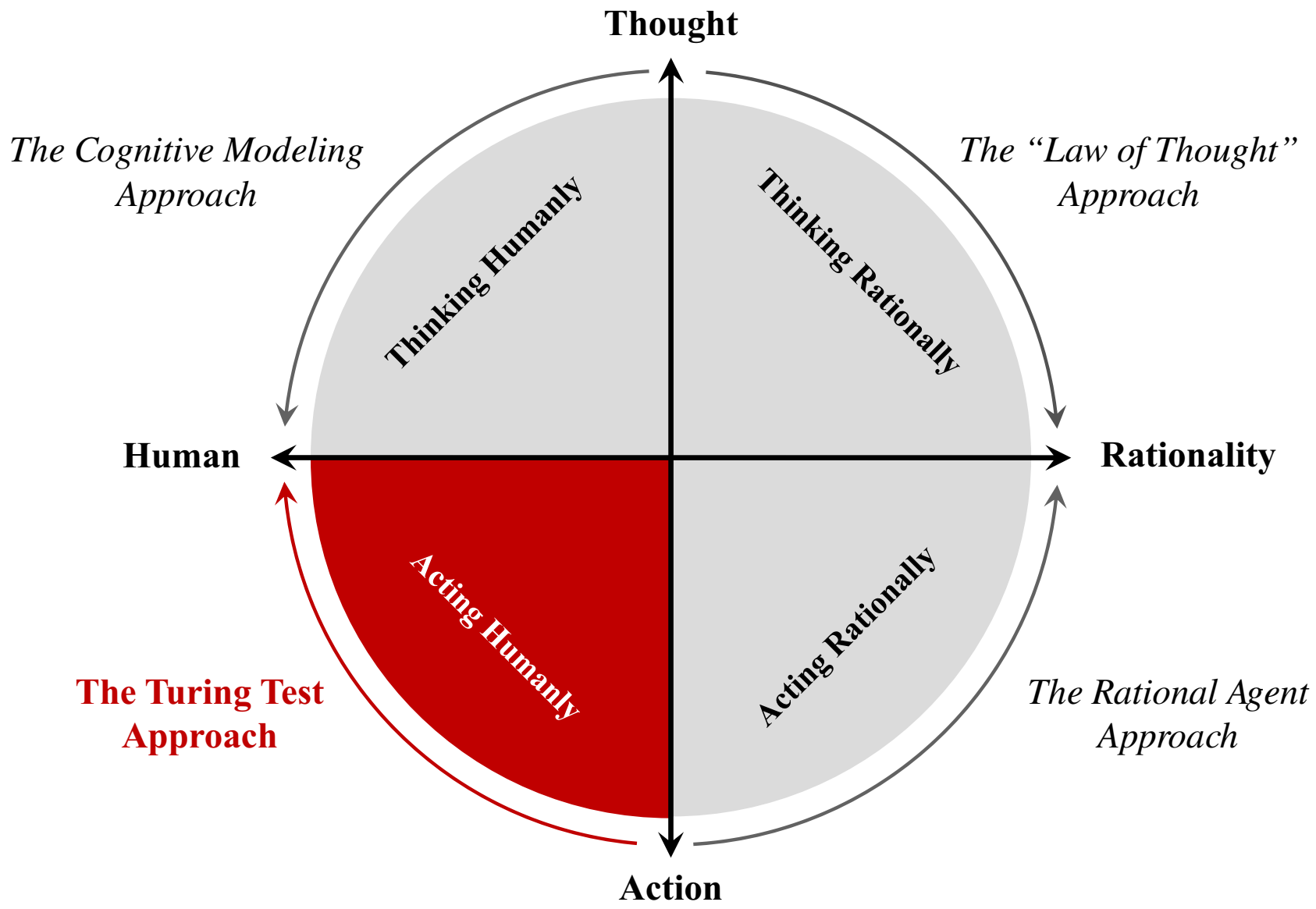


Thinking Humanly	**Thinking Rationally**
"The exciting new effort to make computers think . . . *machines with minds*, in the full and literal sense." (Haugeland, 1985)	"The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985)
"[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman, 1978)	"The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Acting Humanly	**Acting Rationally**
"The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990)	"Computational Intelligence is the study of the design of intelligent agents." (Poole *et al.*, 1998)
"The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	"AI . . . is concerned with intelligent behavior in artifacts." (Nilsson, 1998)

Figure 1.1 Some definitions of artificial intelligence, organized into four categories.

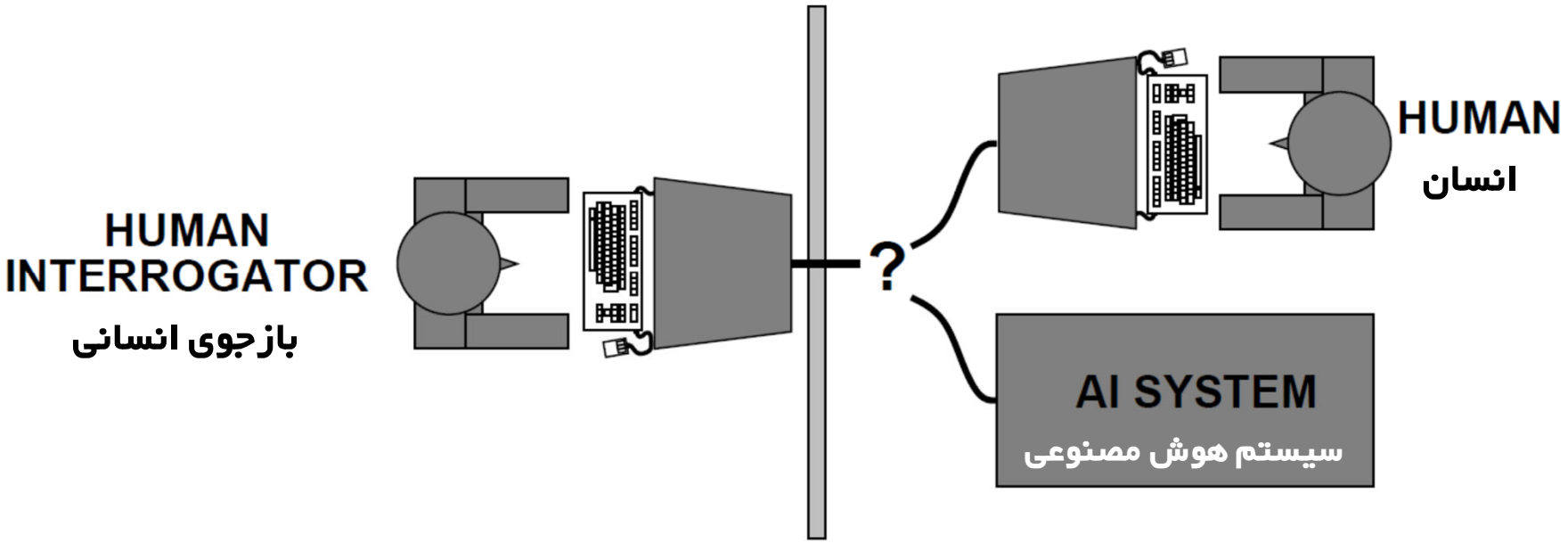
روی‌کردهای تعریف هوش مصنوعی

(۱) روی‌کرد «آزمون تورینگ» / کنش بشرگونه



(۱) رویکرد «آزمون تورینگ» در تعریف هوش مصنوعی

کنش بشرگونه



Turing (1950) "Computing machinery and intelligence":

- "Can machines think?" → "Can machines behave intelligently?"
- Operational test for intelligent behavior: THE IMITATION GAME

ظرفیت‌های لازم برای موفقیت در آزمون تورینگ

TURING TEST

ظرفیت‌های لازم برای آزمون تورینگ			
یادگیری ماشینی *Machine Learning*	استدلال خودکار *Automated Reasoning*	بازنمایی دانایی *Knowledge Representation*	پردازش زبان طبیعی *Natural Language*
یادگیری	**استدلال**	**دانایی**	**درک زبان**
برای وفق‌یابی با شرایط جدید و تشخیص و برون‌یابی الگوها	برای استفاده از اطلاعات ذخیره شده برای پاسخ به پرسش‌ها و استخراج نتایج جدید	برای ذخیره‌سازی آنچه می‌داند یا می‌شنود	برای برقراری ارتباط مؤثر با زبان طبیعی

ظرفیت‌های لازم برای موفقیت در آزمون تورینگ تام

[TOTAL TURING TEST]

ظرفیت‌های لازم برای آزمون تورینگ تام

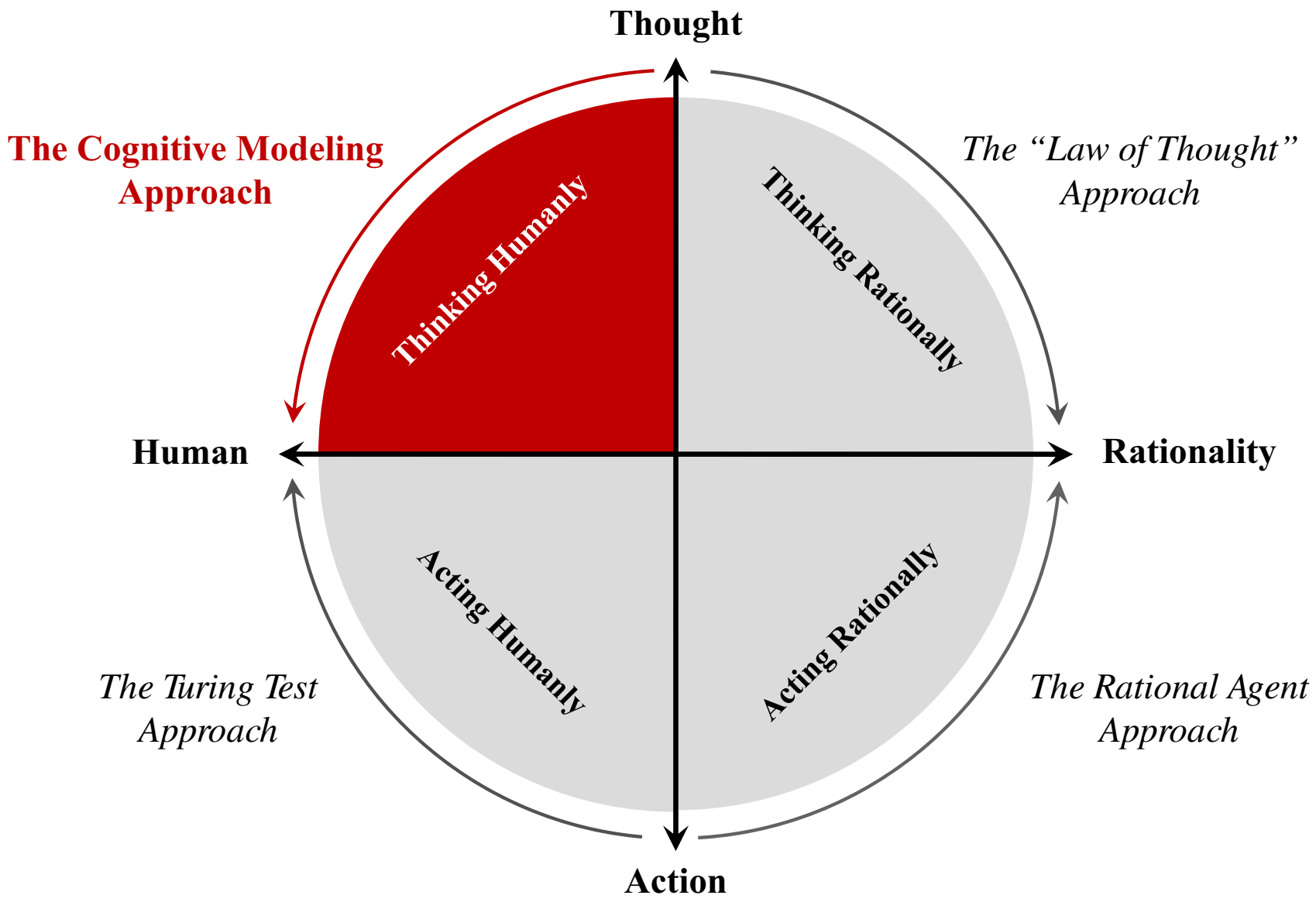
ظرفیت‌های لازم برای آزمون تورینگ

رباتیک *Robotics*	بینایی کامپیوتری *Computer Vision*	یادگیری ماشینی *Machine Learning*	استدلال خودکار *Automated Reasoning*	بازنمایی دانایی *Knowledge Representation*	پردازش زبان طبیعی *Natural Language*
کنش	بینایی	یادگیری	استدلال	دانایی	درک زبان
برای دستکاری اشیا و حرکت در محیط	برای ادراک اشیا	برای وفق‌یابی با شرایط جدید و تشخیص و برون‌یابی الگوها	برای استفاده از اطلاعات ذخیره شده برای پاسخ به پرسش‌ها و استخراج نتایج جدید	برای ذخیره‌سازی آنچه می‌داند یا می‌شنود	برای برقراری ارتباط مؤثر با زبان طبیعی

A **CAPTCHA** (an acronym for "**C**ompletely **A**utomated **P**ublic Turing test to tell **C**omputers and **H**umans **A**part") is a type of challenge-response test used in computing to determine whether or not the user is human.

روی‌کردهای تعریف هوش مصنوعی

(۲) روی‌کرد «مدل‌سازی شناختی» / تفکر بشرگونه



(۲) روی‌کرد «مدل‌سازی شناختی» در تعریف هوش مصنوعی

تفکر بشرگونه

انسان چگونه فکر می‌کند؟

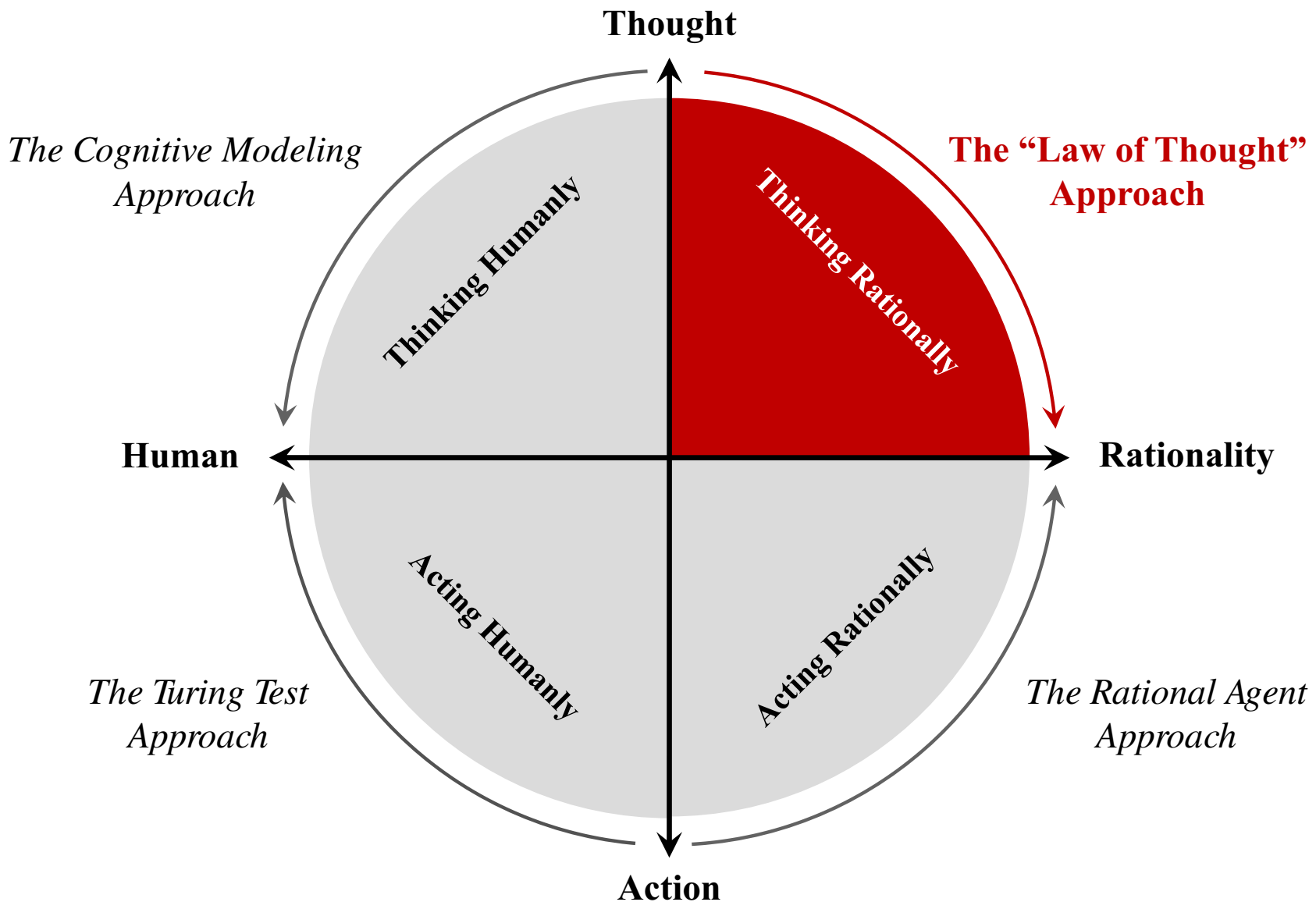
تئوری‌های علمی در مورد فعالیت‌های داخلی مغز و ذهن

بالا به پایین	پایین به بالا
پیش‌بینی و آزمون رفتار سوژه‌ی انسانی	**شناسایی مستقیم داده‌های نورولوژیکی**
علوم شناختی *Cognitive Science*	**علوم اعصاب شناختی** *Cognitive Neuroscience*
سطح انتزاع: **دانایی** *knowledge*	سطح انتزاع: **مدار** *circuit*
شروع از ذهن	شروع از مغز

تبدیل نظریه به برنامه‌ی کامپیوتری
که شبیه انسان فکر می‌کند.

روی‌کردهای تعریف هوش مصنوعی

(۳) روی‌کرد «قوانین تفکر» / تفکر رسیونال



Aristotle (384-322 B.C.)

(۳) روی‌کرد «قوانین تفکر» در تعریف هوش مصنوعی

تفکر رسیونال

منطق *(logic)*

تفکر درست: فرآیند استدلال انکارناپذیر

قیاس صوری *(syllogism)* ← الگوهایی برای ساختار مباحثه

منطق ارسطوئی

↓

منطق ریاضی (قرن نوزدهم)

نمادگذاری و قواعد استخراج افکار

مسیر مستقیم میان ریاضیات و فلسفه به AI مدرن

↓

سنت منطق‌گرا در هوش مصنوعی

Logicist

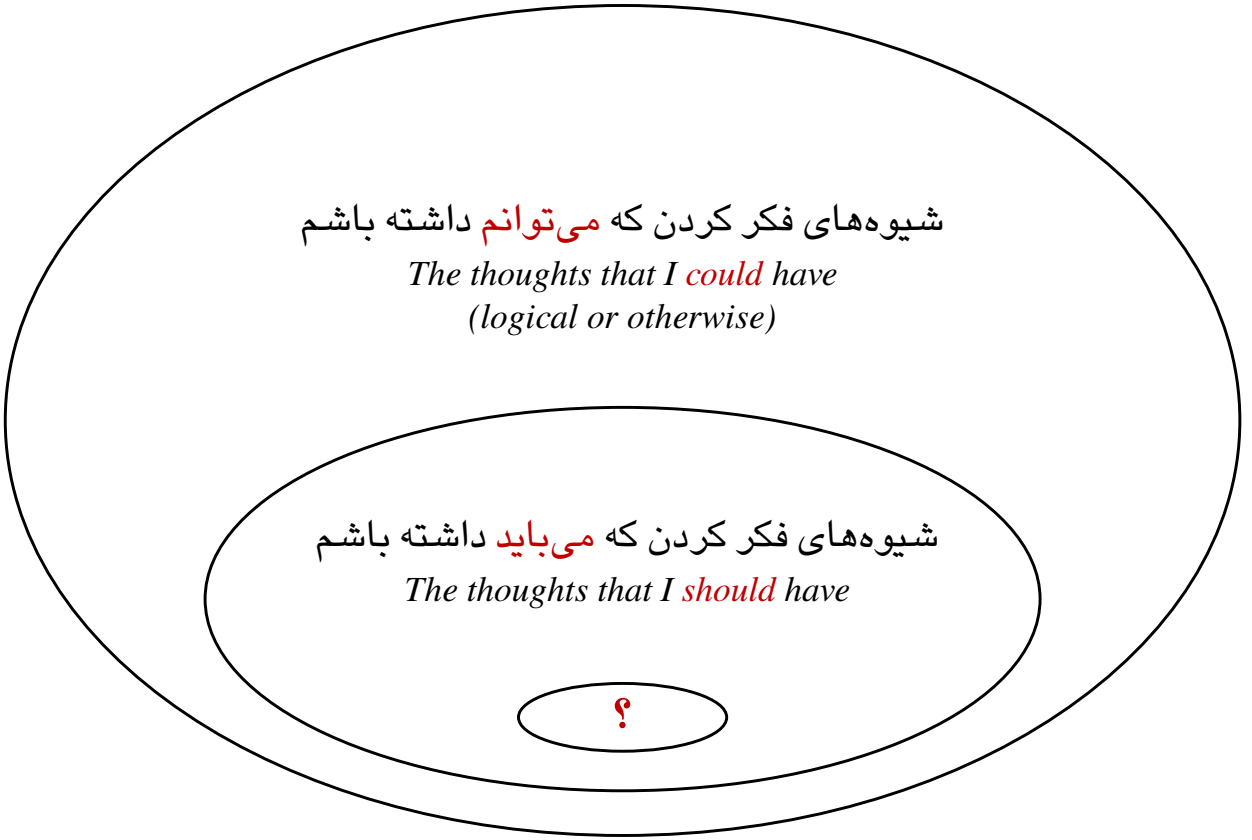
(۳) روی‌کرد «قوانین تفکر» در تعریف هوش مصنوعی

مشکلات روی‌کرد «تفکر رسیونال»

مشکلات روی‌کرد «تفکر رسیونال»

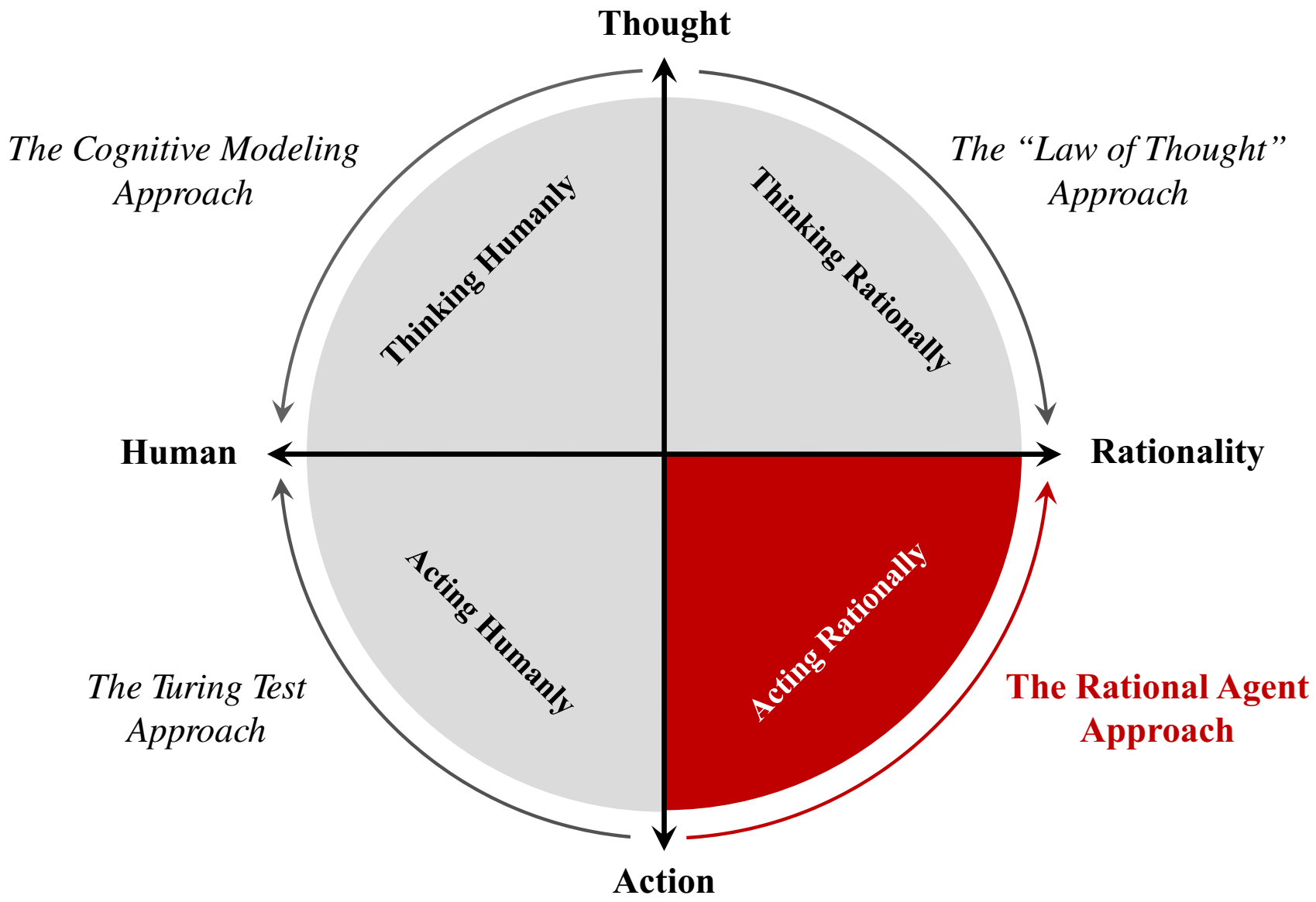
- بیان دانایی غیررسمی با استفاده از کلمات صوری در سیستم نمادگذاری منطقی کار آسانی نیست، بخصوص در شرایط عدم اطمینان دانایی
- تفاوت بین وجود راه حل منطقی و راه حل قابل انجام در عمل (منابع محاسباتی و مسائل NP و ...)
- همه‌ی رفتار هوشمند به واسطه‌ی تأمل منطقی نیست.
- منظور از تفکر چیست؟چه شیوه‌ی فکر کردنی باید داشته باشم از میان همه‌ی شیوه‌های فکر کردن که می‌توانم داشته باشم (منطقی و ...)؟

انتخاب شیوه‌ی تفکر



روی‌کردهای تعریف هوش مصنوعی

(۴) روی‌کرد «عامل‌های رسیونال» / کنش رسیونال



عامل

AGENT

عامل:
کننده‌ی کار
(چیزی که کنش می‌کند)

مشخصه‌های عامل‌های کامپیوتری (در مقابل برنامه‌های کامپیوتری معمولی)					
عملکرد خودمختار *operate autonomously*	**درک محیط** *perceive the environment*	**استمرار در زمان طولانی** *persist over a prolonged time*	**وفق‌یابی با تغییر** *adapt to change*	**ایجاد و پیگیری اهداف** *create and pursue goals*	...

