

実世界システムとしての人工知能

金出武雄

カーネギーメロン大学
ロボット研究所

人工知能における知能とは何か

結構定義は難しい: 曖昧 ないし 哲学的?

- 人が「知能的」と感じるもの?
 - チューリングテスト
 - “Eliza”効果
- 人工的能力が人間以上になると、人は「知能的」と感じない?
 - 仕組みが「完全」に明らかになったとき
 - 「決定的手順 (Deterministic Procedure)」が明らかになったとき
 - 問題の複雑さが計算機で解ける範囲に入ったとき
- 「いまだ機械で出来ない“知能的”な能力」という“定義”?
- 特に定義せず、「問題解決、視覚・音声・言語理解、計画、ゲームなど、人間と“同じような”あるいはそれ以上の“いわゆる知的”活動能力」とくらいにしておくほうが有用

人工知能の歴史：挫折と回帰

- 1966:ピアス勧告 — 「機械翻訳は出来ない」
- 1968:Minsky-Papert — 単層パーセプトロンの限界を指摘
 - 2層非線形要素(シグモイド関数)を含むニューラルネット
- 1973:ライトヒル勧告 — 組み合わせ爆発問題を指摘
 - NP Hard 問題に対する現実的な「近似的」「確率的」解法
 - 「知識」の書き出しと「浅い」探索によるエキスパートシステムの成功
- 1990年代:常識を取り入れる困難さに直面
 - 第五世代プロジェクトの終了と
 - DAPRAが AIへの“Umbrella Funding” を停止
 - 自動運転などの「実世界システム」の開発
 - 機械学習、中でも“Deep Learning”
 - “Big Data” の活用

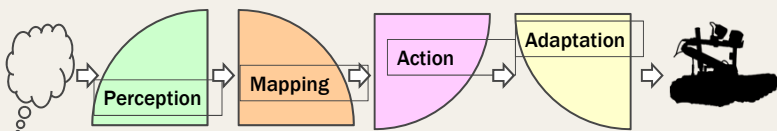
AI ブーム比較

ブーム	“一般”推論機構	人間に勝つゲーム	ロボット	言語・音声・画像	神経回路網？	代表的開発目標AIシステムとその“限界”
1	GPS 述語論理	チェッカー	Hand Eye Project	パターン認識	Perceptron	機械翻訳 複数意味(セマンティックス)計算量の爆発
2	Rule-based Reasoning AIプログラミング言語	チェス	自動運転	Language/Speech/Image “Understanding”	2層非線形NN	エキスパートシステム 知識の獲得 “常識”(Frame Problem)
3	ビッグデータ “推論”	基	Humanoid	画像・音声認識の実用化	Deep NN	QA ?

各段階で指摘された問題点は、いまだ完全に解決されたわけではない(永久に解決されないかもしれない)が、その意味や現実的対処法が少しずつわかってきたということ

Autonomous Robots

実世界とインターアクションするAIシステム
 “Robotics is where AI meets the real world”



Analytics and Machine Learning (“狭い”AI)

How fast and big
a program can
learn

Computing: Processing and Data Kernel

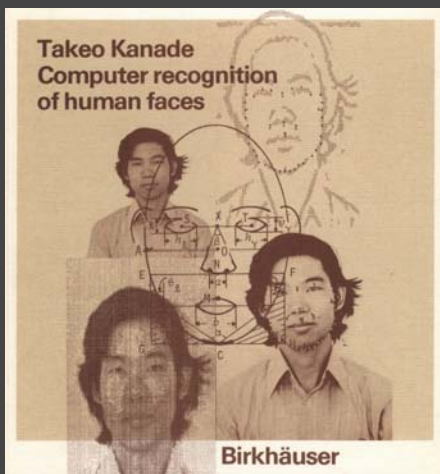
How fast, big,
and organized a
computer can
process and
store