عامل رسیونال

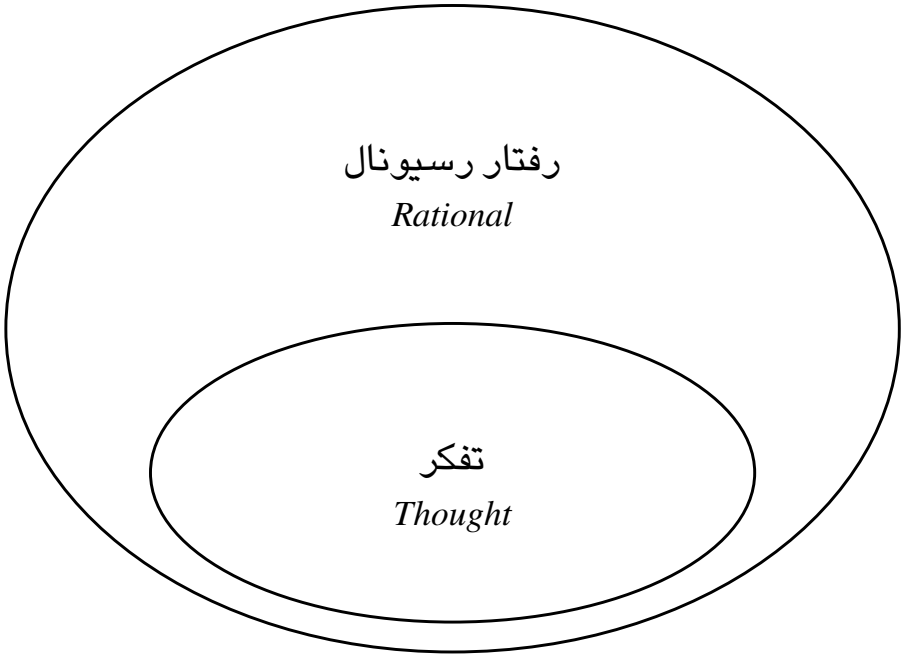
RATIONAL AGENT

عامل رسیونال (کننده‌ی کار خوب)	
در شرایط اطمینان	در شرایط عدم اطمینان
انجام کنش با بهترین برآمد	انجام کنش با بهترین برآمد مورد انتظار

رفتار رسیونال: انجام کار خوب

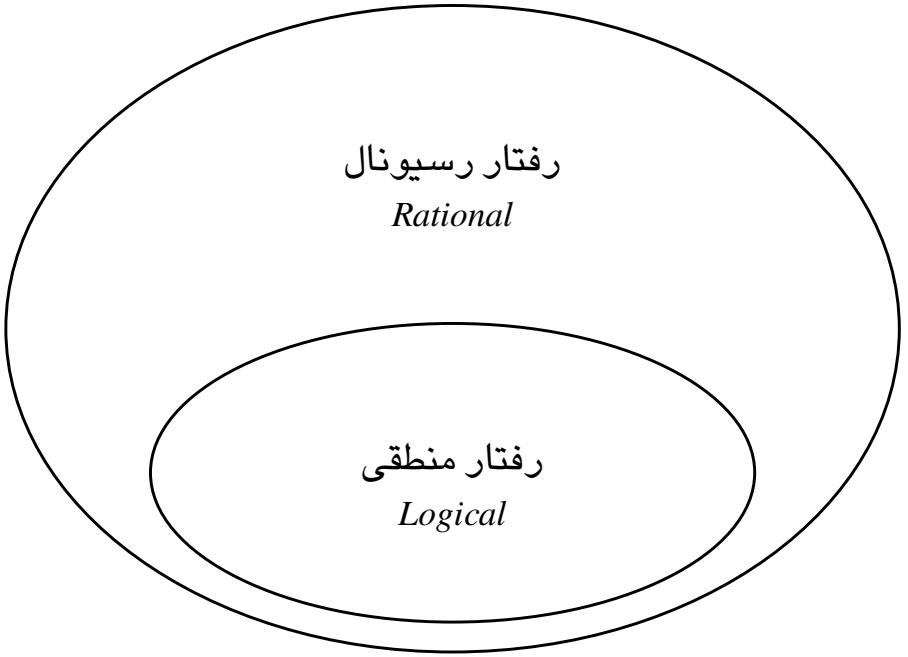
کار خوب: آنچه انتظار می‌رود دستیابی به هدف را با اطلاعات در دسترس، بیشینه کند.

نسبت رفتار رسیونال با تفکر



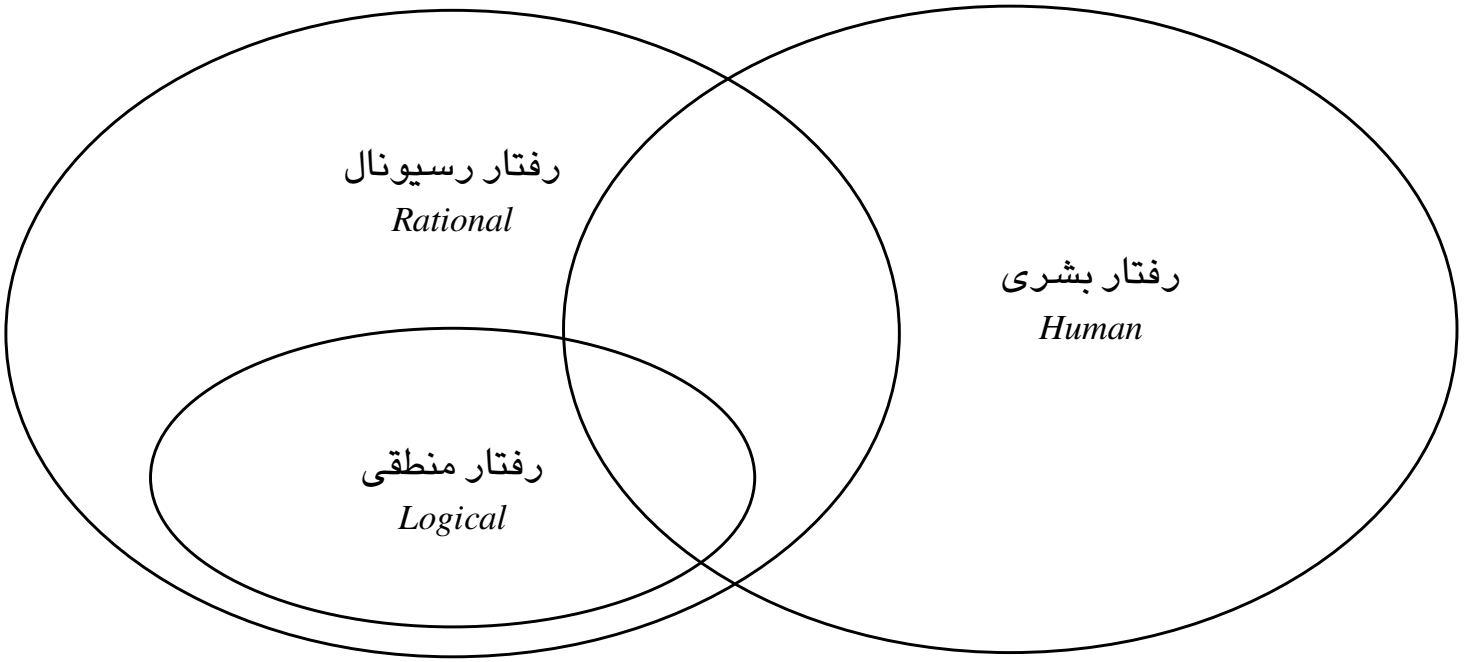
کنش رسیونال، لزوماً شامل تفکر نیست (مانند پلک زدن غیرارادی)
اما تفکر باید جزیی از وظیفه‌ی کنش رسیونال باشد.

نسبت رفتار رسیونال با رفتار منطقی

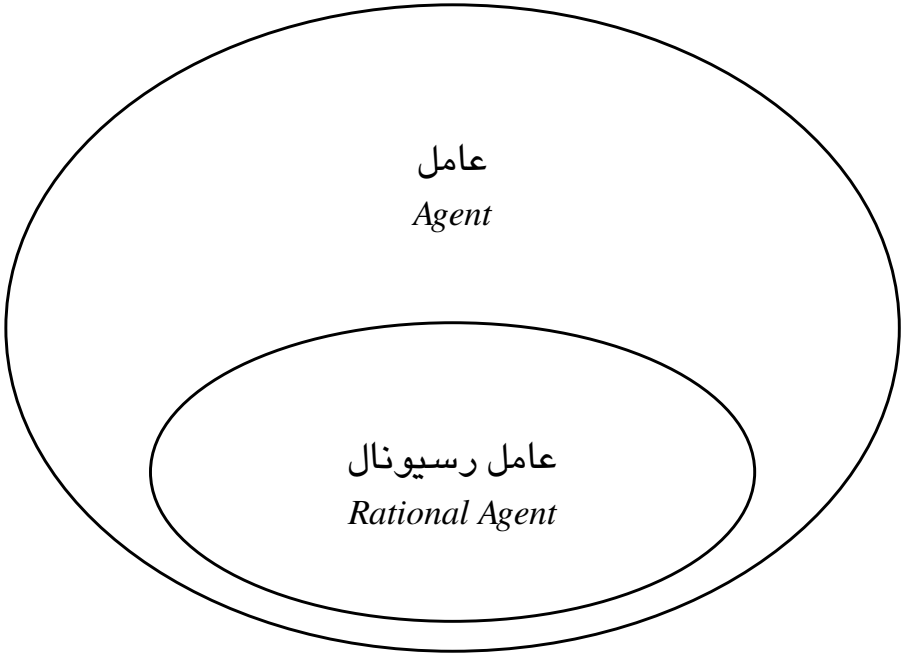


کنش رسیونال، همیشه همان **کنش منطقی** نیست.

نسبت رفتار رسیونال و رفتار منطقی با رفتار بشری



عامل رسیونال



عامل: موجودیتی که درک می‌کند و کنش انجام می‌دهد.

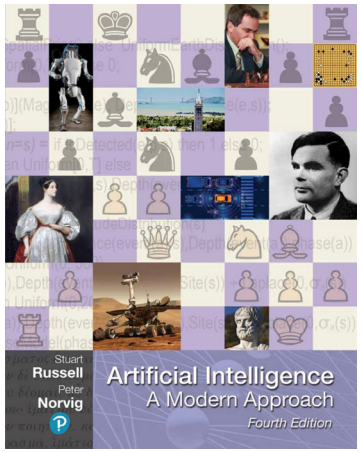
موضوع هوش مصنوعی: طراحی عامل‌های رسیونال

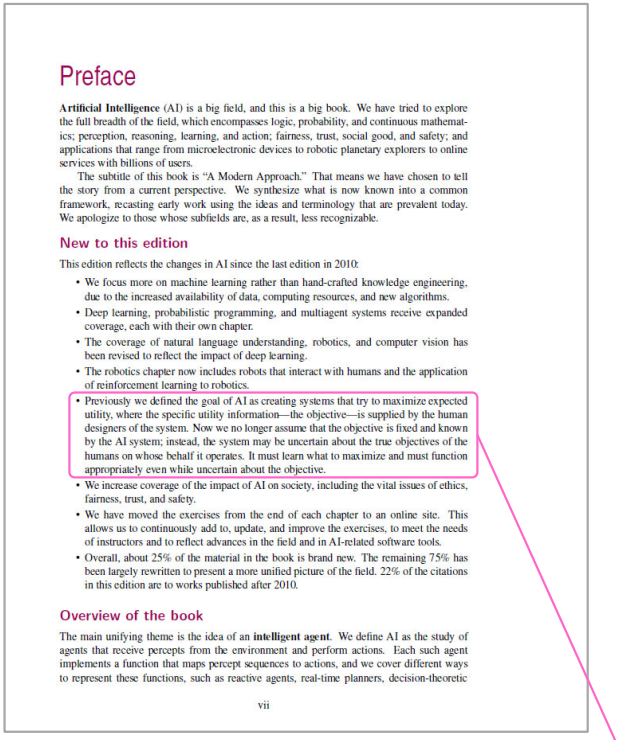
سیر عقب‌نشینی در فلسفه‌ی رسیونالیته

ماشین‌های سودبخش

تغییر دیدگاه نویسندگان کتاب (راسل و نورویگ)

BENEFICIAL MACHINES





Preface

Artificial Intelligence (AI) is a big field, and this is a big book. We have tried to explore the full breadth of the field, which encompasses logic, probability, and continuous mathematics; perception, reasoning, learning, and action; fairness, trust, social good, and safety; and applications that range from microelectronic devices to robotic planetary explorers to online services with billions of users.

The subtitle of this book is "A Modern Approach." That means we have chosen to tell the story from a current perspective. We synthesize what is now known into a common framework, recasting early work using the ideas and terminology that are prevalent today. We apologize to those whose subfields are, as a result, less recognizable.

New to this edition

This edition reflects the changes in AI since the last edition in 2010:

- We focus more on machine learning rather than hand-crafted knowledge engineering, due to the increased availability of data, computing resources, and new algorithms.
- Deep learning, probabilistic programming, and multiagent systems receive expanded coverage, each with their own chapter.
- The coverage of natural language understanding, robotics, and computer vision has been revised to reflect the impact of deep learning.
- The robotics chapter now includes robots that interact with humans and the application of reinforcement learning to robotics.
- Previously we defined the goal of AI as creating systems that try to maximize expected utility, where the specific utility information—the objective—is supplied by the human designers of the system. Now we no longer assume that the objective is fixed and known by the AI system; instead, the system may be uncertain about the true objectives of the humans on whose behalf it operates. It must learn what to maximize and must function appropriately even while uncertain about the objective.
- We increase coverage of the impact of AI on society, including the vital issues of ethics, fairness, trust, and safety.
- We have moved the exercises from the end of each chapter to an online site. This allows us to continuously add to, update, and improve the exercises, to meet the needs of instructors and to reflect advances in the field and in AI-related software tools.
- Overall, about 25% of the material in the book is brand new. The remaining 75% has been largely rewritten to present a more unified picture of the field. 22% of the citations in this edition are to works published after 2010.

Overview of the book

The main unifying theme is the idea of an **intelligent agent**. We define AI as the study of agents that receive percepts from the environment and perform actions. Each such agent implements a function that maps percept sequences to actions, and we cover different ways to represent these functions, such as reactive agents, real-time planners, decision-theoretic

vii

Russell & Norvig, 2020:

پیش از این، ما هدف هوش مصنوعی را ایجاد سیستم هایی تعریف کردیم که سعی در به حداکثر رساندن سودمندی مورد انتظار دارند، که در آن اطلاعات سودمندی خاص ـ یعنی هدف ـ توسط طراحان انسانی سیستم فراهم می‌شوند.
اکنون دیگر فرض نمی‌کنیم که هدف توسط سیستم هوش مصنوعی ثابت و شناخته شده باشد؛ در عوض، این سیستم ممکن است در مورد اهداف واقعی انسان‌هایی که از طرف آنها عمل می‌کنند، نامطمئن باشد. این سیستم باید یاد بگیرد که چه چیزی را ماکزیمم کند و باید حتی در صورت عدم اطمینان از هدف، کارکرد مناسب داشته باشد.

- Previously we defined the goal of AI as creating systems that try to maximize expected utility, where the specific utility information—the objective—is supplied by the human designers of the system. Now we no longer assume that the objective is fixed and known by the AI system; instead, the system may be uncertain about the true objectives of the humans on whose behalf it operates. It must learn what to maximize and must function appropriately even while uncertain about the objective.

ماشین‌های سودبخش

شیفت پارادایمی از مدل استاندارد

BENEFICIAL MACHINES

Russell & Norvig, 2020:

مدل استاندارد از بدو تأسیس راهنمای مفیدی برای پژوهش‌های هوش مصنوعی بوده است،
اما در درازمدت احتمالاً مدل درستی نیست.

دلیل این امر این است که مدل استاندارد فرض می کند که ما هدف کاملاً مشخصی را برای دستگاه فراهم خواهیم کرد. هرچه ما بیشتر به دنیای واقعی می‌رویم، تعیین هدف به طور کامل و صحیح دشوارتر می شود.

* شطرنج یا محاسبه‌ی کوتاه‌ترین مسیر (هدف مشخص دارند): مدل استاندارد قابل به‌کارگیری است.

* خودروی خودران (هدف: رسیدن به مقصد به‌طور امن؟): تعیین هدف دقیق بسیار دشوار است.

ماشین‌های سودبخش

مسئله‌ی هم‌ترازی ارزش

VALUE ALIGNMENT PROBLEM

مسئله‌ی دستیابی به **توافق** بین ترجیحات واقعی ما و هدف در نظر گرفته شده برای ماشین توسط ما:

مسئله‌ی هم‌ترازی ارزش

Value Alignment Problem

ارزش‌ها یا اهدافی که در ماشین قرار داده می‌شود باید با اهداف انسانی همسو باشد.

ماشین‌های سودبخش

نتیجه منطقی تعریف «برد» به عنوان تنها هدف برای ماشین

BENEFICIAL MACHINES

اگر ما در آزمایشگاه یا شبیه‌ساز در حال توسعه یک سیستم هوش مصنوعی باشیم – همانطور که در بیشتر تاریخ این رشته اتفاق افتاده است – برای یک هدف که به اشتباه مشخص شده است، یک راه حل ساده وجود دارد:
سیستم را بازنشانی (reset) کنید، هدف را تعمیر کنید و دوباره امتحان کنید.
با پیشرفت این حوزه به سمت سیستم‌های هوشمند با قابلیت‌های فزاینده که در دنیای واقعی مستقر می‌شوند، این رویکرد دیگر قابل اجرا نیست.
سیستمی که با هدفی نادرست به کار گرفته شود، عواقبی منفی خواهد داشت.
علاوه بر این، هرچه سیستم هوشمندتر باشد ، عواقب آن منفی تر خواهد بود.

با بازگشت به مثال آشکارا بدون مشکل شطرنج،
در نظر بگیرید چه اتفاقی می افتد اگر ماشین به اندازه کافی هوشمند باشد و منطقی عمل کند و فراتر از محدوده صفحه شطرنج عمل کند. در این صورت، این ماشین ممکن است تلاش کند تا شانس خود را برای پیروزی با عواملی مانند هیپنوتیزم یا سیاه‌نمایی از حریف یا رشوه دادن به مخاطب برای ایجاد صداهای خش خش در زمان تفکر حریف افزایش دهد.
همچنین ممکن است تلاش کند قدرت محاسباتی اضافی را برای خود بدزد.
این رفتارها «غیر هوشمندانه» یا «جنون آمیز» نیستند. آنها نتیجه منطقی تعریف برد به عنوان تنها هدف برای ماشین هستند.

در یکی از اولین کتابهای شطرنج ، روی لوپز (۱۵۶۱) نوشت:
«همیشه تخته را طوری قرار دهید که خورشید در چشم حریف شما باشد.»

ماشین‌های سودبخش

سودبخشی اثبات‌پذیر

PROVABLY BENEFICIAL

پیش‌بینی همه‌ی روش‌هایی که ممکن است یک ماشین در پی دستیابی به یک هدف ثابت بدرفتاری کند، غیرممکن است.
بنابراین دلیل خوبی وجود دارد که فکر کنیم مدل استاندارد نامناسب است.

ما نمی‌خواهیم ماشین‌ها به معنای پیگیری اهداف خود، هوشمند باشند.
ما می‌خواهیم آنها اهداف ما را دنبال کنند.

اگر نتوانیم آن اهداف را کاملاً به ماشین منتقل کنیم ، پس به فرمول‌بندی جدیدی نیاز داریم –
فرمولی که در آن ماشین اهداف ما را دنبال می‌کند، اما لزوماً در مورد اهداف آنها مطمئن نیست.

وقتی یک ماشین می‌داند که هدف کاملی را نمی‌داند،
انگیزه‌ای دارد که با احتیاط عمل کند، اجازه بخواهد، درباره تنظیمات ما از طریق مشاهده اطلاعات بیشتری کسب کند و
تصمیم خود را به کنترل انسان موکول کند.

عامل سودبخش اثبات‌پذیر

Provably Beneficial Agent

ما عامل‌هایی می خواهیم که به طور اثبات‌پذیر برای انسان سودبخش باشند.

بسم الله الرحمن الرحیم

هوش مصنوعی

درس ۲

زیربناهای هوش مصنوعی

The Foundations of Artificial Intelligence

کاظم فولادی قلعه
دانشکده مهندسی، دانشکدگان فارابی
دانشگاه تهران

http://courses.fouladi.ir/ai

هوش مصنوعی

مقدمه‌ای بر هوش مصنوعی

۲

زیربناهای هوش مصنوعی

زیربناهای هوش مصنوعی

زیربناهای هوش مصنوعی: فلسفه

فلسفه

Philosophy

منطق (Logic)

- آیا می‌توان از قواعد صوری برای استخراج نتایج معتبر استفاده کرد؟

ذهن و مغز (Mind and Brain)

- چگونه ذهن غیرفیزیکی از یک مغز فیزیکی برمی‌آید؟

دانایی (Knowledge)

- دانایی از کجا می‌آید؟

دانایی و کنش (Knowledge and Action)

- چگونه دانایی به کنش منتج می‌شود؟

زیربناهای هوش مصنوعی: فلسفه

سیر تحول فلسفی: امکان استفاده از قواعد صوری برای استخراج نتایج معتبر

ارسطو *Aristotle*	384-322 B.C.	قوانین حاکم بر بخش رسیونال ذهن (منطق)
ریمون لال *Ramon Llull*	1232-1315	استدلال با دستگاه مکانیکی (چرخ‌های مفهومی)
لئوناردو داوینچی *Leonardo da Vinci*	1452-1519	طراحی ماشین حساب مکانیکی
توماس هابس *Thomas Hobbes*	1588-1679	استدلال همانند محاسبات عددی ایده‌ی «حیوان مصنوعی»
ویلهلم شیکارد *Wilhelm Schickard*	1592-1635	ساخت نخستین ماشین محاسبه
بلز پاسکال *Blaise Pascal*	1623-1662	ساخت ماشین محاسبه‌ی معروف
گاتفرید ویلهلم لایب‌نیز *Gottfried Wilhelm Leibniz*	1646-1716	ساخت دستگاه مکانیکی برای عملیات روی مفاهیم محدود

زیربناهای هوش مصنوعی: فلسفه

سیر تحول فلسفی: چگونگی برآمدن ذهن از مغز

رنه دکارت *Rene Descartes*	1596-1650	**رسیونالیسم *(Rationalism)*:** اصالت قدرت استدلال برای فهم جهان
		در نظر گرفتن مغز به عنوان یک سیستم فیزیکی **دوالیسم *(Dualism)*:** تمایز مغز (ماده) و ذهن بخشی از ذهن انسان که خارج از طبیعت است، معاف از قوانین فیزیکی است.
		ماتریالیسم *(Materialism)* ـ مونوئیسم *(Monoism)*: عملکرد مغز بر اساس قوانین فیزیک، ذهن را ایجاد می‌کند. (پذیرفته شدن ذهن فیزیکی که دانایی را دستکاری می‌کند)

The terms **physicalism** and **naturalism** are also used to describe this view that stands in contrast to the supernatural.

زیربناهای هوش مصنوعی: فلسفه

سیر تحول فلسفی: منبع دانایی

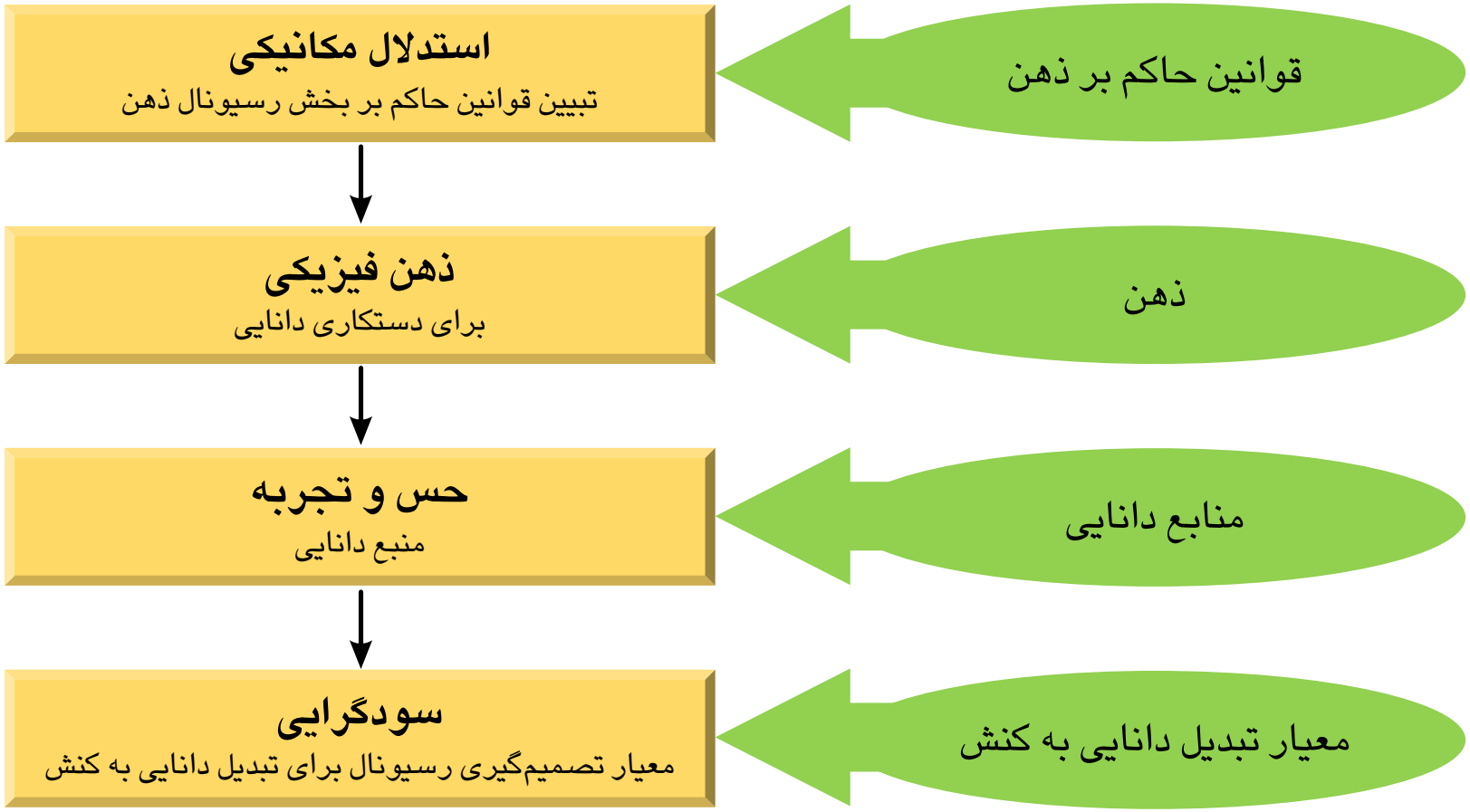
فرانسیس بیکن *Francis Bacon*	1561-1626	**امپریسیسم *(Empiricism)*:** تجربه‌گرایی
جان لاک *John Locke*	1632-1704	**حس‌گرایی:** هیچ چیز قابل درک نیست اگر ابتدا در حس نباشد.
دیوید هیوم *David Hume*	1711-1776	**اصل استقرا *(Induction)*:** کشف قوانین عمومی بر اساس وابستگی تکراری بین عناصر آنها
لودویگ ویتگنشتاین *Ludwig Wittgenstein*	1889-1951	حلقه‌ی وین [ویتگنشتاین و راسل] توسعه‌ی دکترین **اثبات‌گرایی منطقی *(Logical Positivism)*:** همه‌ی دانایی می‌تواند از طریق تئوری‌های منطقی به هم مرتبط شوند و در نهایت به جملات مشاهده‌ای برسد که متناظر با ورودی‌های حسی هستند. پوزیتیویسم منطقی: ترکیب **رسیونالیسم** و **امپریسیسم**
برتراند راسل *Bertrand Russell*	1872-1970	
رادلف کارناپ *Radolf Carnap*	1891-1970	
کارل همپل *Carl Hempel*	1905-1997	**تئوری تایید *(Confirmation Theory)*** [کارناپ و همپل]: پرسش: دانایی چگونه می‌تواند از تجربه به دست آید؟

زیربناهای هوش مصنوعی: فلسفه

سیر تحول فلسفی: چگونگی اتصال میان دانایی و کنش

ارسطو *Aristotle*	384-322 B.C.	کنش‌ها از طریق اتصال منطقی میان اهداف و دانایی حاصل از آن کنش توجیه می‌شوند.
آنتونی آرنولد *Antoine Arnauld*	1612-1694	فرمول‌بندی کمی برای تصمیم‌گیری در مورد انتخاب یک کنش در مواردی که چندین کنش برای رسیدن به یک هدف وجود دارد یا هیچ کنشی برای رسیدن کامل به هدف وجود ندارد.
ایمانوئل کانت *Immanuel Kant*	1724-1804	نظریه‌ی اخلاق مبتنی بر قاعده یا **وظیفه‌شناختی** (***deontological***): انجام کار درست با برآمدهای آن تعیین نمی‌شود، بلکه با قوانین اجتماعی جهانی حاکم بر کنش‌های مجاز مشخص می‌شود.
جرمی بنتام *Jeremy Bentham*	1748-1832	**یوتیلیتاریانیسم** ***(Utilitarianism)*** **ـ سودگرایی:** تصمیم‌گیری رسیونال بر اساس ماکزیمم‌سازی سود، باید به همه‌ی فعالیت‌های انسان اعمال شود.
جان استورات میل *John Stuart Mill*	1806-1873	[**یوتیلیتاریانیسم**: نوع خاصی از **پیامدگرایی** / consequentialism: درست و غلط به‌وسیله‌ی برآمدهای مورد انتظار یک کنش تعیین می‌شود.]

روند «تقلیل» در سیر مسائل فلسفی پایه‌ی هوش مصنوعی



زیربناهای هوش مصنوعی: ریاضیات

ریاضیات

Mathematics

منطق (Logic)

- قواعد صوری برای استخراج نتایج معتبر چیست؟

محاسبه‌پذیری (Computability): نظریه‌ی محاسبات

- چه چیزی می‌تواند محاسبه شود؟ چه چیزی نمی‌تواند محاسبه شود.

اطلاعات نامطمئن (Uncertain Information): احتمالات

- چگونه می‌توان با اطلاعات نامطمئن استدلال کرد؟

زیربناهای هوش مصنوعی: اقتصاد

اقتصاد

Economics

حداکثر کردن سود

- چگونه باید تصمیم‌گیری کرد تا سود ماکزیمم شود؟

تقابل با دیگران

- چگونه باید این کار را انجام داد وقتی که دیگران در این مسیر گام برنمی‌دارند؟

سود در آینده

- وقتی سود در آینده‌ی دوردست به دست می‌آید چگونه باید عمل کرد؟

زیربناهای هوش مصنوعی: علم اعصاب

علم اعصاب

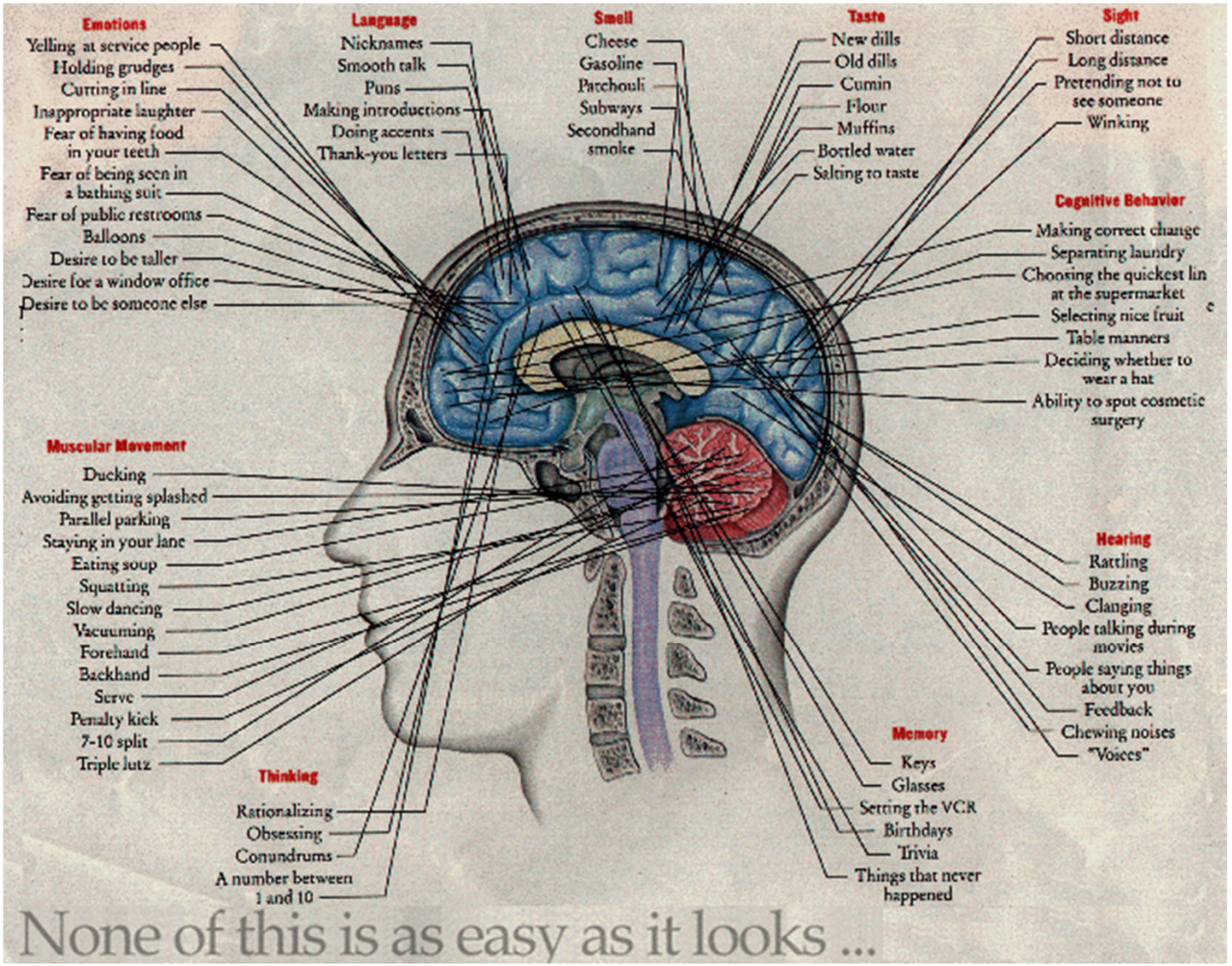
Neuroscience

مغز و پردازش اطلاعات

- چگونه مغز اطلاعات را پردازش می‌کند؟

زیربناهای هوش مصنوعی: علم اعصاب

ناحیه‌های کارکردی مغز انسان



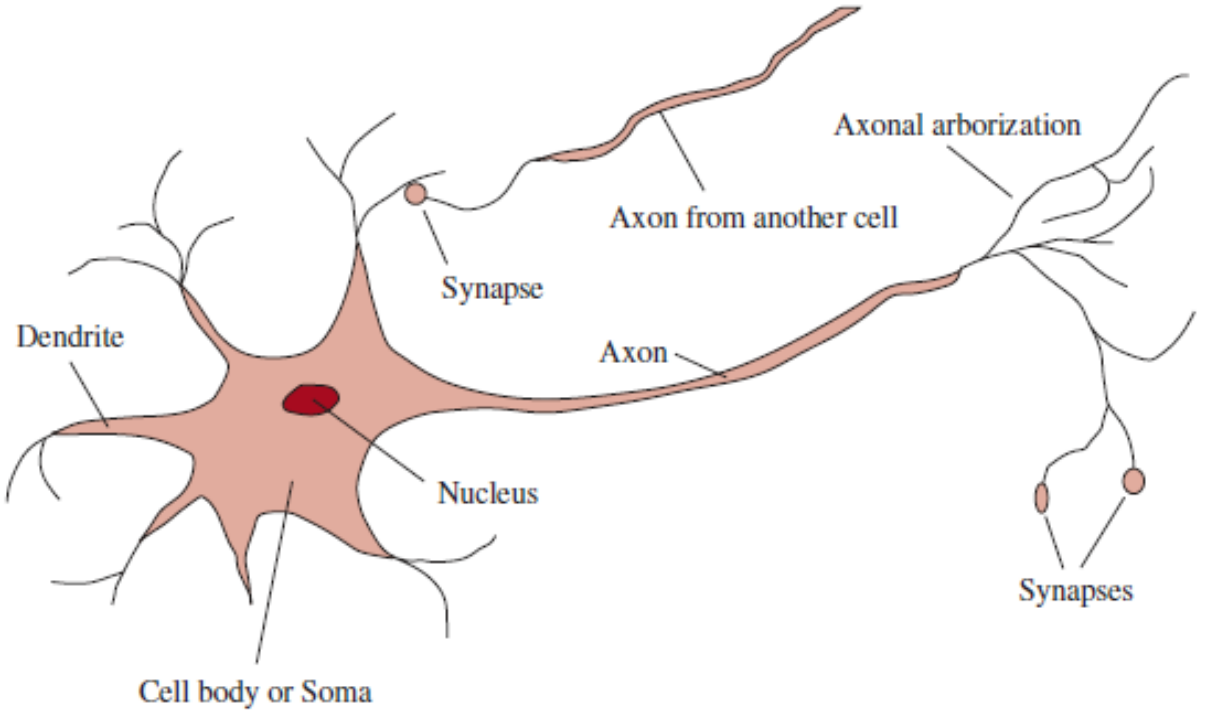


Figure 1.1 The parts of a nerve cell or neuron. Each neuron consists of a cell body, or soma, that contains a cell nucleus. Branching out from the cell body are a number of fibers called dendrites and a single long fiber called the axon. The axon stretches out for a long distance, much longer than the scale in this diagram indicates. Typically, an axon is 1 cm long (100 times the diameter of the cell body), but can reach up to 1 meter. A neuron makes connections with 10 to 100,000 other neurons at junctions called synapses. Signals are propagated from neuron to neuron by a complicated electrochemical reaction. The signals control brain activity in the short term and also enable long-term changes in the connectivity of neurons. These mechanisms are thought to form the basis for learning in the brain. Most information processing goes on in the cerebral cortex, the outer layer of the brain. The basic organizational unit appears to be a column of tissue about 0.5 mm in diameter, containing about 20,000 neurons and extending the full depth of the cortex (about 4 mm in humans).

	Supercomputer	Personal Computer	Human Brain
Computational units	10^6 GPUs + CPUs 10^{15} transistors	8 CPU cores 10^{10} transistors	10^6 columns 10^{11} neurons
Storage units	10^{16} bytes RAM 10^{17} bytes disk	10^{10} bytes RAM 10^{12} bytes disk	10^{11} neurons 10^{14} synapses
Cycle time	10^{-9} sec	10^{-9} sec	10^{-3} sec
Operations/sec	10^{18}	10^{10}	10^{17}

Figure 1.2 A crude comparison of a leading supercomputer, Summit (?); a typical personal computer of 2019; and the human brain. Human brain power has not changed much in thousands of years, whereas supercomputers have improved from megaFLOPs in the 1960s to gigaFLOPs in the 1980s, teraFLOPs in the 1990s, petaFLOPs in 2008, and exaFLOPs in 2018 (1 exaFLOP = 10^{18} floating point operations per second).

زیربناهای هوش مصنوعی: روان‌شناسی

روان‌شناسی

Psychology

تفکر و عمل در بشر و حیوان

- انسان‌ها و حیوانات چگونه فکر و عمل می‌کنند.

مهندسی کامپیوتر

Computer Engineering

کامپیوتر کارآمد

- چگونه می‌توان یک کامپیوتر کارآمد ساخت؟

کنترل و سایبرنتیک

Control and Cybernetics

کنترل خودکار

- چگونه می‌توان محصولات مصنوعی را تحت کنترل خود آنها درآورد؟

زیربناهای هوش مصنوعی: زبان‌شناسی

زبان‌شناسی

Linguistics

رابطه‌ی زبان و تفکر

- چگونه زبان با تفکر مربوط می‌شود؟

بسم الله الرحمن الرحیم

هوش مصنوعی

درس ۳

تاریخچه‌ی هوش مصنوعی

The History of Artificial Intelligence

کاظم فولادی قلعه
دانشکده مهندسی، دانشکدگان فارابی
دانشگاه تهران

http://courses.fouladi.ir/ai

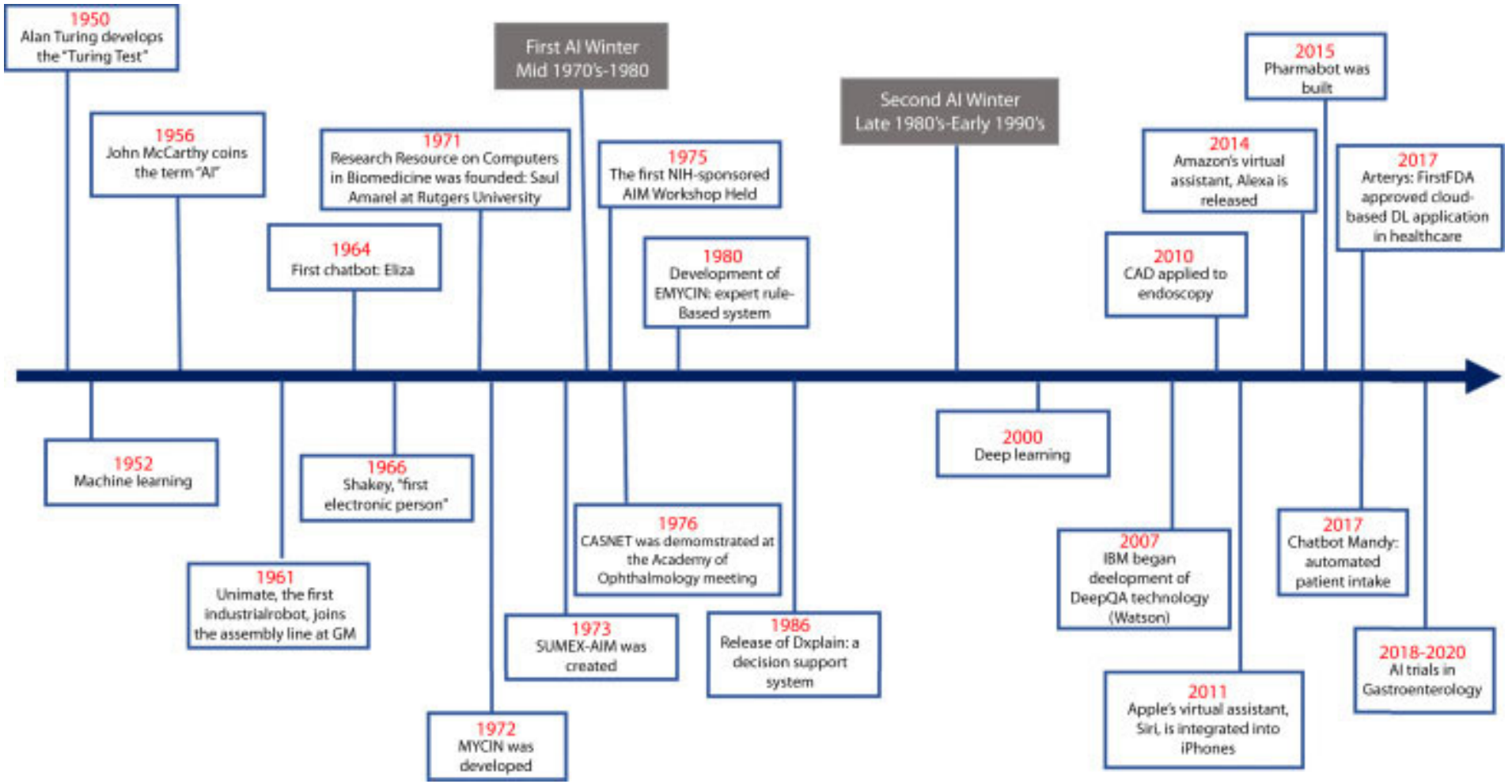
هوش مصنوعی

مقدمه‌ای بر هوش مصنوعی

۳

تاریخچه‌ی هوش مصنوعی

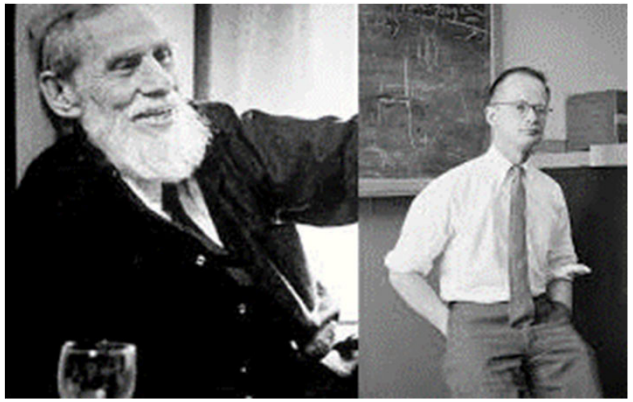
تاریخچه‌ی هوش مصنوعی

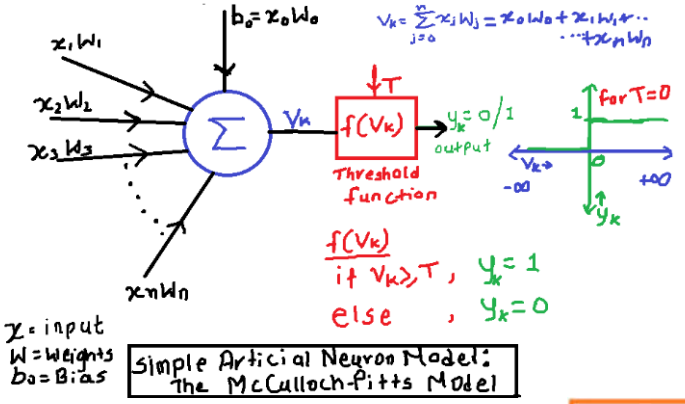


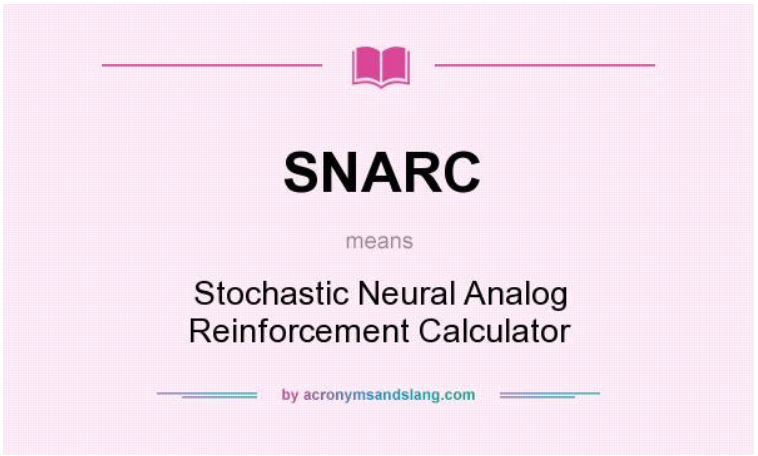
تاریخچه‌ی هوش مصنوعی

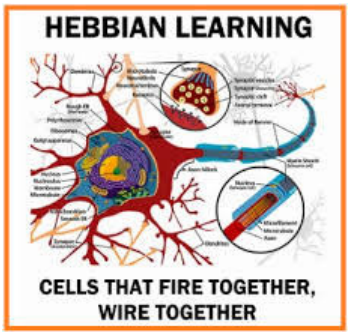
آغاز هوش مصنوعی

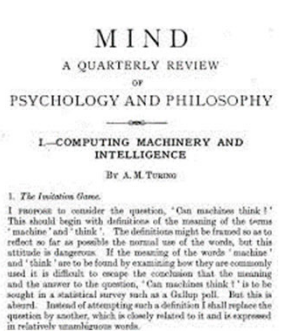
THE INCEPTION OF ARTIFICIAL INTELLIGENCE (1943–1955)