明らかな進歩

わたしが係わった例：

- 顔認識
- 自動運転

顔認識 — 40年前



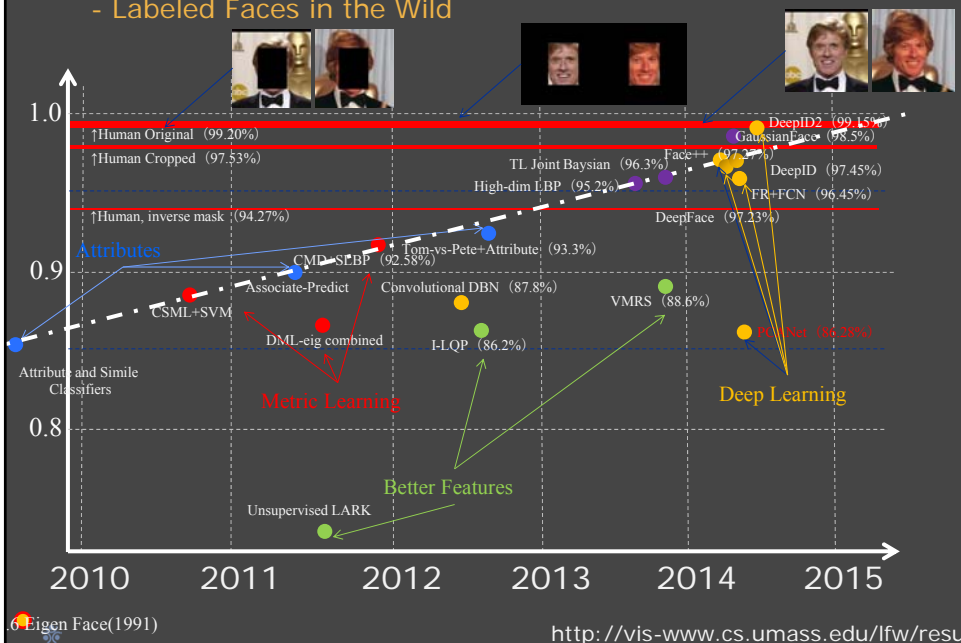
金出博士論文 (1973)

Processed 1000 images
collected at World Expo 70

The "world-first complete"
face recognition program
(1991 NSF report).

Face Recognition Today

- Labeled Faces in the Wild



Autonomous Driving – 30 Years Ago



Navlab 1: 1985-
Park pathway
~20 km/hr, 10km
Cameras, Laser range
Sun 3s and 4s



Today: Autonomous Car Becoming Real

Google Car on Public Roads



Uber Automated Taxi Project (2016)

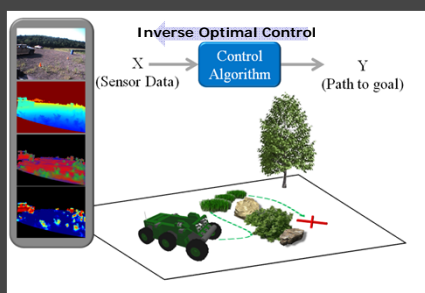


現在：本当にシーンを“理解”する車にむけて



Courtesy of Munoz, Hebert

Learning by showing : ストラテジーの学習



Inverse optimal control method – learning the strategy, not the result

[Bagnell, Hebert]

Courtesy of Bagnell



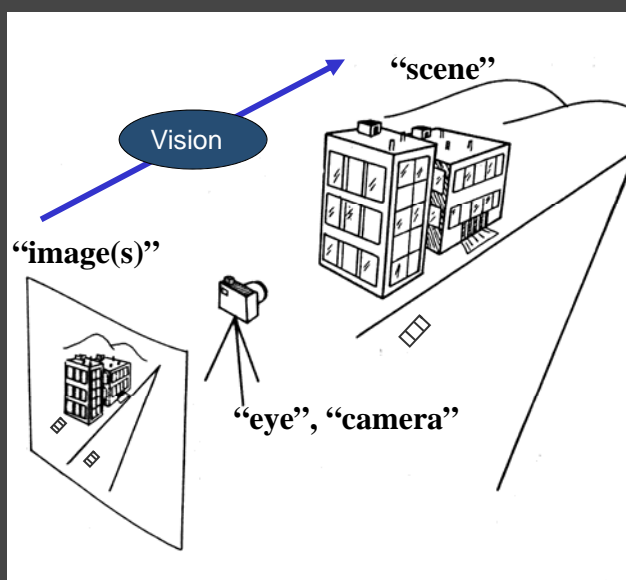
実世界に根ざした 人工知能研究の本質はなにか

コンピュータビジョンを例に使いながら

本質的困難さ、
解決されなければならないこと、

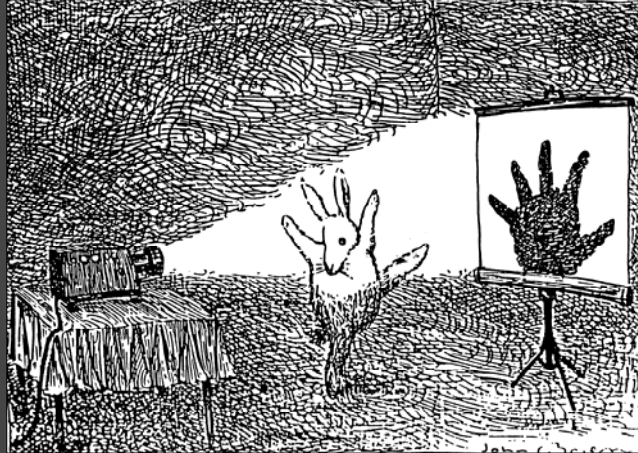
を議論する

What is Vision?