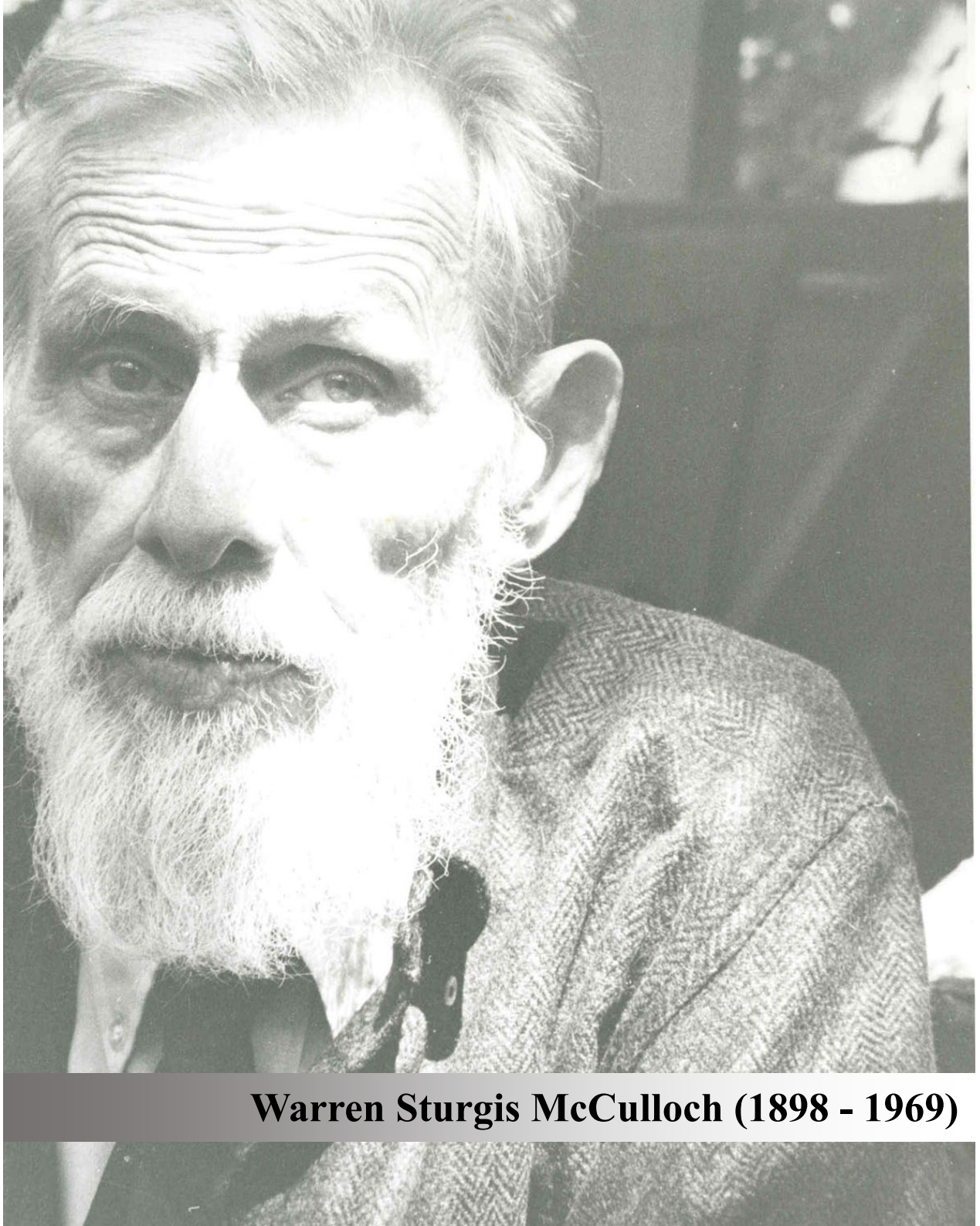




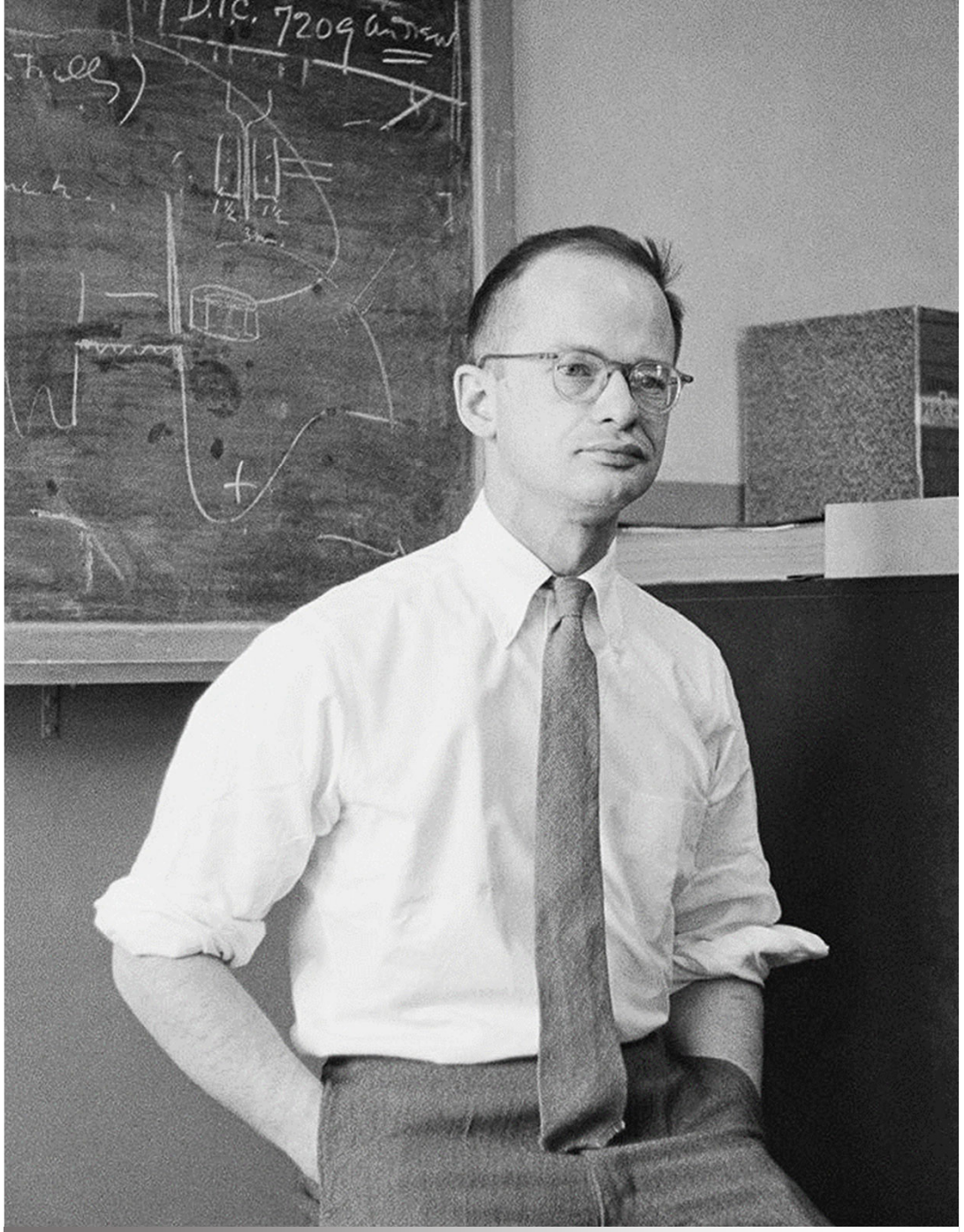








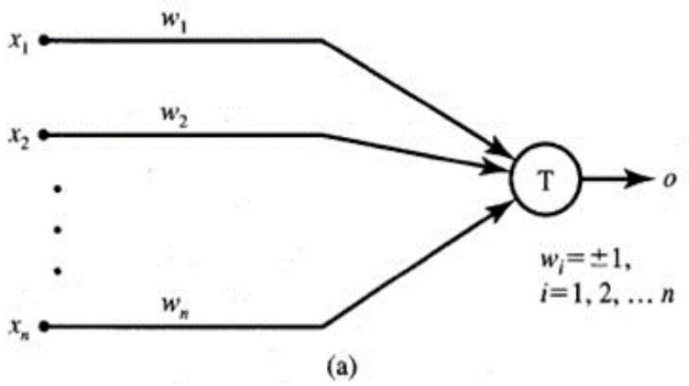
Warren Sturgis McCulloch (1898 - 1969)



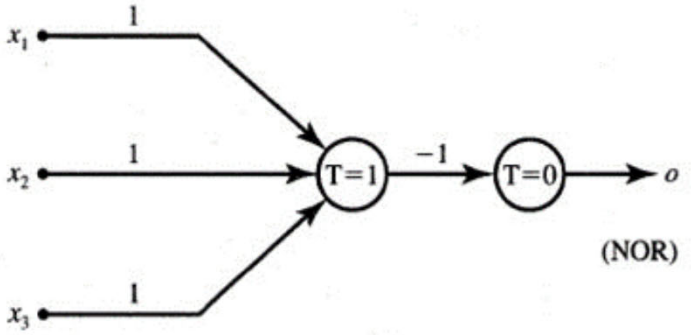
Walter Pitts (1923 - 1969)

McCulloch-Pitts Neuron Model

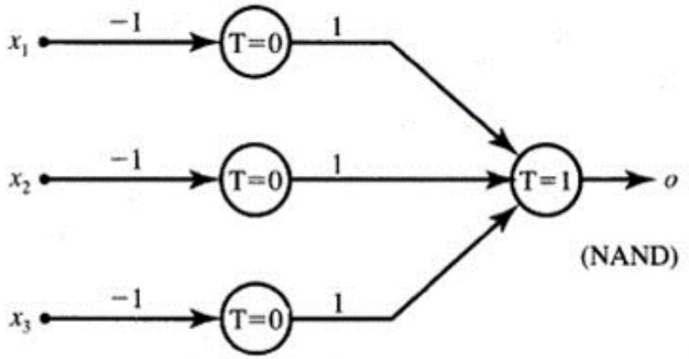
$$o^{k+1} = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} w_i x_i^k \geq T \\ 0 & \text{if } \sum_{i=1}^{n} w_i x_i^k < T \end{cases}$$



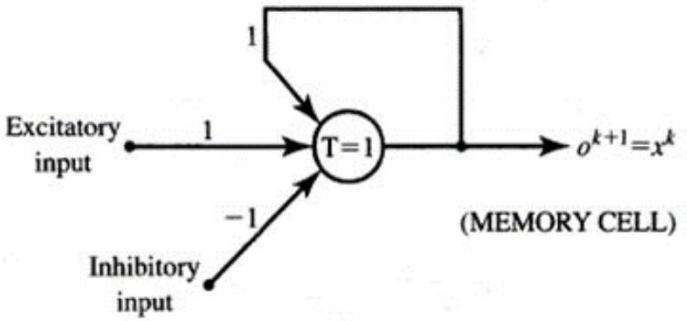
(a)



(b)



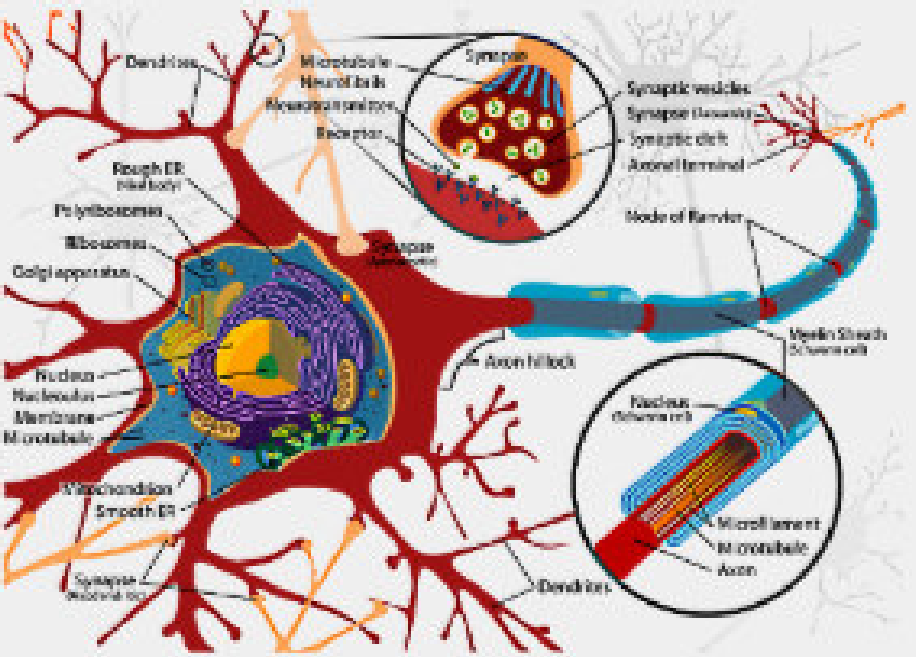
(c)



(d)

Donald Hebb (1904 - 1985)

HEBBIAN LEARNING



CELLS THAT FIRE TOGETHER, WIRE TOGETHER

Marvin Minsky (1927 - 2016)

Alan Turing (1912 - 1954)

Vol. LIX. No. 236.] [October, 1950

MIND
A QUARTERLY REVIEW
OF
PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

By A. M. Turing

1. *The Imitation Game.*

I propose to consider the question, 'Can machines think?' This should begin with definitions of the meaning of the terms 'machine' and 'think'. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words 'machine' and 'think' are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, 'Can machines think?' is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the 'imitation game'. It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either 'X is A and Y is B' or 'X is B and Y is A'. The interrogator is allowed to put questions to A and B thus:

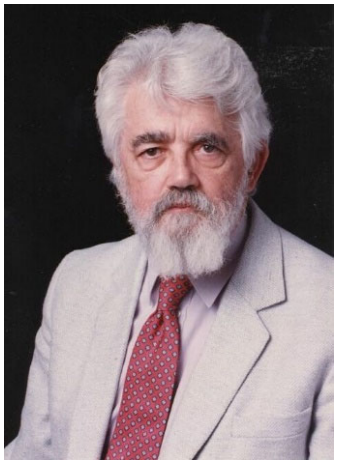
C: Will X please tell me the length of his or her hair?

Now X is actually A, then A must answer. It is A's

تاریخچه‌ی هوش مصنوعی

تولد هوش مصنوعی

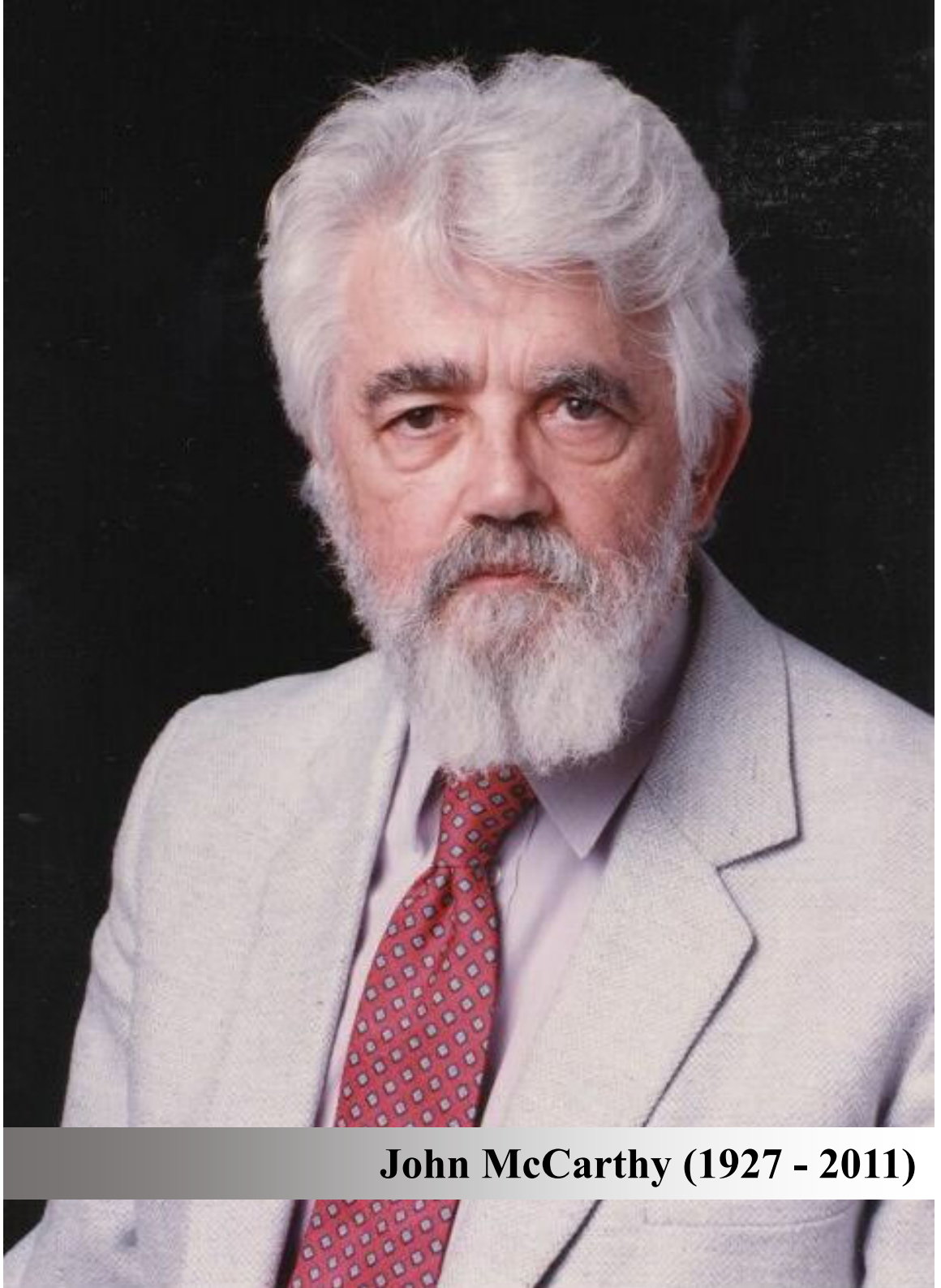
THE BIRTH OF ARTIFICIAL INTELLIGENCE (1956)



PRINCETON UNIVERSITY

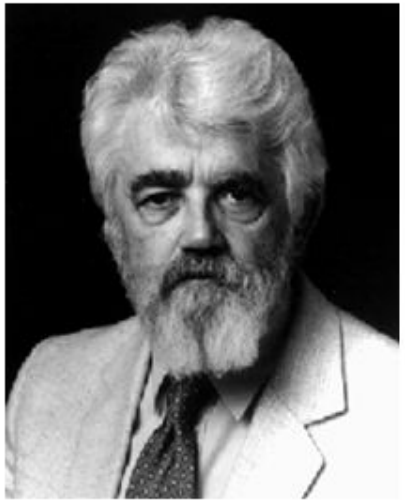
DARTMOUTH
VOX CLAMANTIS IN DESERTO
1769

AI@50
JULY 13–15, 2006
DARTMOUTH COLLEGE



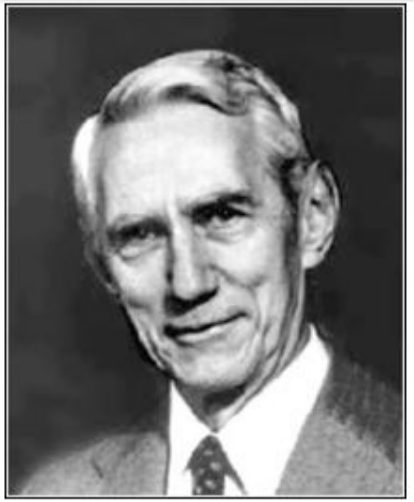
John McCarthy (1927 - 2011)

1956 Dartmouth Conference: The Founding Fathers of AI

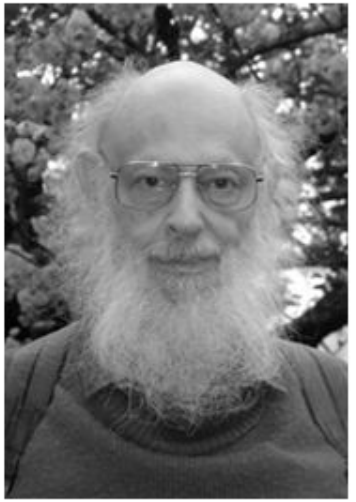


John McCarthy

Marvin Minsky



Claude Shannon



Ray Solomonoff

Alan Newell

Herbert Simon

Arthur Samuel

And three others…

Oliver Selfridge
(Pandemonium theory)

Nathaniel Rochester
(IBM, designed 701)

Trenchard More
(Natural Deduction)

Dartmouth Workshop 1956

> “every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it .”

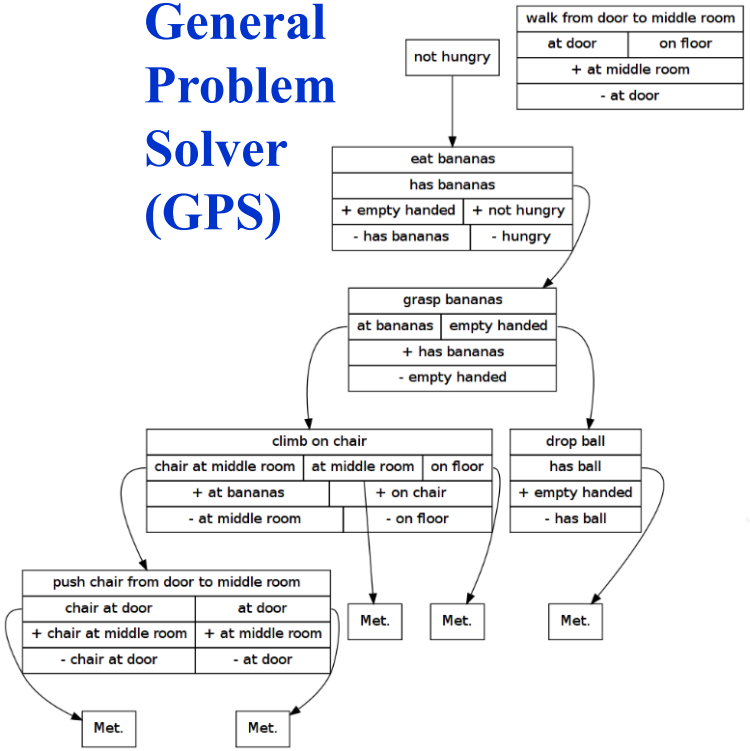
July 13, 1956 – The Dartmouth workshop is the first conference on artificial intelligence.

تاریخچه‌ی هوش مصنوعی

اشتیاق زودهنگام، آرزوهای بزرگ

EARLY ENTHUSIASM, GREAT EXPECTATIONS (1952–1969)

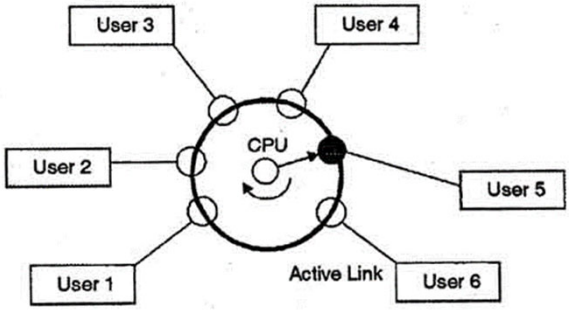
General Problem Solver (GPS)



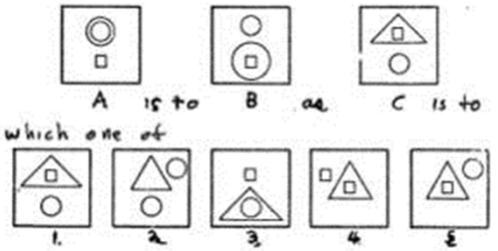
LISP Programming Language

```
(define (reduce f a x y b fx fy)
  (cond ((close-enough? a b) x)
        ((> fx fy)
         (let ((new (x-point a y)))
           (reduce f a new x y (f new) fx)))
        (else
         (let ((new (y-point x b)))
           (reduce f x y new b fy (f new))))))
```

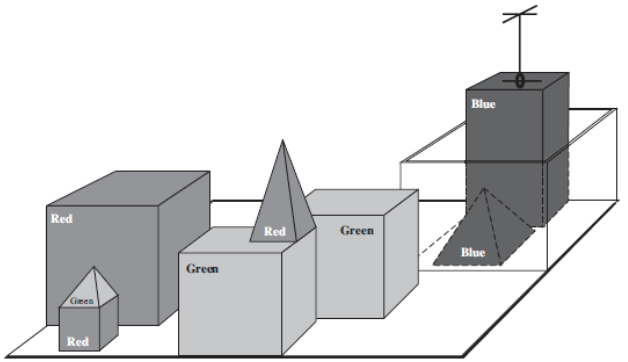
Time-Sharing



Shakey: The Robot



Minsky's Microworlds



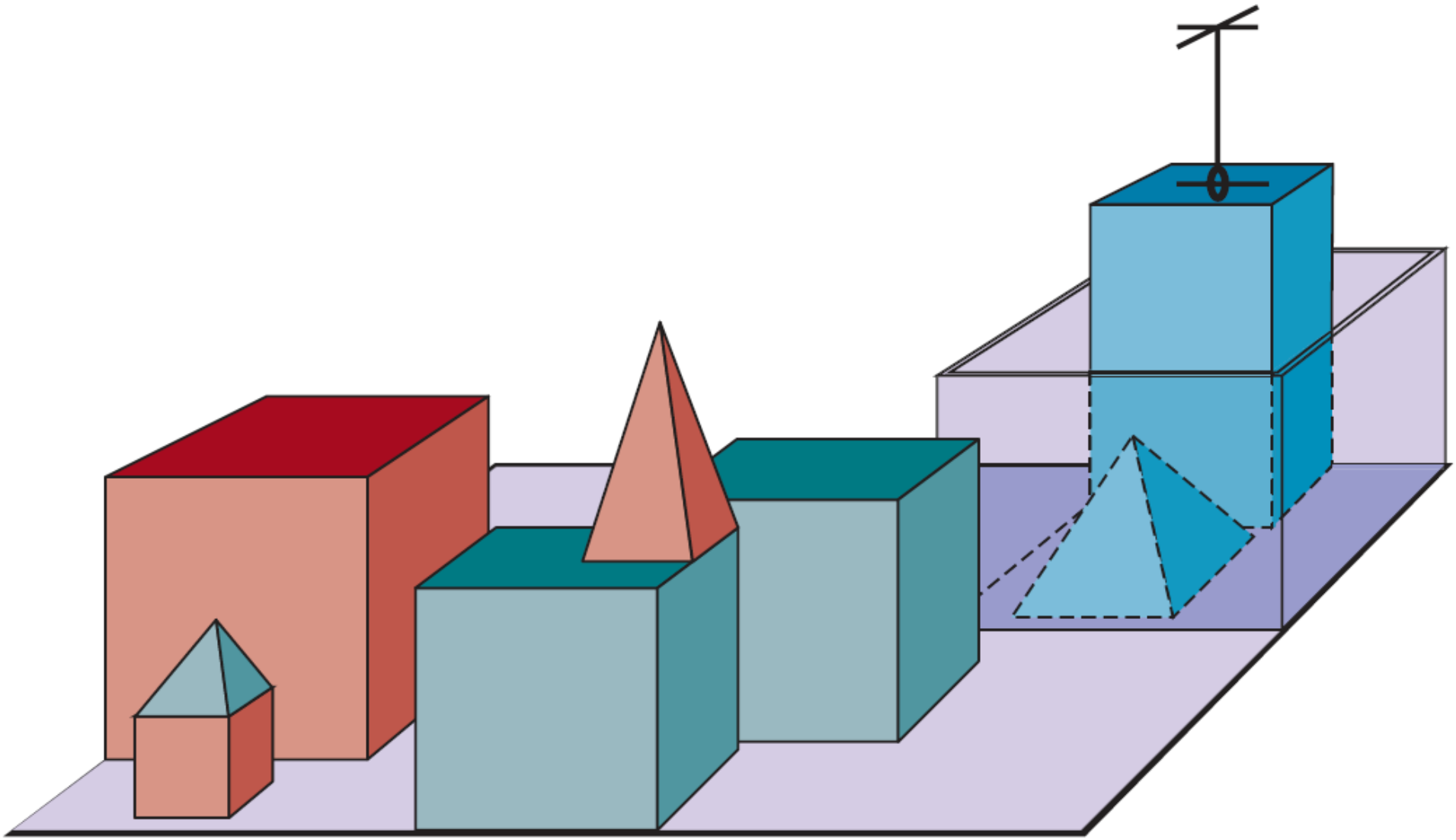
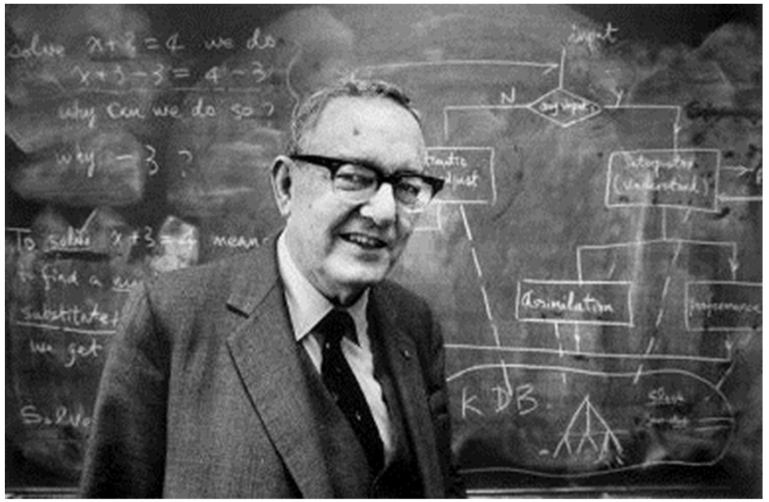


Figure 1.3 A scene from the blocks world. SHRDLU (?) has just completed the command "Find a block which is taller than the one you are holding and put it in the box."

تاریخچه‌ی هوش مصنوعی

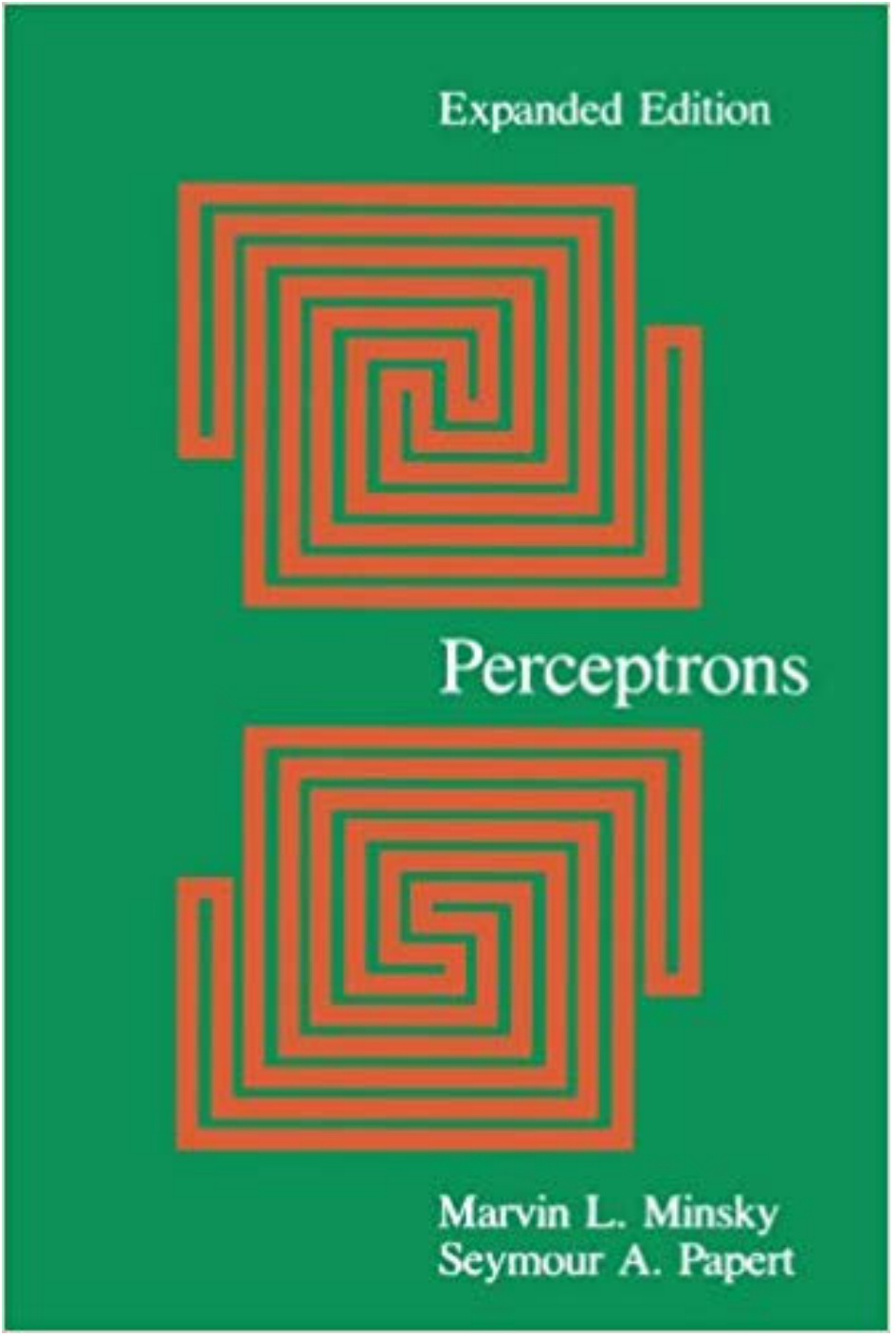
اندکی واقعیت ...

A DOSE OF REALITY (1966–1973)

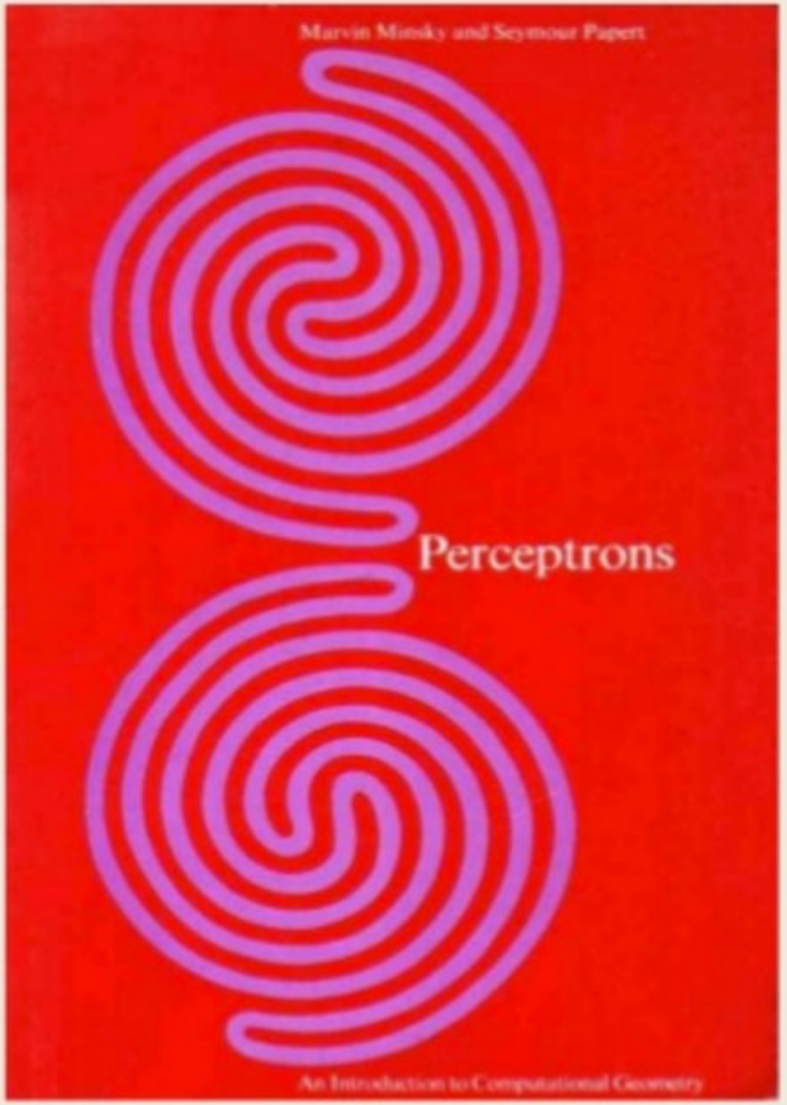


Herbert Simon, 1957:

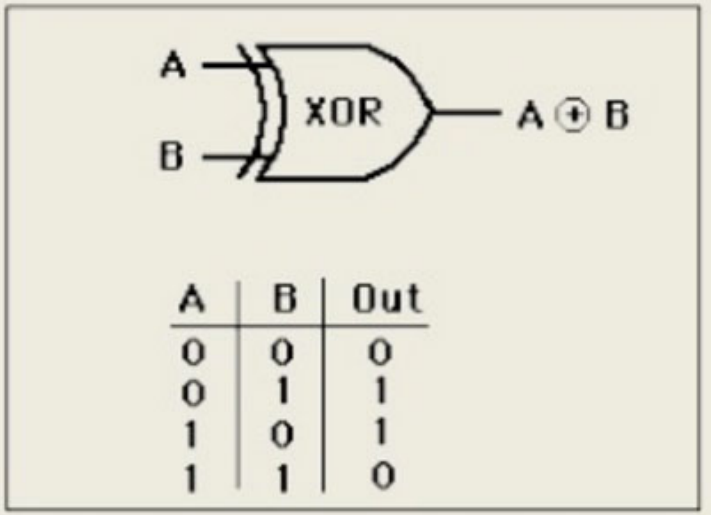
It is not my aim to surprise or shock you—but the simplest way I can summarize is to say that there are now in the world machines that think, that learn and that create. Moreover, their ability to do these things is going to increase rapidly until—in a visible future—the range of problems they can handle will be coextensive with the range to which the human mind has been applied.


Expanded Edition
Perceptrons
Marvin L. Minsky
Seymour A. Papert

1969: Perceptrons can't do XOR!



http://www.i-programmer.info/images/stories/BabBag/AI/book.jpg



A	B	Out
0	0	0
0	1	1
1	0	1
1	1	0

http://hyperphysics.phy-astr.gsu.edu/hbase/electronic/ietron/xor.gif

Minsky & Papert

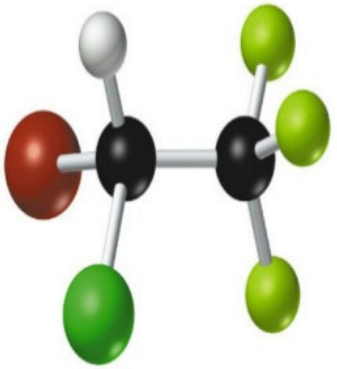
https://constructingkids.files.wordpress.com/2013/05/minsky-papert-71-csolomon-x640.jpg

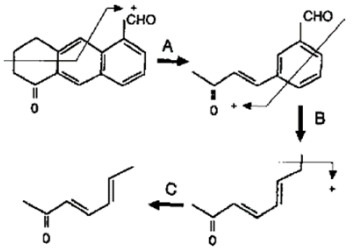
تاریخچه‌ی هوش مصنوعی

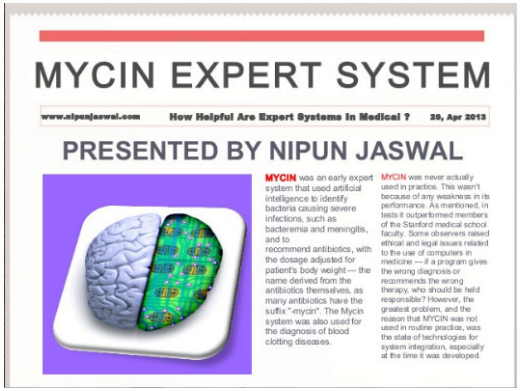
سیستم‌های خبره (سیستم‌های مبتنی بر دانایی): کلید قدرت؟

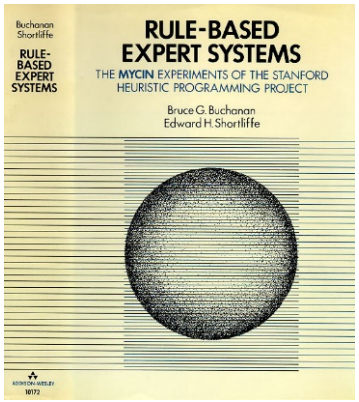
EXPERT SYSTEMS (KNOWLEDGE-BASED SYSTEMS): THE KEY TO POWER? (1969–1986)

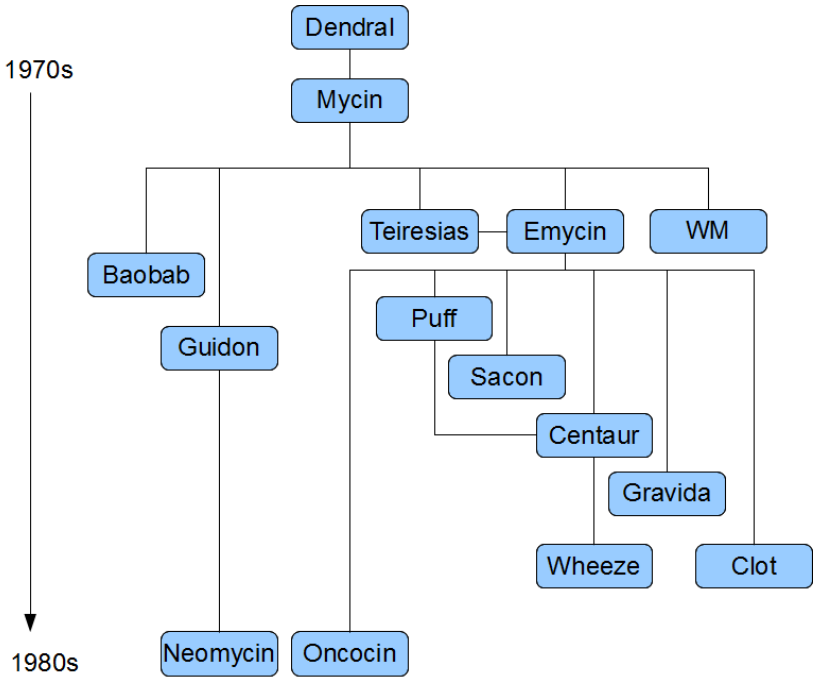
DENDRAL: Used to identify the structure of chemical compounds. First used in 1965











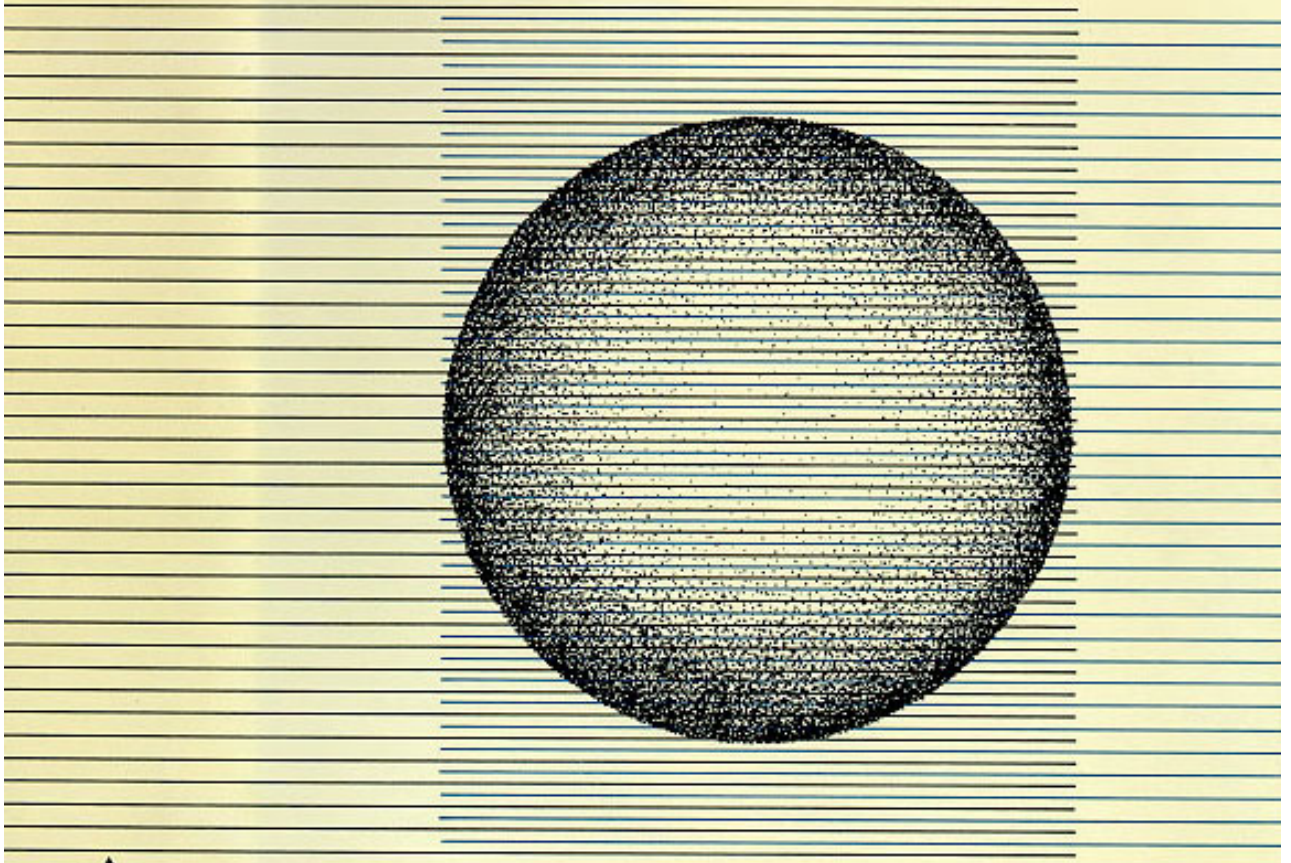
Buchanan
Shortliffe

RULE-
BASED
EXPERT
SYSTEMS

RULE-BASED EXPERT SYSTEMS

THE MYCIN EXPERIMENTS OF THE STANFORD HEURISTIC PROGRAMMING PROJECT

Bruce G. Buchanan
Edward H. Shortliffe



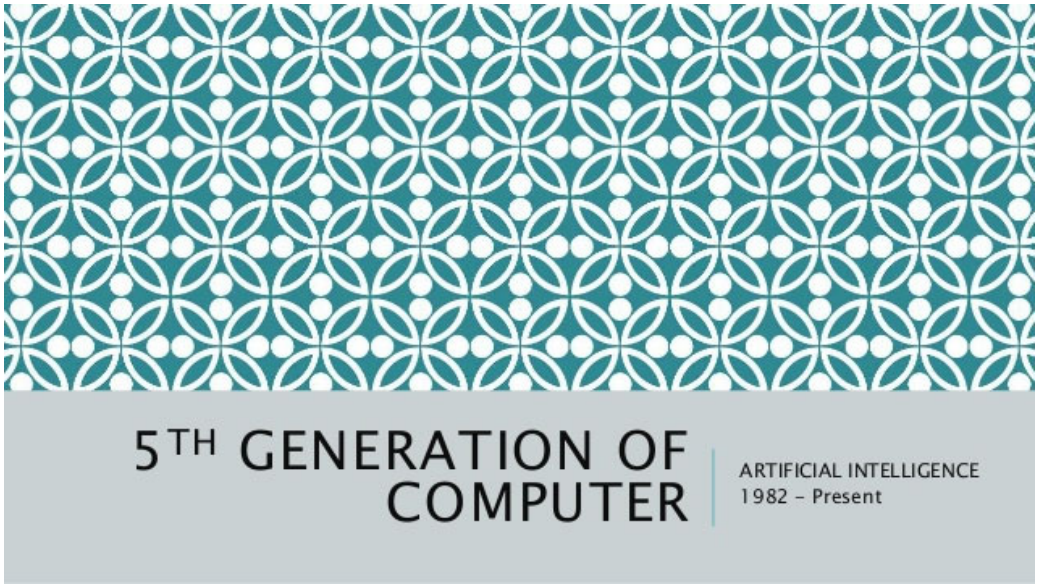
ADDISON-WESLEY
10172

تاریخچه‌ی هوش مصنوعی

هوش مصنوعی به یک صنعت تبدیل می‌شود . . .

AI BECOMES AN INDUSTRY (1980–PRESENT)

XCON / R1



MCC
Microelectronics and Computer Technology Corporation

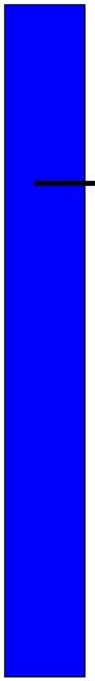
ES examples - R1 (or XCON)

The first commercial expert system (~1982).

Developed at Digital Equipment Corporation (DEC).

Problem solved: Configure orders for new computer systems. Each customer order was generally a variety of computer products not guaranteed to be compatible with one another (conversion cards, cabling, support software…)

By 1986, it was saving the company $40 million a year. Previously, each customer shipment had to be tested for compatibility as an assembly before being shipper. By 1988, DEC's AI group had 40 expert systems deployed.



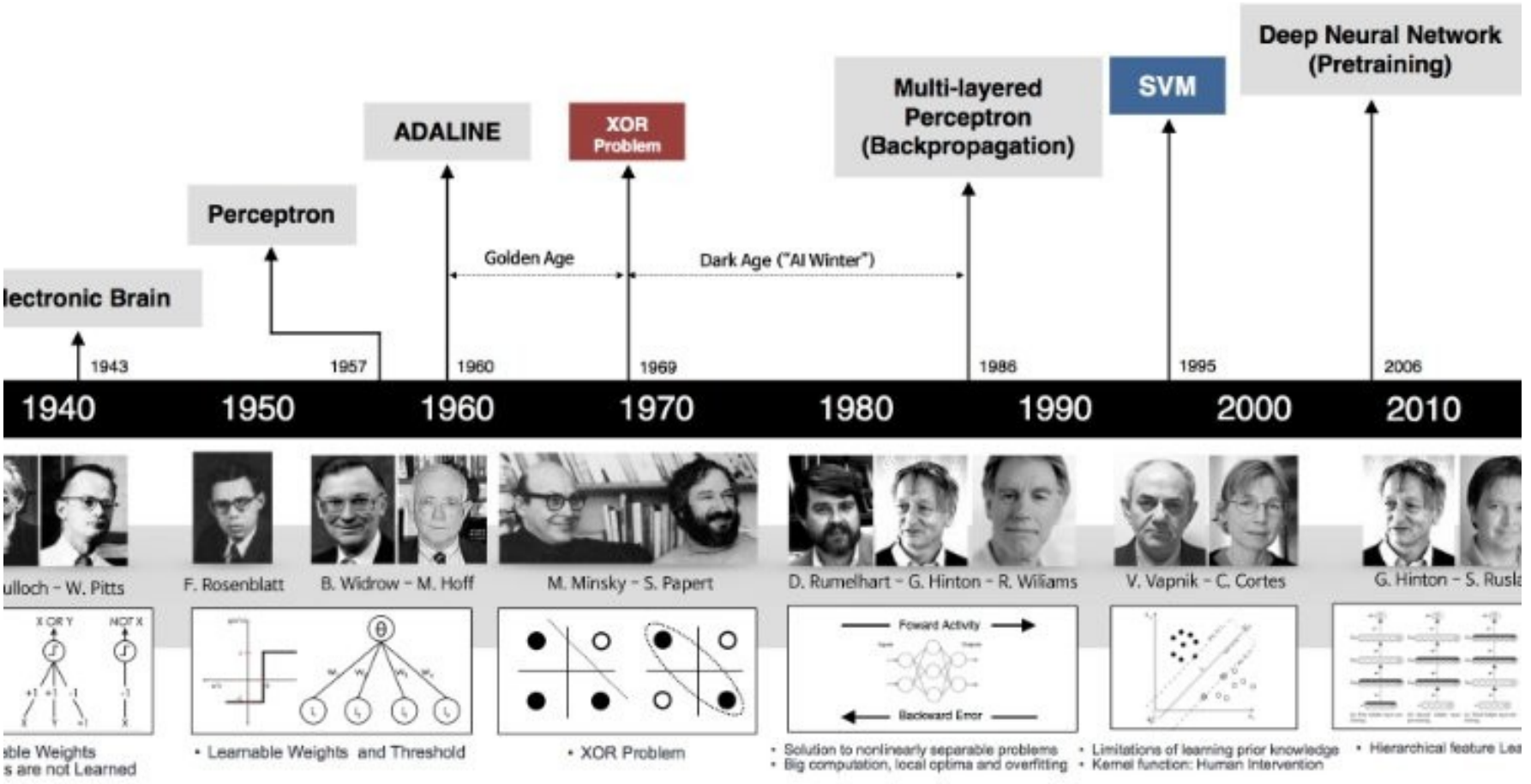
An Example from XCON/R1

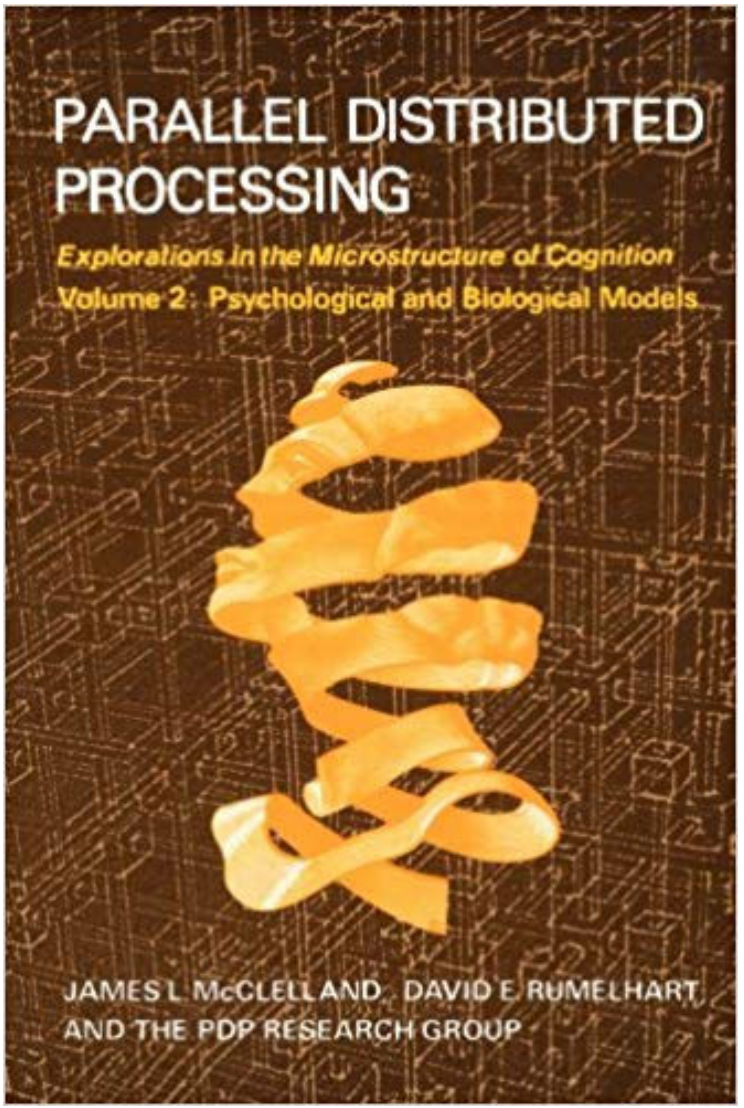
- IF
  - The current context is assigning devices to Unibus modules, and
  - There is an unassigned dual-port disk drive, and
  - The type of controller it requires is known, and
  - There are two such controllers, neither of which has any devices assigned to it, and
  - The number of devices that these controllers can support is known,
- THEN
  - Assign the disk drive to each of the controllers, and
  - Note that the two controllers have been associated and each supports one drive.