Guessing reality from projection

--- Obviously, a difficult problem

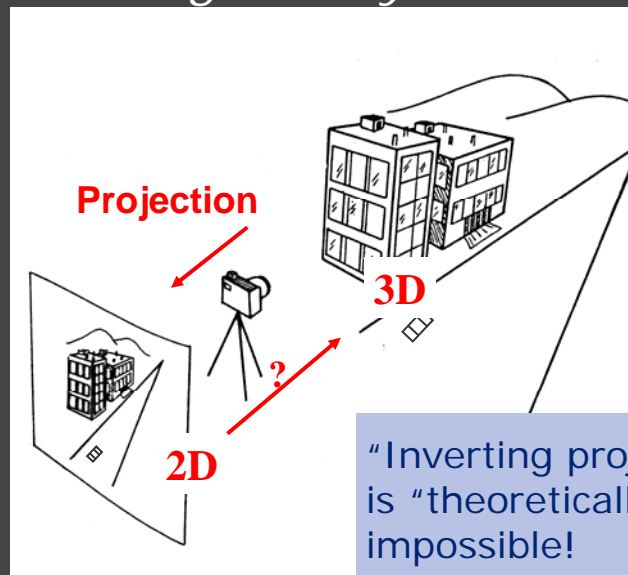


Computer Vision was one of the earliest problems in Artificial Intelligence: e.g. Blocks World, Hand-eye Project
It turns out to be difficult - Minsky's episode



Why vision is difficult:

2D-3D Degeneracy



ほとんどの知能問題は“解“がよく定義されていないか、複数答えがある

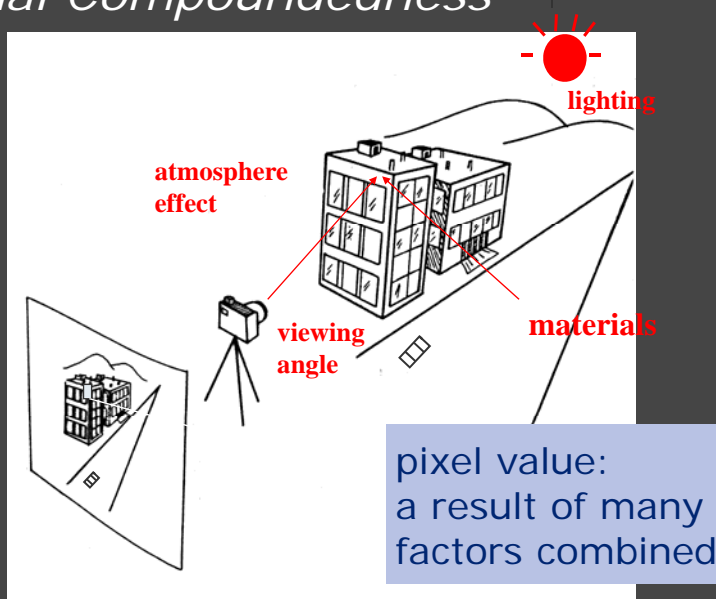
- 1 2 5 6 、 、
の答えは、実は何を入れても「正解」
- “Time flies like an arrow.”の意味
- 「宣教師とライオン」のクイズ
- 最も“簡単な“、“美しい“答え、答えのあるべき範囲などを陰に規定している
- 一般に「枠組み問題」という

曖昧さを取り除く方法の研究

- 物理現象(光の反射や撮像)をモデル化する
- 人が使っているらしい Implicit (陰)に仮定されている条件を見つけ出す。
- 「普通」の条件 (Probably Situation) に起こっている性質を使って曖昧さを取り除く
→ Domain Knowledge (領域特有の知識) の獲得
→ 人間の解答のデータ (Big Data) をまねする —
現象論的知能 (aka 学習)

Why vision is difficult:

Signal Compoundedness

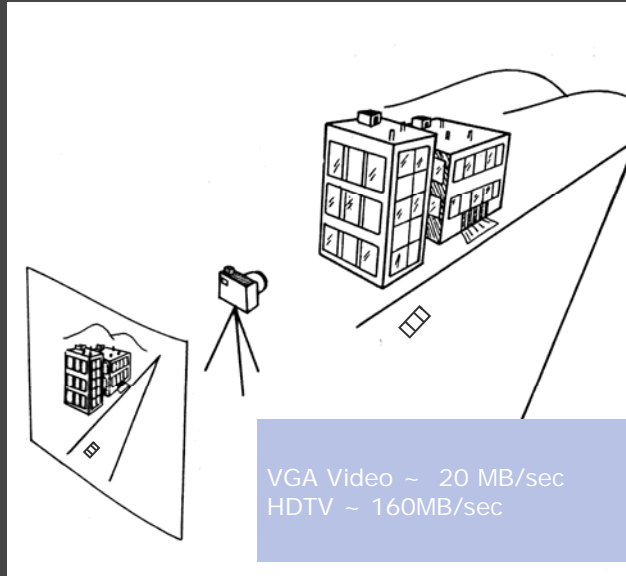


物理世界の法則に従わなければならない

- 物理法則はほとんどが、“Forward” モデル(こうすれば何が起こる)。ただし、解釈やActionの計画には “Backward” モデル(何を起すにはどうするか)がいる
- 計測(センシング)には必ず、Uncertaintyが伴う
- Signal to Symbol Mismatch (ほとんどの知識はシンボルを使って書かれている。観測をそのレベルに信頼性良く上げることが難しい)

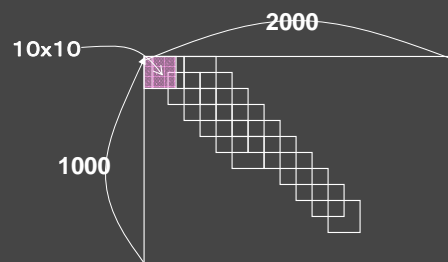
Why vision is difficult:

Sheer number: *Data and Computation*



A Simple Fact

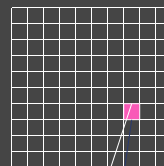
- A large image consists of a large number of small images.
- A HDTV 2000x1000-pixel image contains
~ 2×10^6 10x10-pixel images



Another simple fact

The number of possible
10x10-pixel images is:
(assuming 8-bits/pixel)

$$256^{10 \times 10} = 10^{240}$$



8 bits
→ $2^8 = 256$
possibilities

$$>> 10^6 \times 10^{10} \times 365 \times 24 \times 60 \times 60 \times 30 \times 10^6 = 10^{31}$$

**Human kind has not yet seen all of
10x10-pixel images**

(Kanade's theorem?)

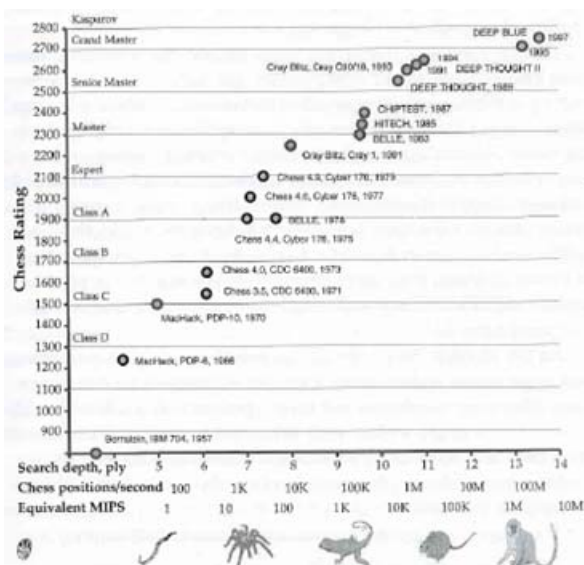
→ We need fairly bold assumptions!



有限問題(多くのゲーム)でも難しい

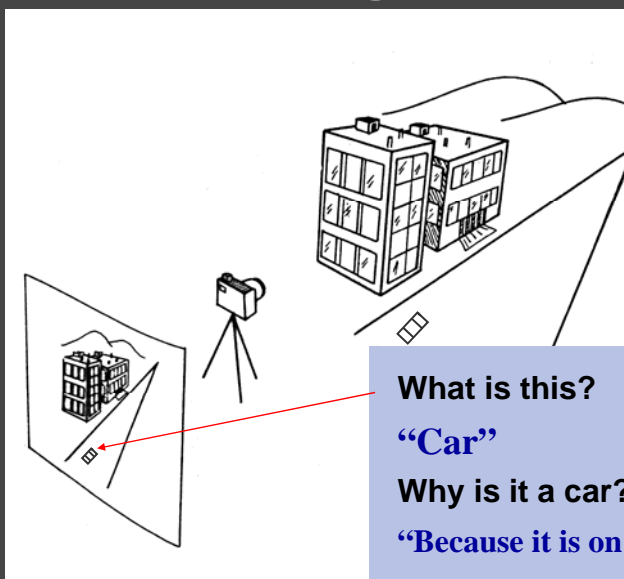
- 解の総数が有限の問題 → 必ず解けるはず
- しかし、NP困難(あるいはNP完全)な問題がほとんど → 問題の大きさに対して 指数関数的に解の困難さが増す(例:トラベリングセールスマン問題)
- しかし、ある程度以上の計算能力があれば、現実的に(近似的、確率的に)解く方法が見つけられている
 - 新しい最適化手法、ランダムアルゴリズム、整数計画法など
- 計算機の能力が上れば解ける問題が急速に増える