تاریخچه‌ی هوش مصنوعی

بازگشت شبکه‌های عصبی ...

THE RETURN OF NEURAL NETWORKS (1986–PRESENT)

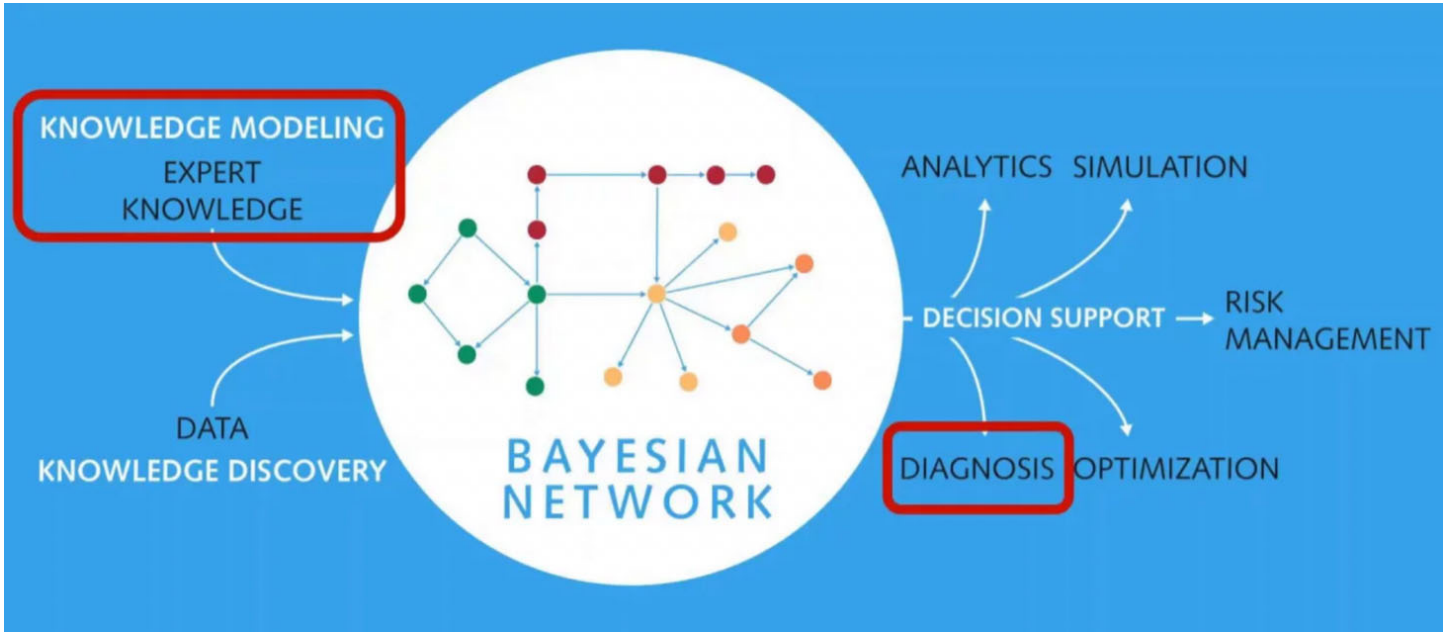


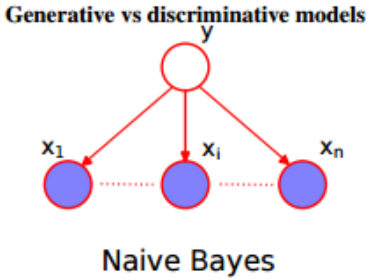

PARALLEL DISTRIBUTED PROCESSING
Explorations in the Microstructure of Cognition
Volume 2: Psychological and Biological Models
JAMES L. McCLELLAND, DAVID E. RUMELHART
AND THE PDP RESEARCH GROUP

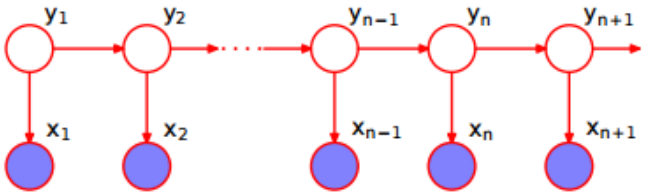
تاریخچه‌ی هوش مصنوعی

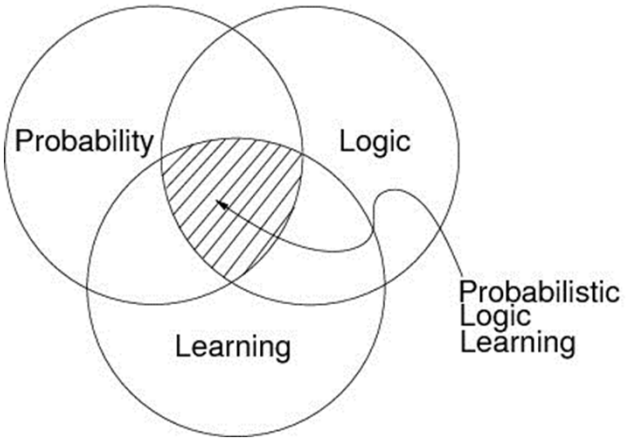
استدلال احتمالاتی و یادگیری ماشینی

PROBABILISTIC REASONING AND MACHINE LEARNING (1987–PRESENT)







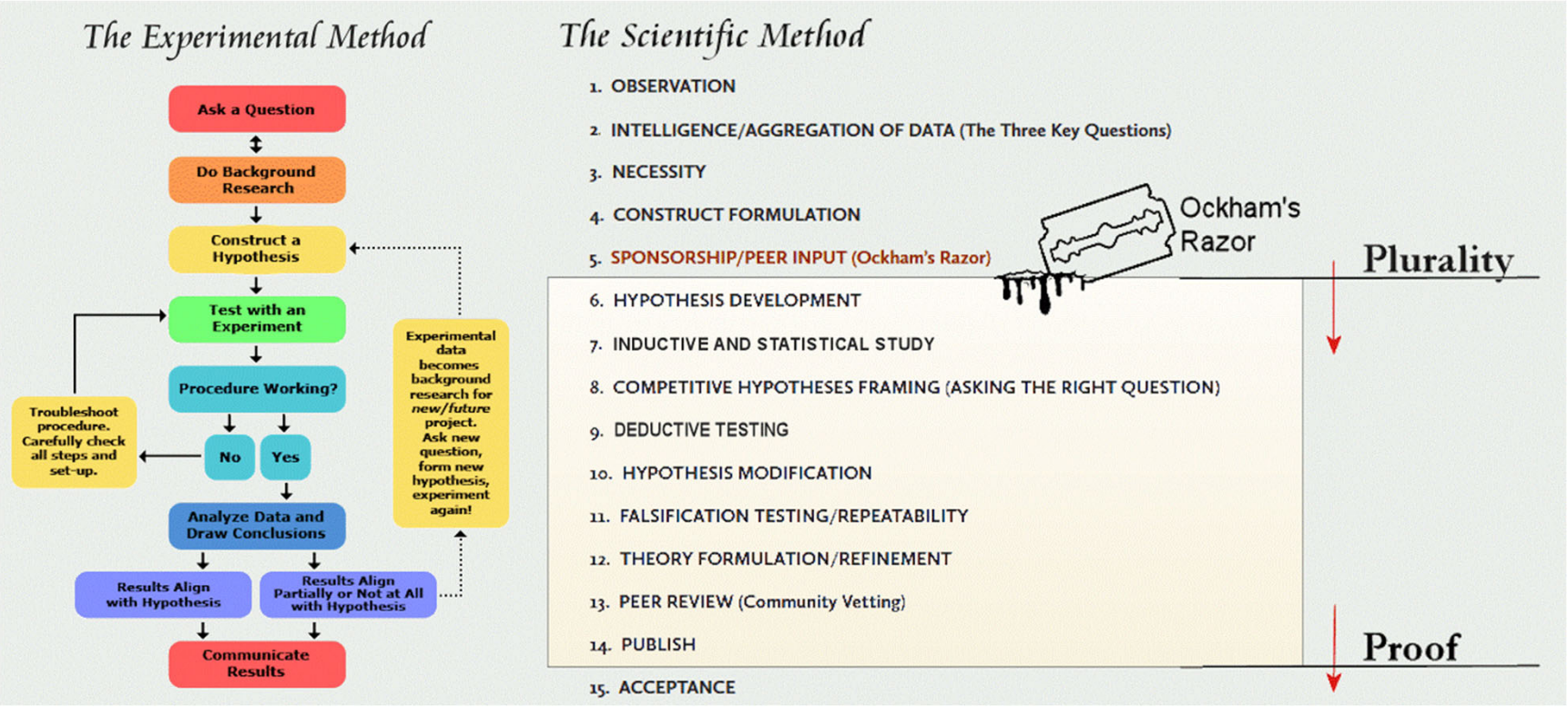


Judea Pearl (1936 -)

تاریخچه‌ی هوش مصنوعی

پذیرش روش علمی در هوش مصنوعی ...

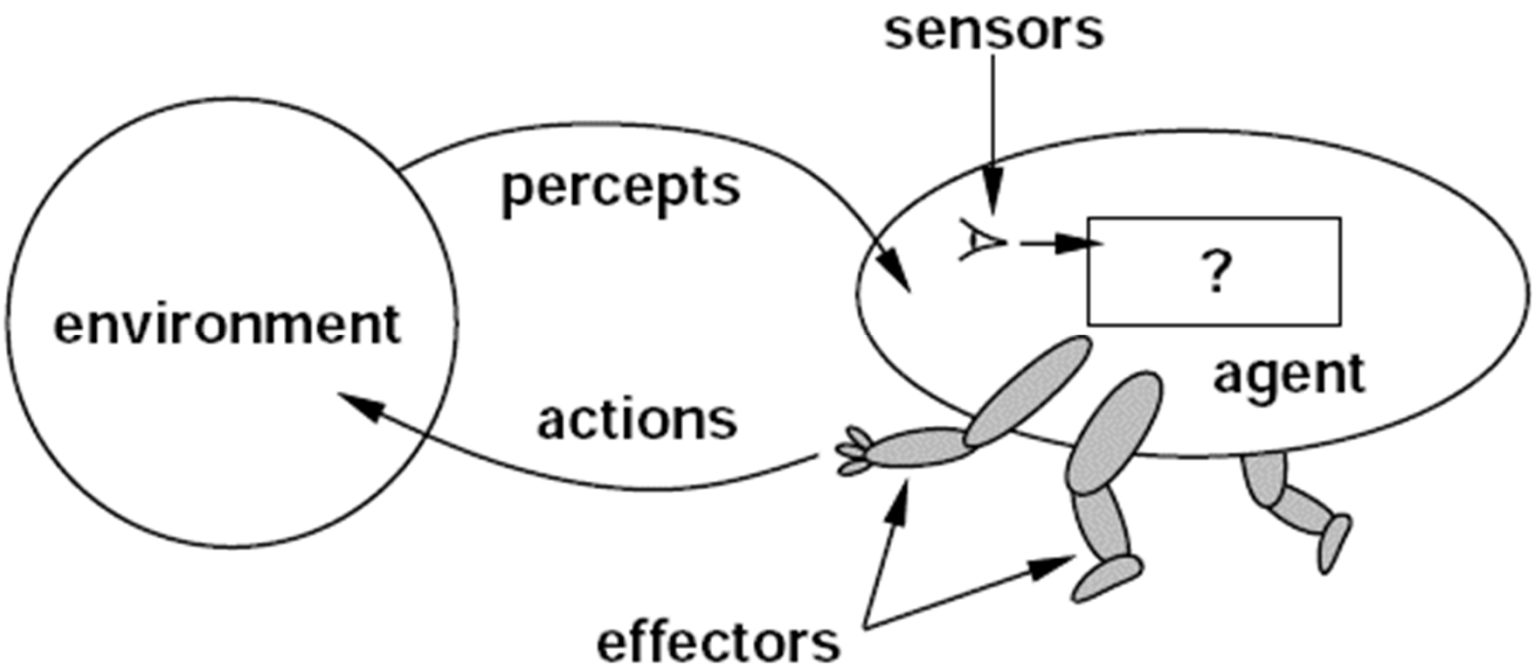
AI ADOPTS THE SCIENTIFIC METHOD (1987–PRESENT)



تاریخچه‌ی هوش مصنوعی

ظهور عامل‌های هوشمند ...

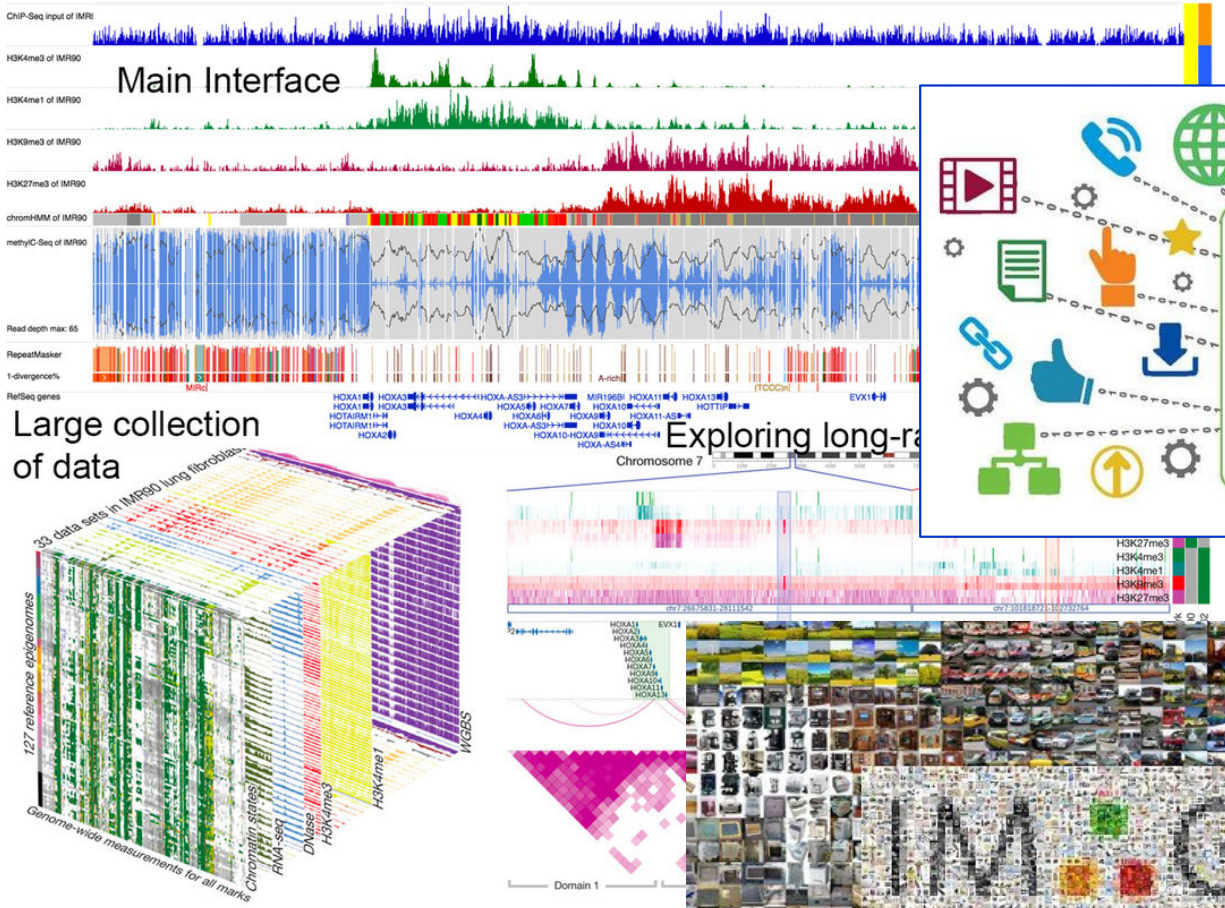
THE EMERGENCE OF INTELLIGENT AGENTS (1995–PRESENT)

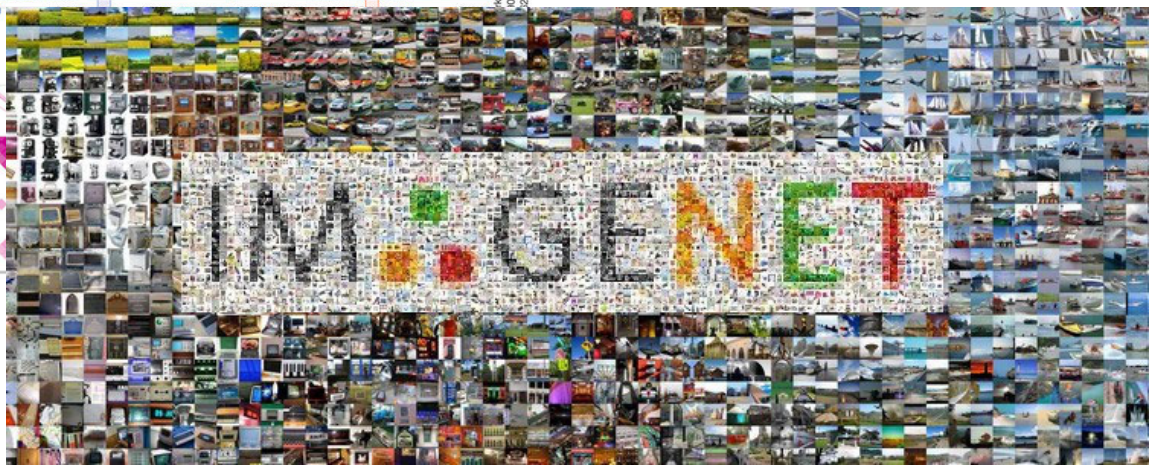


تاریخچه‌ی هوش مصنوعی

کلان داده‌ها: موجود شدن مجموعه داده‌های بسیار بزرگ ...

BIG DATA: THE AVAILABILITY OF VERY LARGE DATA SETS (2001–PRESENT)





1959

Data set of
1,000 cases

1965

Data sets of
10,000 cases
(e.g. observations)

1978

technology to analyse
500 to 5,000 variables
at the same time

1981

Statistics for **20,000 samples,** as histograms with 800 cells

1981

Software to handle
88,000 variables

1982

Moderate data sets have less than **500 cases** and large more than **2,000**

1986

Very large data sets referred as **11,000 cases**

1987

50,000 points, but the representation of 1 million or more data is feasible

1990

5,000 cases with 6 variables for immediate evaluation (c. 3 seconds.)

1991

1 million cases a very large data file

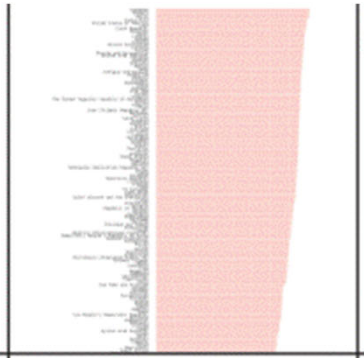
1996

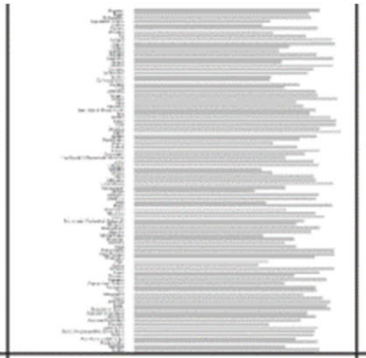
Large data sets have **60,000 or more observations**

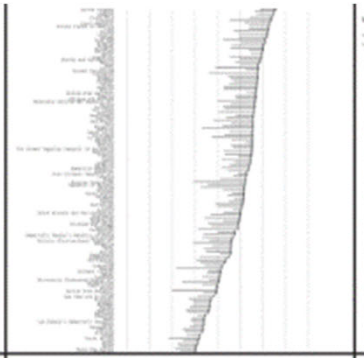
2006

1 million (bytes, cases, variables, combinations or tests) as a guideline to refer to a dataset as 'large'

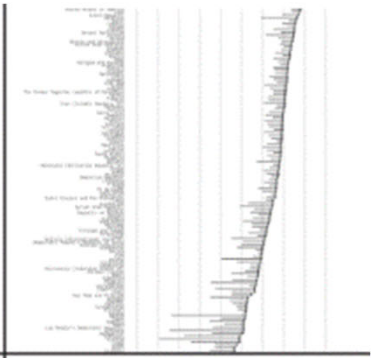

LIFE EXPECTANCY AT BIRTH, 2000-2015

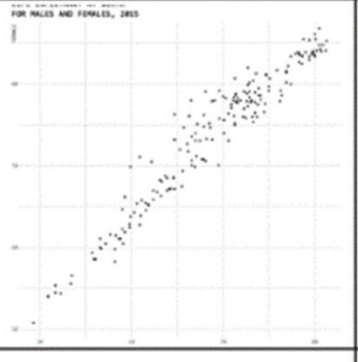


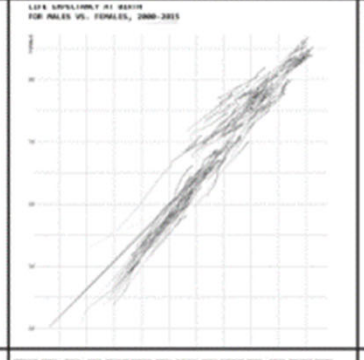


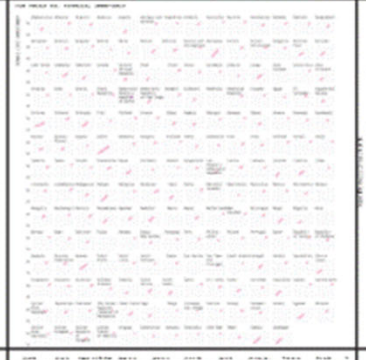


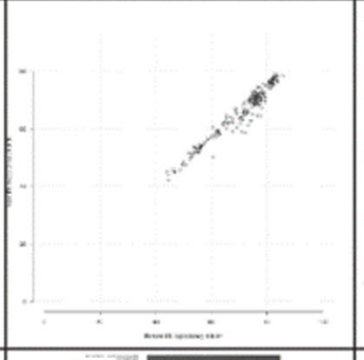

LIFE EXPECTANCY FOR MALE AND FEMALE, 2000-2015

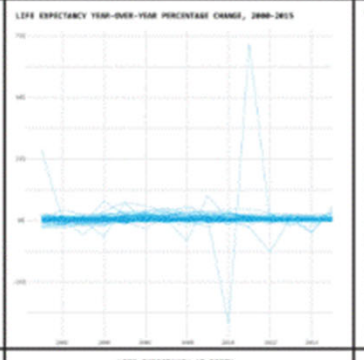


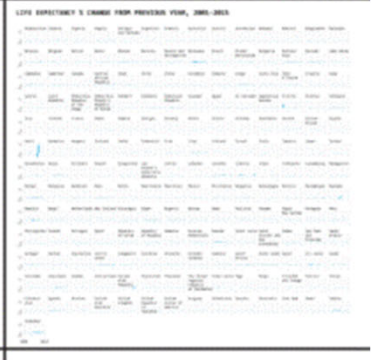

FOR MALES AND FEMALES, 2015


LIFE EXPECTANCY AT BIRTH
FOR MALES VS. FEMALES, 2000-2015

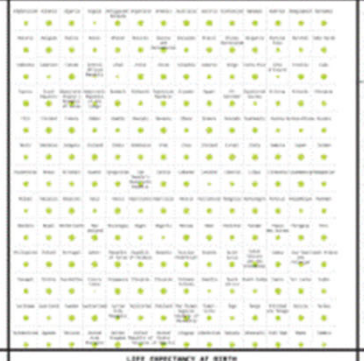


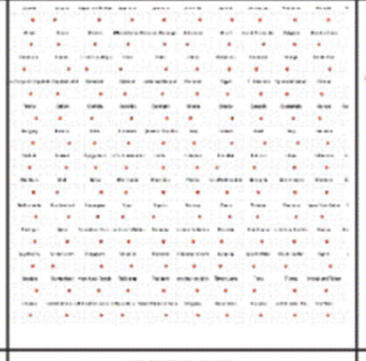


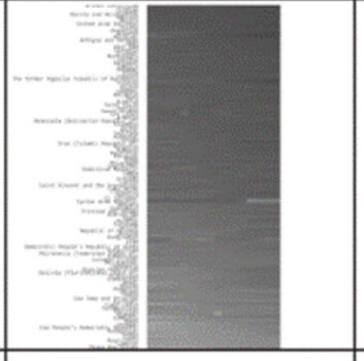

LIFE EXPECTANCY YEAR-OVER-YEAR PERCENTAGE CHANGE, 2000-2015

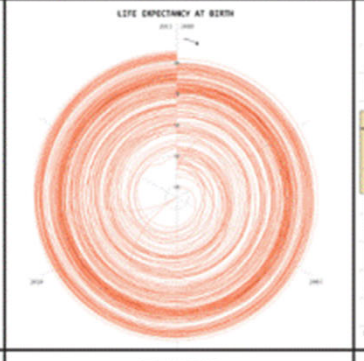

LIFE EXPECTANCY % CHANGE FROM PREVIOUS YEAR, 2001-2015


LIFE EXPECTANCY CHANGE FROM PREVIOUS YEAR, 2001-2015
Botswana
Eritrea
Haiti
Lesotho
Liberia
Myanmar
Rwanda
Zambia

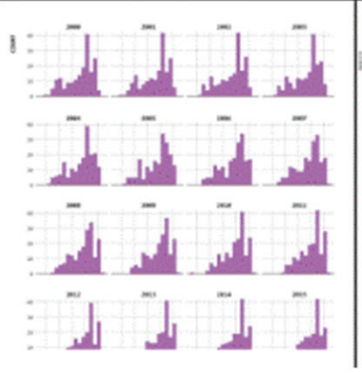


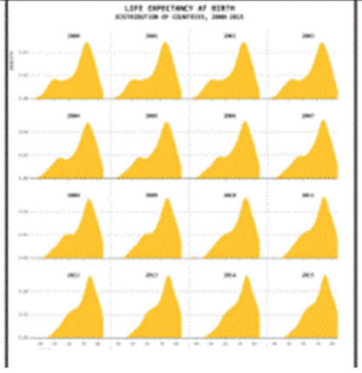


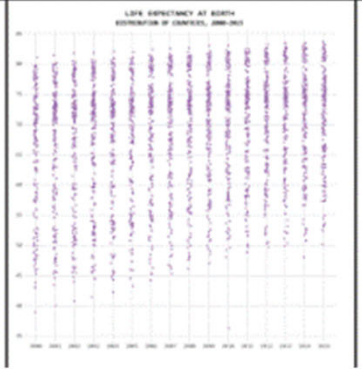


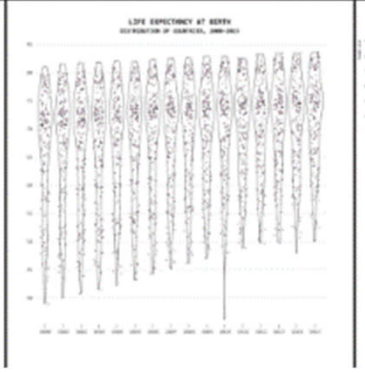

LIFE EXPECTANCY AT BIRTH

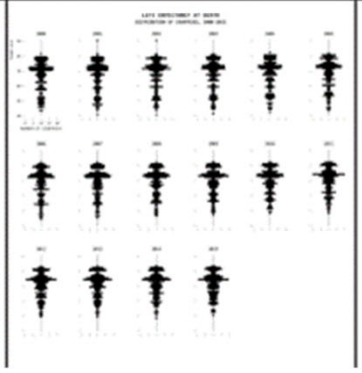

LIFE EXPECTANCY AT BIRTH
DISTRIBUTION OF COUNTRIES, 2000-2015

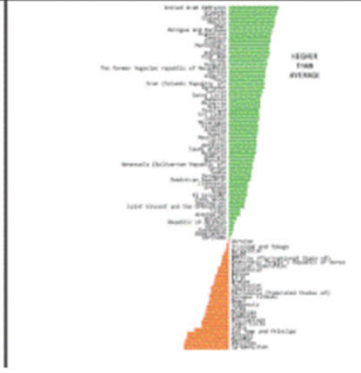



LIFE EXPECTANCY AT BIRTH
DISTRIBUTION OF COUNTRIES, 2000-2015


LIFE EXPECTANCY AT BIRTH
DISTRIBUTION OF COUNTRIES, 2000-2015


LIFE EXPECTANCY AT BIRTH
DISTRIBUTION OF COUNTRIES, 2000-2015

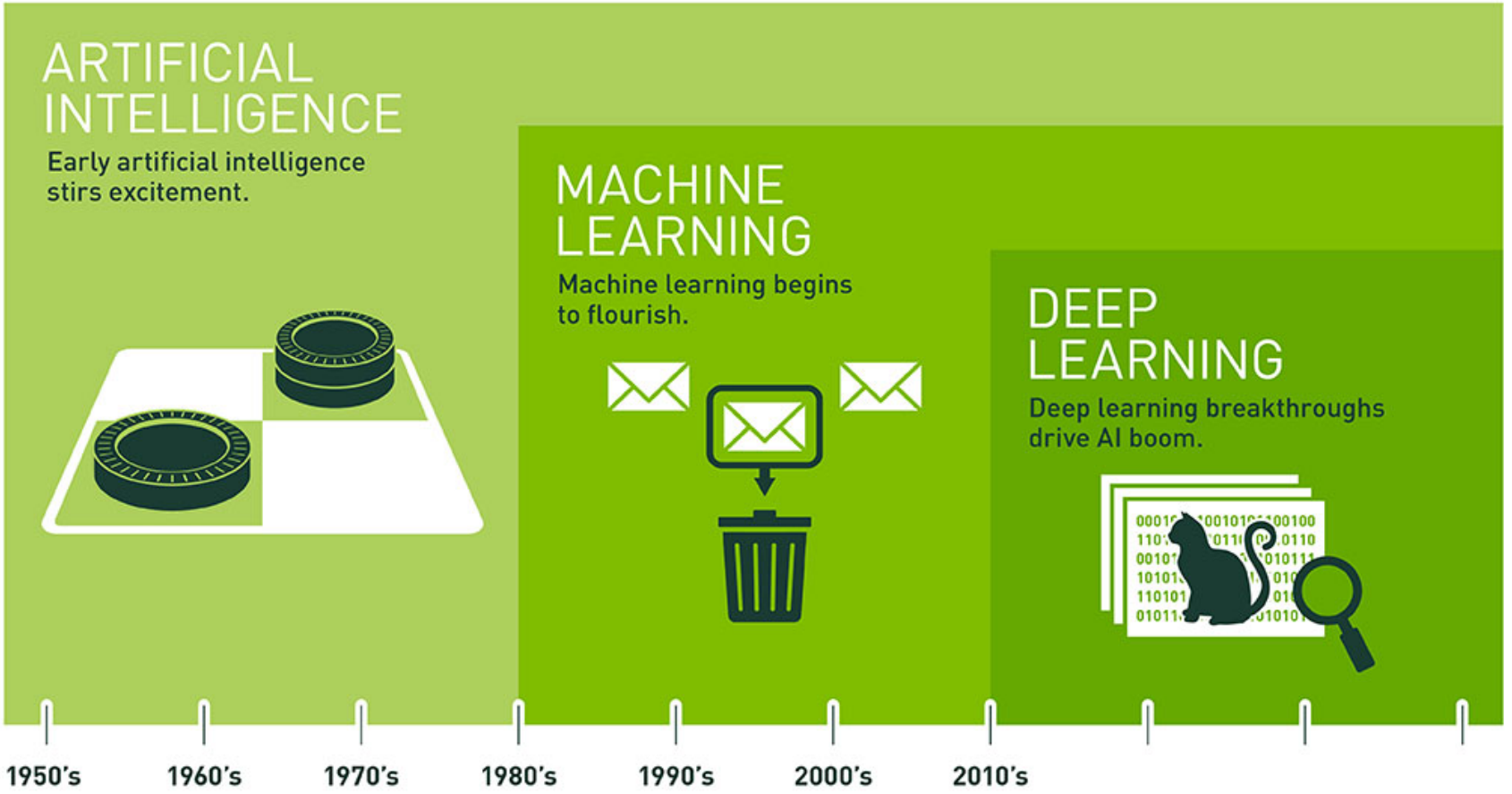

LIFE EXPECTANCY AT BIRTH
DISTRIBUTION OF COUNTRIES, 2000-2015


HIGHER THAN AVERAGE

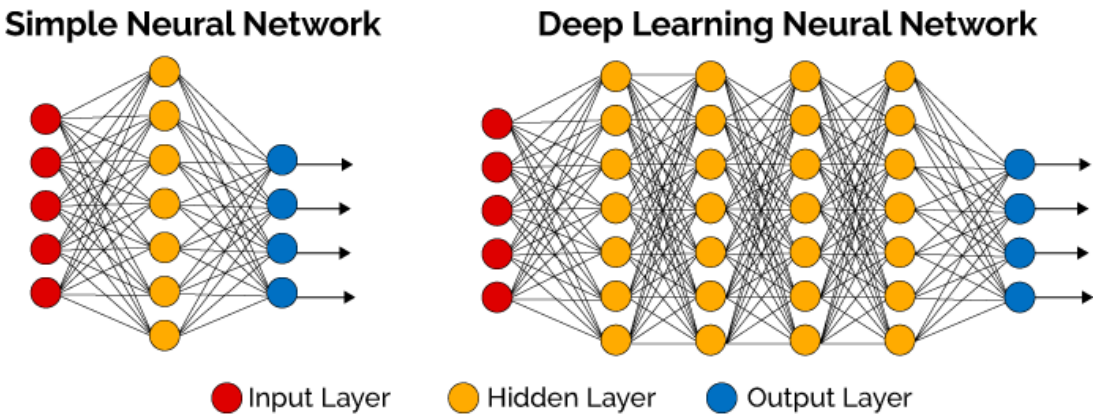
تاریخچه‌ی هوش مصنوعی

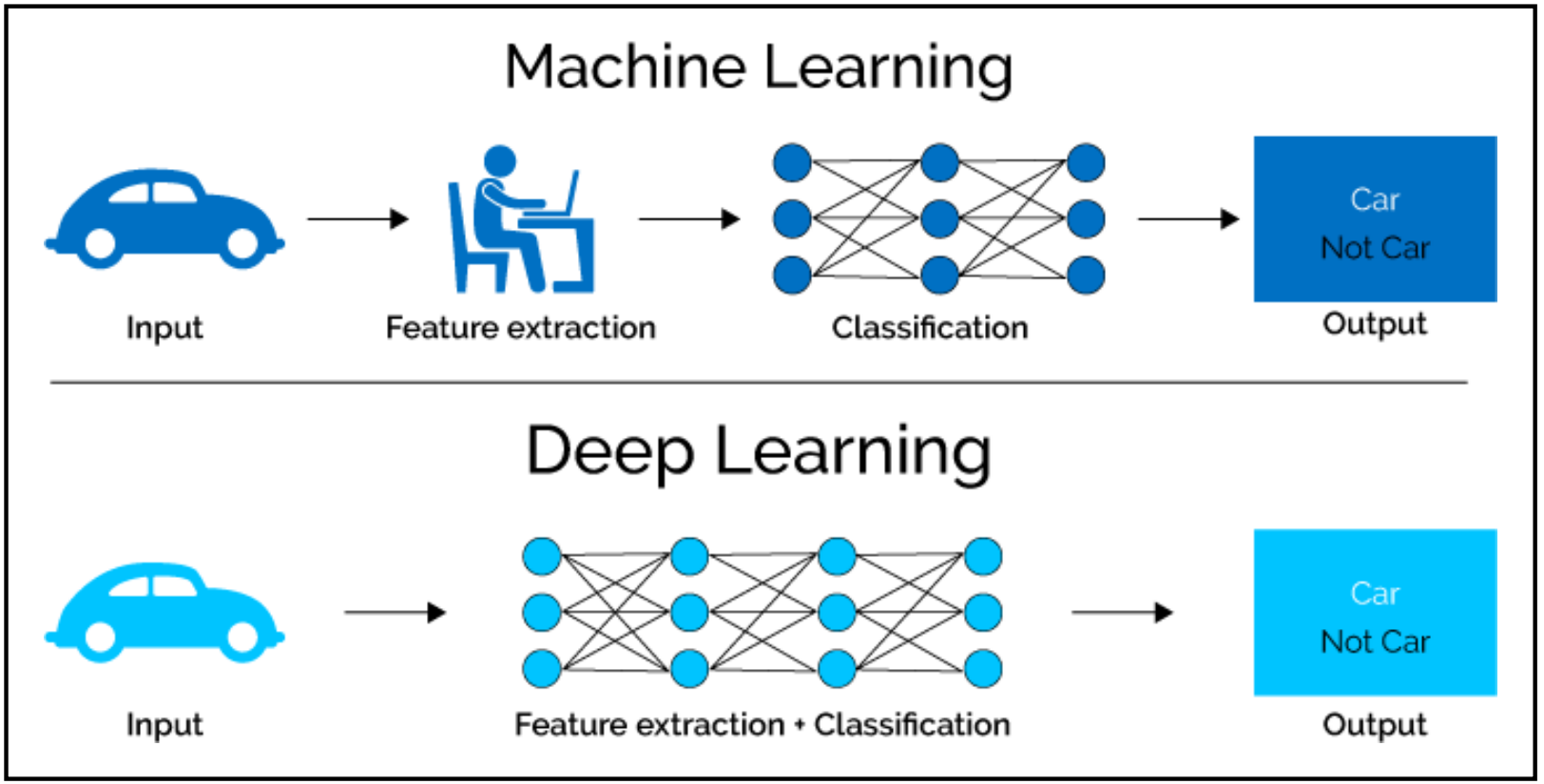
یادگیری عمیق

DEEP LEARNING (2011–PRESENT)



Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

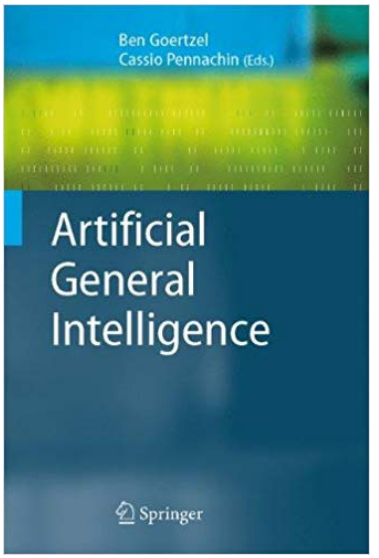

Simple Neural Network
Deep Learning Neural Network
Input Layer
Hidden Layer
Output Layer

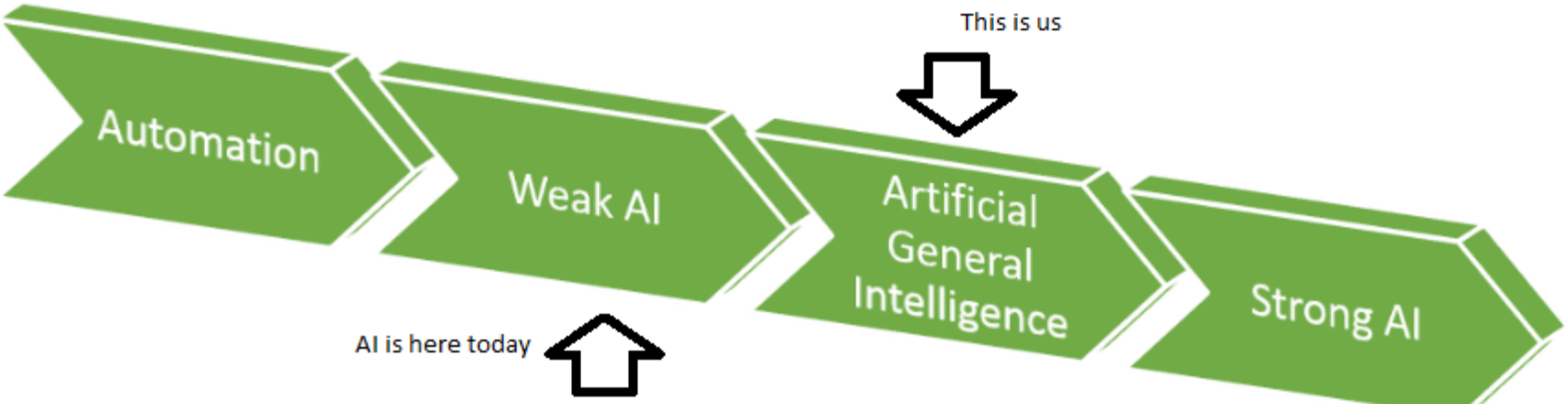

Machine Learning
Input
Feature extraction
Classification
Car
Not Car
Output
Deep Learning
Input
Feature extraction + Classification
Car
Not Car
Output

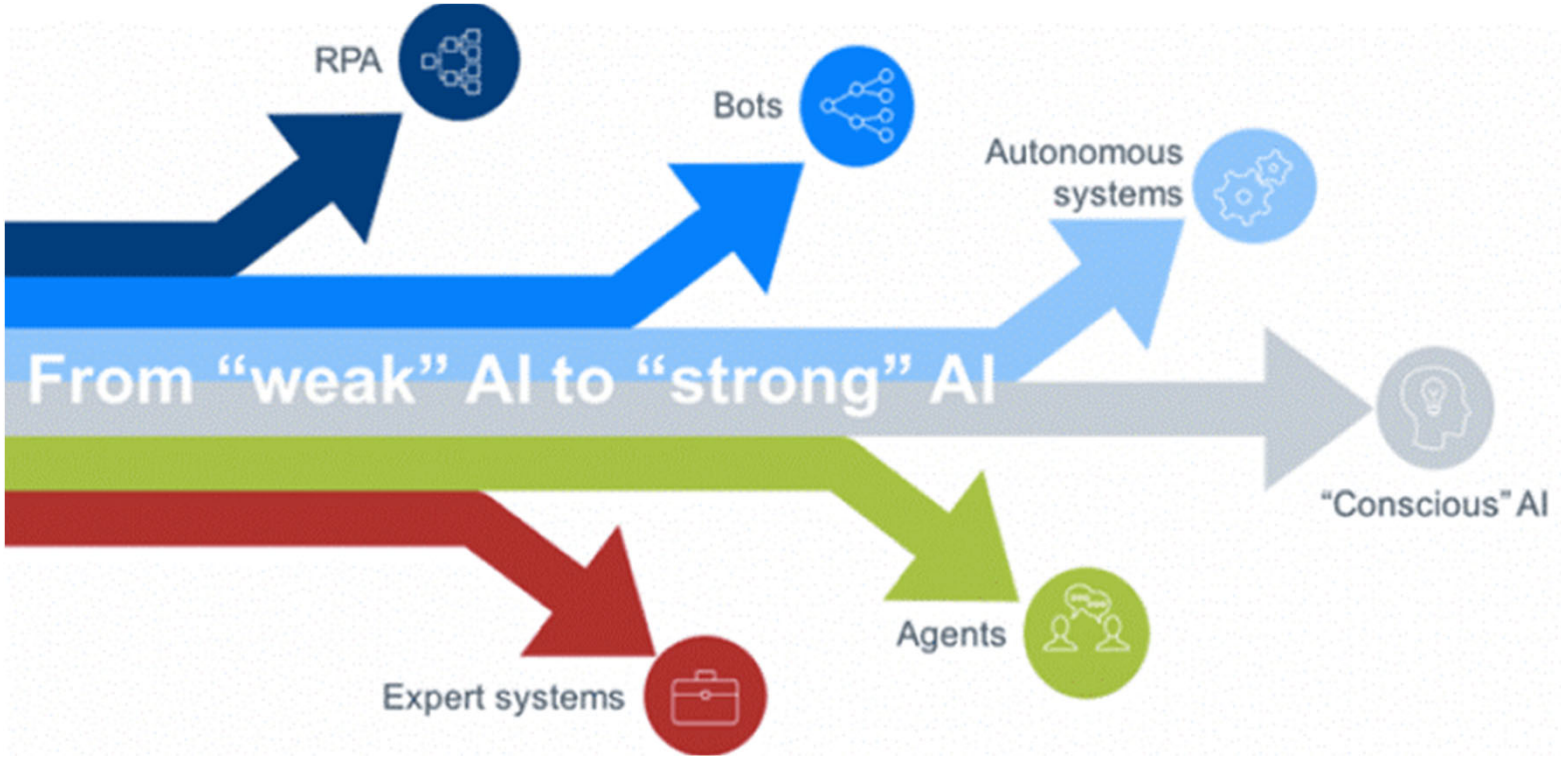
تاریخچه‌ی هوش مصنوعی

بازگشت هدف «هوش مصنوعی در سطح انسان» ...

HUMAN-LEVEL AI BACK ON THE AGENDA (2003–PRESENT)






RPA
Bots
Autonomous systems
From “weak” AI to “strong” AI
“Conscious” AI
Agents
Expert systems

Google DeepMind

Towards General Artificial Intelligence

Dr. Demis Hassabis, Founder & CEO, DeepMind
Twitter: @demishassabis

CENTER FOR
Brains
Minds+
Machines

April 20, 2016

Towards General
Artificial Intelligence

Demis Hassabis
Google DeepMind

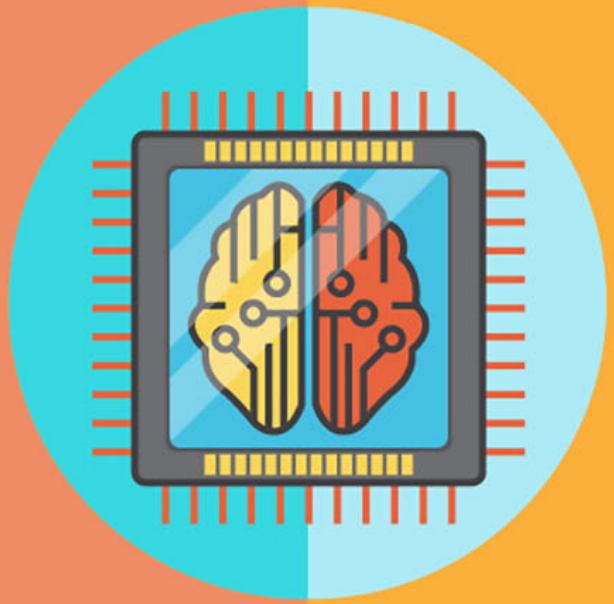
ARTIFICIAL NARROW INTELLIGENCE

VS

ARTIFICIAL GENERAL INTELLIGENCE

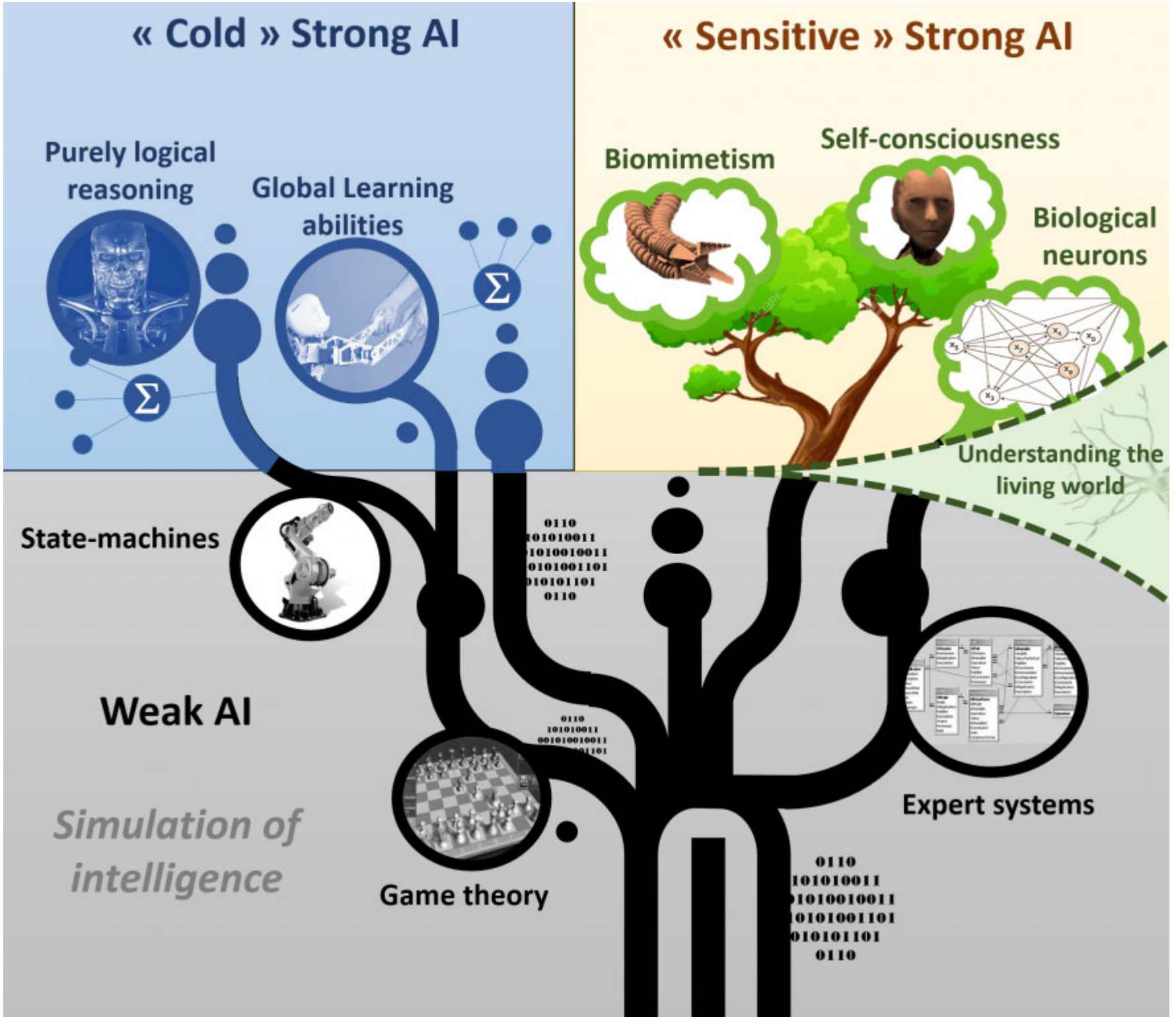
IDEA

Machine's ability to perform a single task extremely well, even better than humans.



IDEA

Machines can be made to think and function as human mind.


« Cold » Strong AI
« Sensitive » Strong AI
Purely logical reasoning
Global Learning abilities
Biomimetism
Self-consciousness
Biological neurons
Understanding the living world
State-machines
Weak AI
Simulation of intelligence
Game theory
Expert systems

THE RISE OF AI

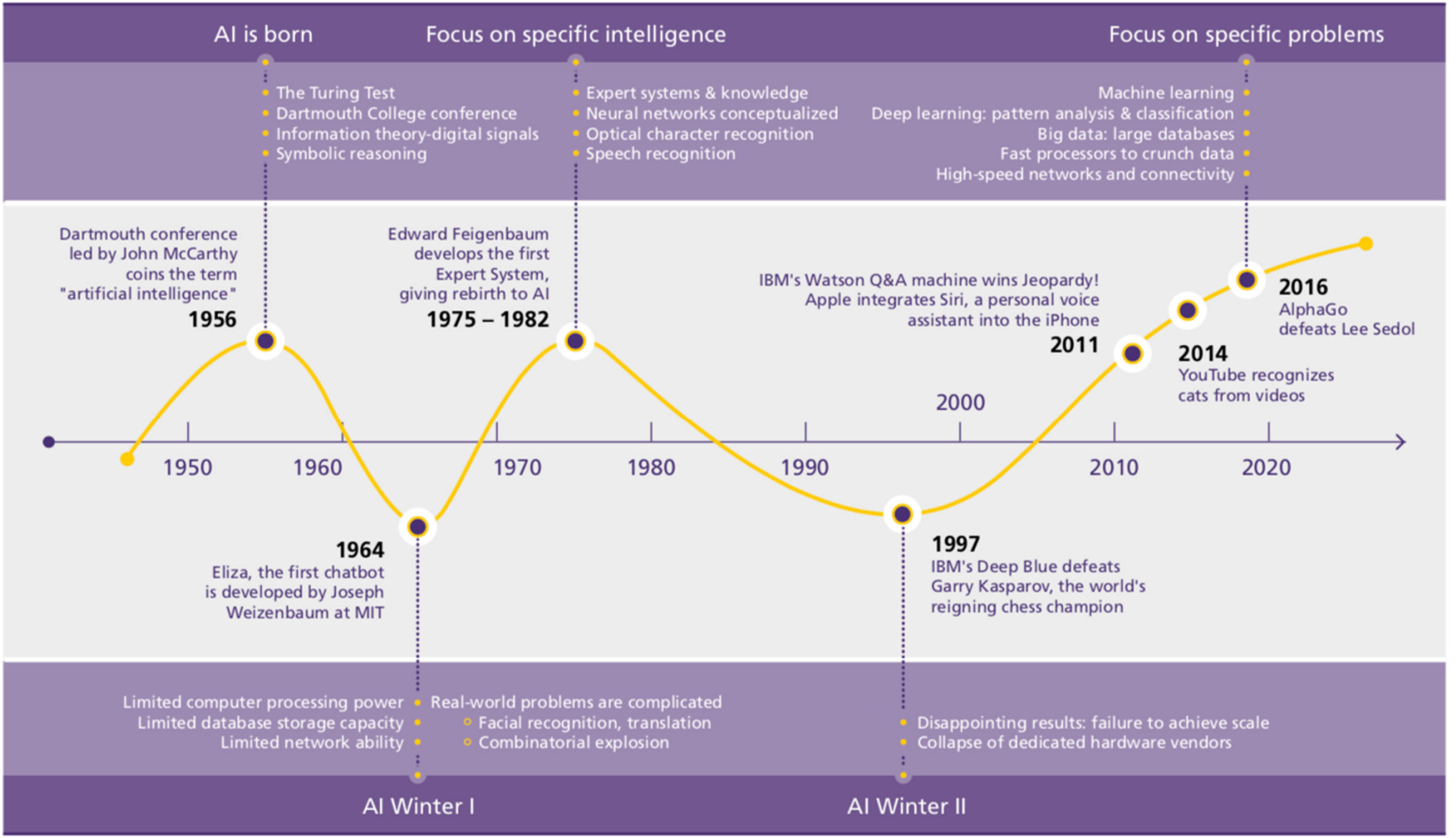
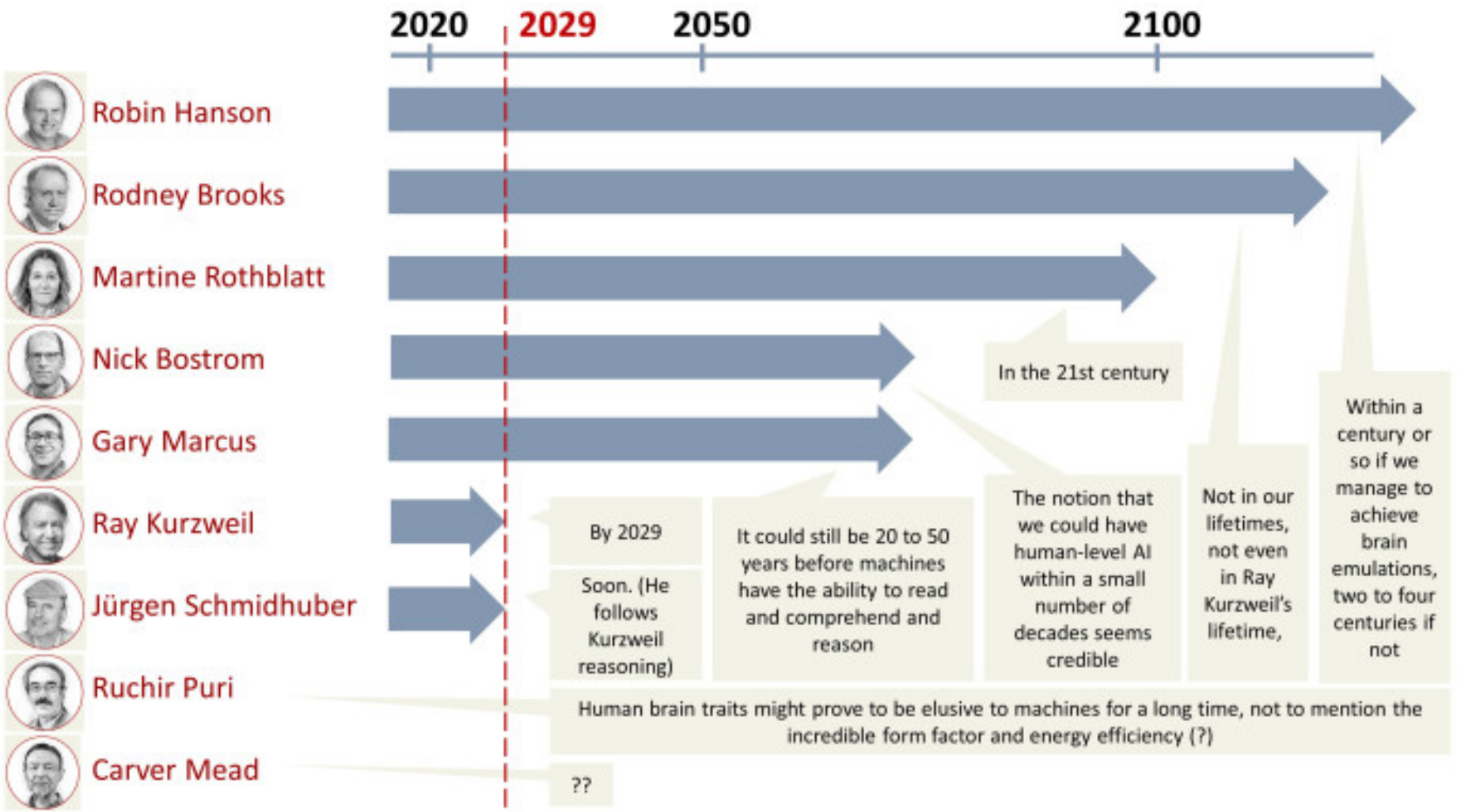


Figure 1: An AI timeline; **Source:** Lavenda, D./Marsden, P.

source dhl via @mikequindazzi


2020
2029
2050
2100
Robin Hanson
Rodney Brooks
Martine Rothblatt
Nick Bostrom
Gary Marcus
Ray Kurzweil
Jürgen Schmidhuber
Ruchir Puri
Carver Mead
In the 21st century
Within a century or so if we manage to achieve brain emulations, two to four centuries if not
Not in our lifetimes, not even in Ray Kurzweil's lifetime,
The notion that we could have human-level AI within a small number of decades seems credible
It could still be 20 to 50 years before machines have the ability to read and comprehend and reason
By 2029
Soon. (He follows Kurzweil reasoning)
Human brain traits might prove to be elusive to machines for a long time, not to mention the incredible form factor and energy efficiency (?)
??

هوش مصنوعی

درس ۴

مرزهای دانش کنونی هوش مصنوعی

AI: The State of the Art

کاظم فولادی قلعه
دانشکده مهندسی، دانشکدگان فارابی
دانشگاه تهران

http://courses.fouladi.ir/ai

هوش مصنوعی

مقدمه‌ای بر هوش مصنوعی

۴

مرزهای دانش در هوش مصنوعی

Stanford University SUNetID Login

One Hundred Year Study on Artificial Intelligence (AI100)

Home History People Framing 2016 Report Next Report AI Index Publications Press Coverage

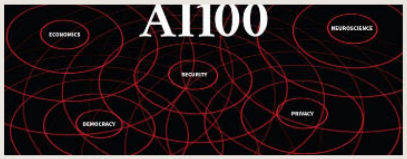
100 Year Study on AI kicks off its next Report cycle.

A set of study-workshops will inform the AI100 2020 Report.

Download the 2016 Report here

Learn more about our next Report here

About



The One Hundred Year Study on Artificial Intelligence, or AI100, is a 100-year effort to study

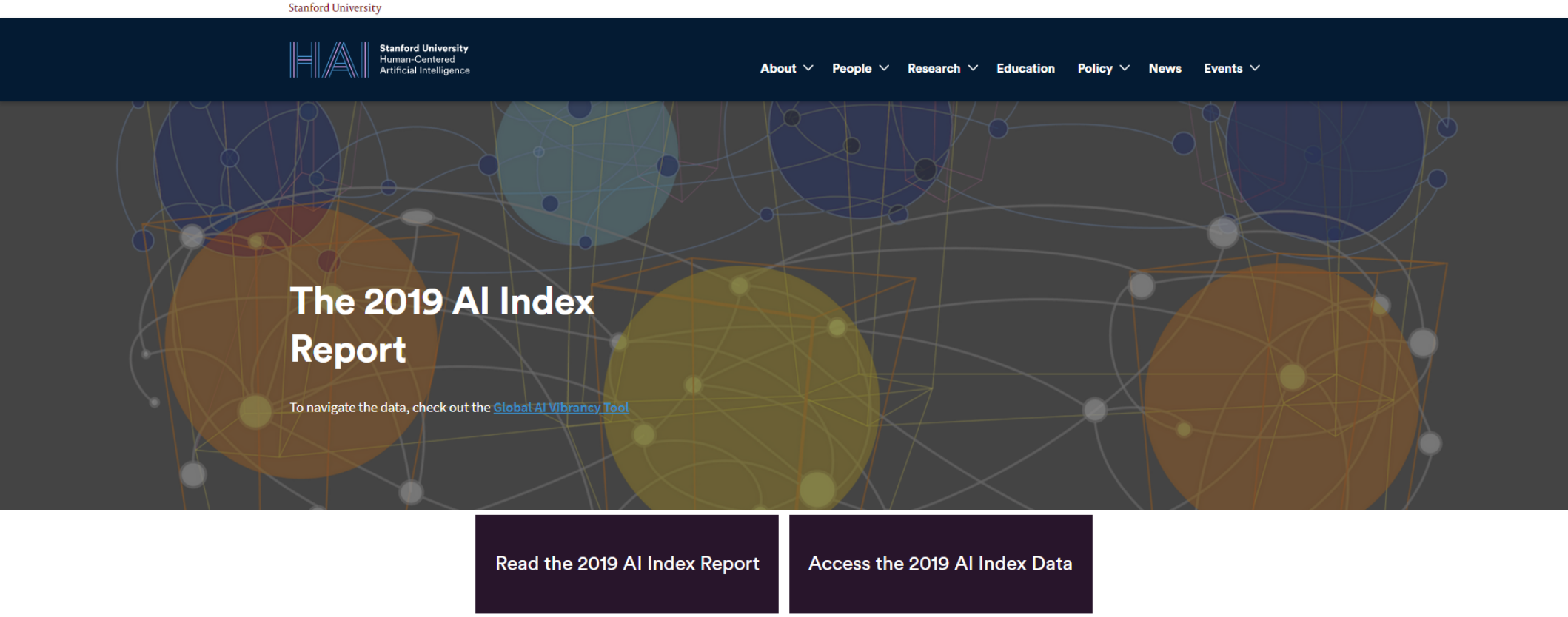
Reflections and Framing

One Hundred Year Study of Artificial Intelligence (AI) has its roots in a one-year study on Long-term

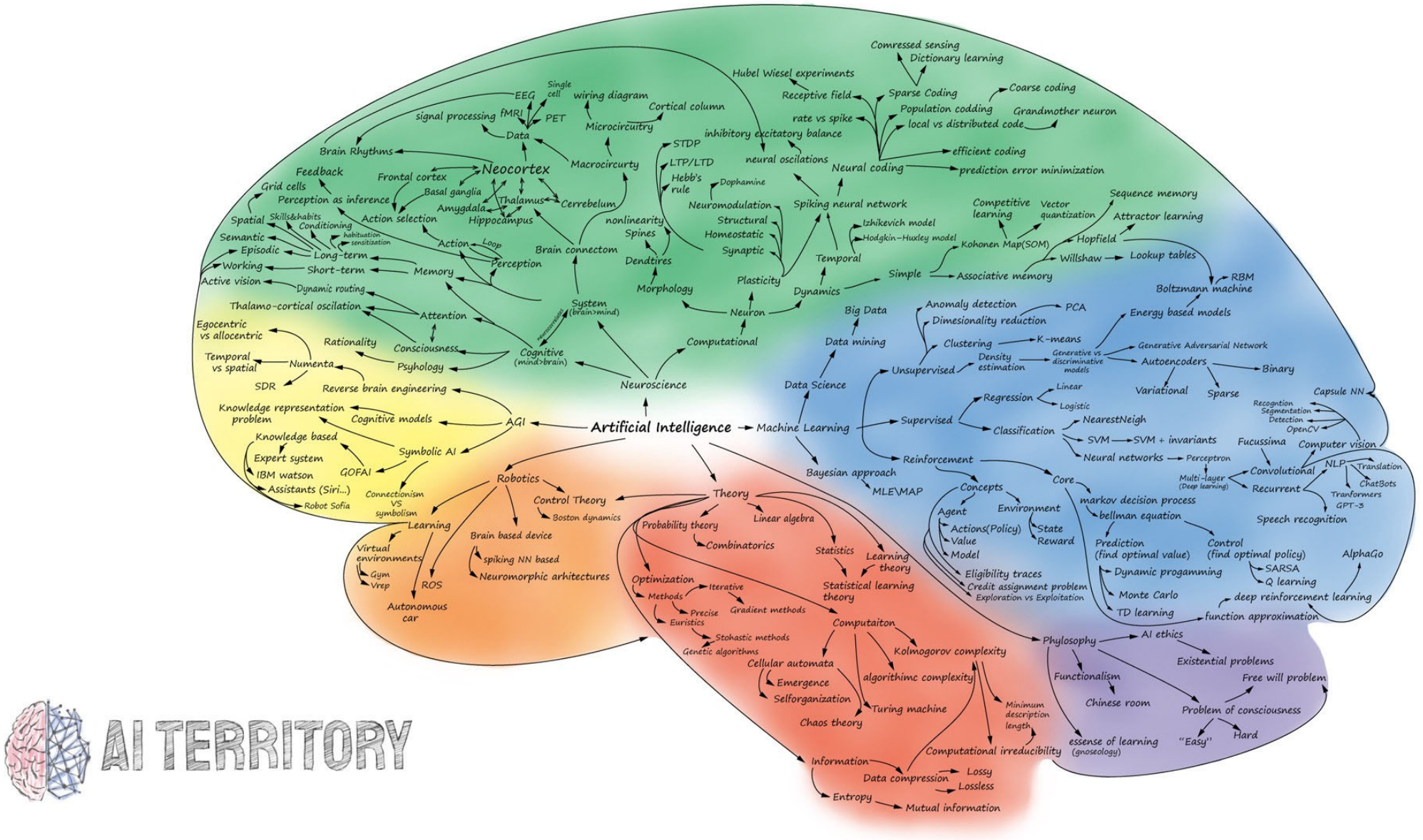
Committees and Panels

Seven researchers formed the first AI100 standing committee. It and subsequent committees

https://hai.stanford.edu/research/ai-index-2019



https://hai.stanford.edu/research/ai-index-2019



AI TERRITORY

مرزهای دانش در هوش مصنوعی

سال ۲۰۱۰

مبارزه با هرزنامه‌ها *Spam Fighting*	انجام بازی *Game Playing*	طرح‌ریزی و زمان‌بندی خودمختار *Autonomous Planning and Scheduling*	بازشناسی گفتار *Speech Recognition*	وسایل نقلیه رباتیک *Robotic Vehicles*
کنترل خودمختار *Autonomous Control*	تشخیص پزشکی *Medical Diagnosis*	ترجمه‌ی ماشینی *Machine Translation*	رباتیک *Robotics*	طرح‌ریزی آماد *Logistics Planning*

مرزهای دانش در هوش مصنوعی

سال ۲۰۲۰

بازشناسی گفتار *Speech Recognition*	ترجمه‌ی ماشینی *Machine Translation*	طرح‌ریزی و زمان‌بندی خودمختار *Autonomous Planning and Scheduling*	جابه‌جایی با پا *Legged Locomotion*	وسایل نقلیه رباتیک *Robotic Vehicles*
علوم اقلیم *Climate Science*	پزشکی *Medicine*	فهم تصویر *Image Understanding*	انجام بازی *Game Playing*	توصیه‌گرها *Recommenders*

Defense Advanced Research Projects Agency

هوش مصنوعی

مقدمه‌ای بر هوش مصنوعی

۵

ریسک‌ها و منافع هوش مصنوعی

ریسک‌ها و منافع هوش مصنوعی

سال ۲۰۲۰

RISKS AND BENEFITS OF AI

سلاح‌های خودمختار مرگبار *Lethal Autonomous Weapons*	مراقبت و تعقیب *Surveillance And Persuasion*	اتخاذ تصمیم سوگیری‌شده *Biased Decision Making*
تأثیر بر استخدام‌ها *Impact on Employment*	کاربردهای ایمنی-حیاتی *Safety-Critical Applications*	امنیت سایبری *Cybersecurity*

3 Types of Artificial Intelligence

Artificial Narrow Intelligence (ANI)

Stage-1

Machine Learning

- Specialises in one area and solves one problem


Siri


Alexa


Cortana

Artificial General Intelligence (AGI)

Stage-2

Machine Intelligence

- Refers to a computer that is as smart as a human across the board

Artificial Super Intelligence (ASI)

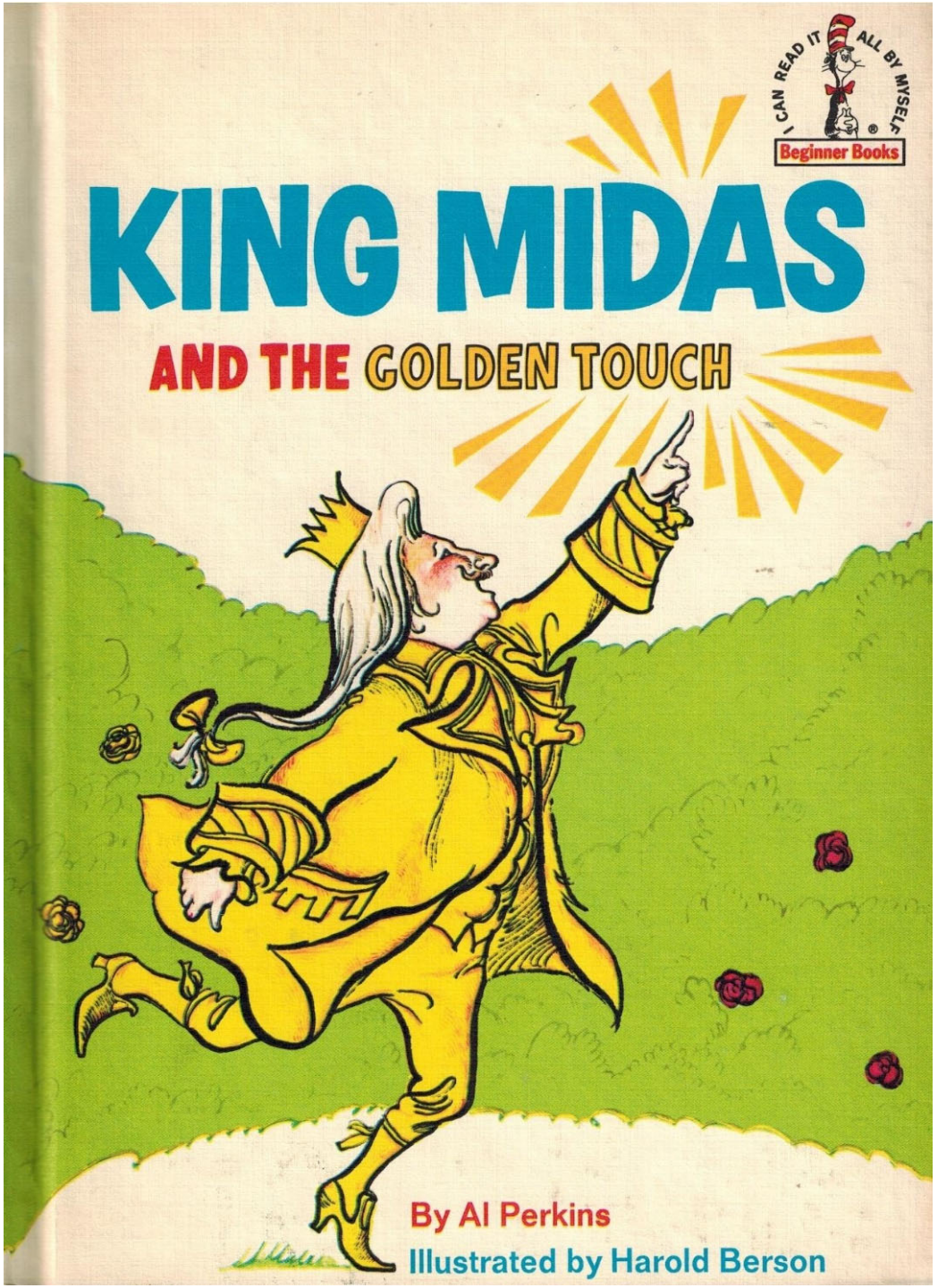
Stage-3

Machine Consciousness

- An intellect that is much smarter than the best human brains in practically every field

Weak AI	Strong AI
It is a narrow application with a limited scope.	It is a wider application with a more vast scope.
This application is good at specific tasks.	This application has an incredible human-level intelligence.
It uses supervised and unsupervised learning to process data.	It uses clustering and association to process data.
Example: Siri, Alexa.	Example: Advanced Robotics

Source: https://www.mygreatlearning.com/blog/what-is-artificial-intelligence/


I CAN READ IT ALL BY MYSELF
Beginner Books
KING MIDAS
AND THE GOLDEN TOUCH
By Al Perkins
Illustrated by Harold Berson


AI

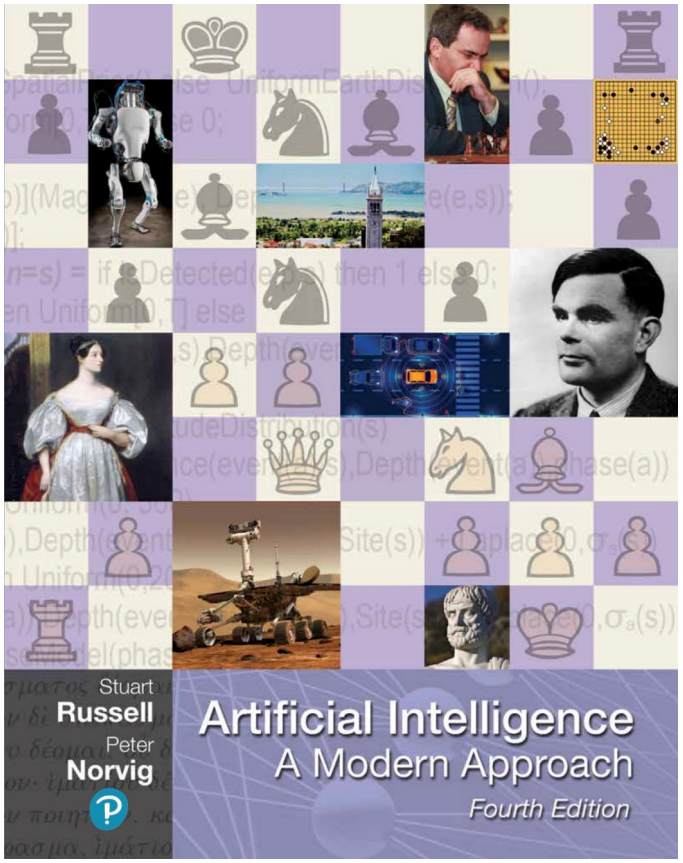
هوش مصنوعی

مقدمه‌ای بر هوش مصنوعی

۶

منابع، مطالعه، تکلیف

منبع اصلی



Stuart Russell and Peter Norvig,
Artificial Intelligence: A Modern Approach,
4th Edition, Prentice Hall, 2020.

Chapter 1

CHAPTER 1

INTRODUCTION

In which we try to explain why we consider artificial intelligence to be a subject most worthy of study, and in which we try to decide what exactly it is, this being a good thing to decide before embarking.

We call ourselves *Homo sapiens*—man the wise—because our **intelligence** is so important to us. For thousands of years, we have tried to understand *how we think and act*—that is, how our brain, a mere handful of matter, can perceive, understand, predict, and manipulate a world far larger and more complicated than itself. The field of **artificial intelligence**, or AI, is concerned with not just understanding but also *building* intelligent entities—machines that can compute how to act effectively and safely in a wide variety of novel situations. Intelligence Artificial intelligence

Surveys regularly rank AI as one of the most interesting and fastest-growing fields, and it is already generating over a trillion dollars a year in revenue. AI expert Kai-Fu Lee predicts that its impact will be "more than anything in the history of mankind." Moreover, the intellectual frontiers of AI are wide open. Whereas a student of an older science such as physics might feel that the best ideas have already been discovered by Galileo, Newton, Curie, Einstein, and the rest, AI still has many openings for full-time masterminds.

AI currently encompasses a huge variety of subfields, ranging from the general (learning, reasoning, perception, and so on) to the specific, such as playing chess, proving mathematical theorems, writing poetry, driving a car, or diagnosing diseases. AI is relevant to any intellectual task; it is truly a universal field.

1.1 What Is AI?

We have claimed that AI is interesting, but we have not said what it *is*. Historically, researchers have pursued several different versions of AI. Some have defined intelligence in terms of fidelity to *human* performance, while others prefer an abstract, formal definition of intelligence called **rationality**—loosely speaking, doing the "right thing." The subject matter itself also varies: some consider intelligence to be a property of internal *thought processes* and *reasoning*, while others focus on intelligent *behavior*, an external characterization.[1] Rationality

From these two dimensions—human vs. rational[2] and thought vs. behavior—there are four possible combinations, and there have been adherents and research programs for all

[1] In the public eye, there is sometimes confusion between the terms "artificial intelligence" and "machine learning." Machine learning is a subfield of AI that studies the ability to improve performance based on experience. Some AI systems use machine learning methods to achieve competence, but some do not.

[2] We are not suggesting that humans are "irrational" in the dictionary sense of "deprived of normal mental clarity." We are merely conceding that human decisions are not always mathematically perfect.