チェスマシンの計算能力とレーティング

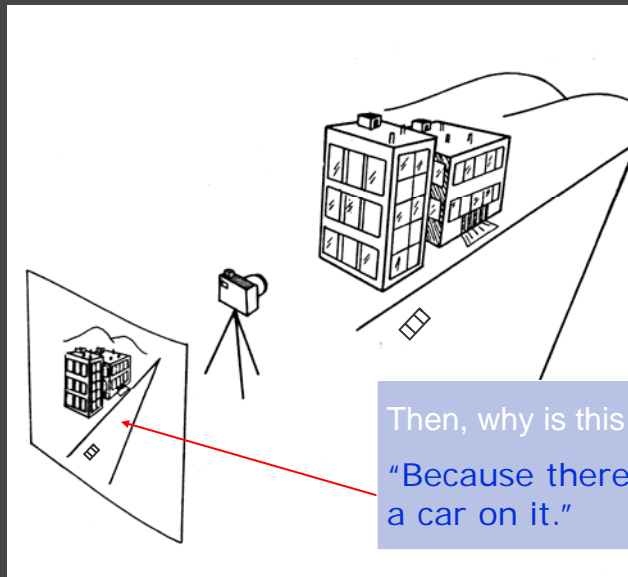
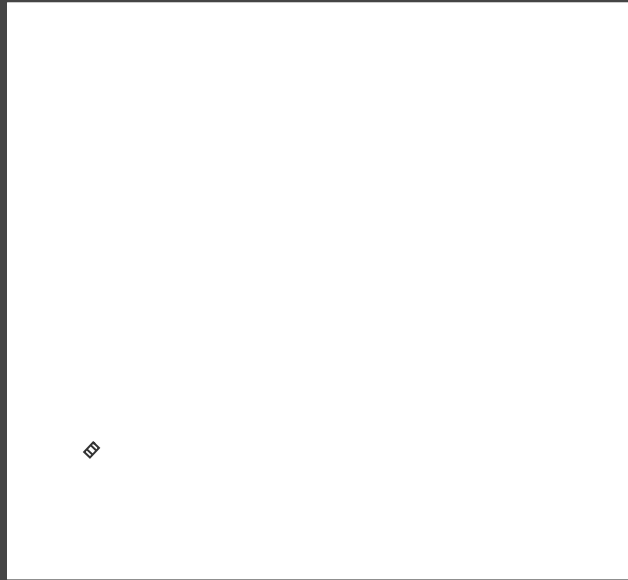


Why vision is difficult:

Context, Knowledge



Context is indeed playing a role



Then, why is this road?
"Because there is
a car on it."

一般的な知識の獲得

- 知識は必ずしも論理的ではないHeuristicなものであり、人から学ぶ必要がある。
 - 初期の頃は「人手」による個々のコーディング
 - Cyc Project 人手による集積(84から)
 - 機械学習の応用(現在)
 - インターネットは「知識」の宝庫
 - Internet からの知識の自動獲得
- 例: NELL and NEIL (Never Ending Language and Image Learner)

NEIL

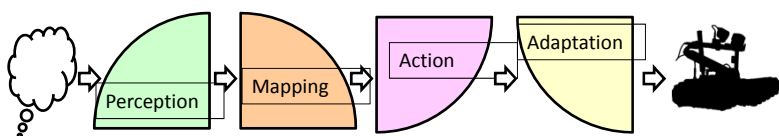
**NEVER ENDING IMAGE LEARNER
BY DR. ABHINAV GUPTA**

Use images on the web
To Automatically build world's largest
Visual Knowledge Base
Running 24 hours a day 7days a week
To learn Concepts and Relationships
Necessary for Visual Understanding

新しいAutonomyは「人」をふくむ

- 人にアダプトする車
 - 助ける必要があるときだけ助けるアシスト技術
 - 人と協同する生産システム
-
- → 人の Functional・Behavioral モデル
 - Digital Human?

Autonomous Robots (再) Total Integrated System



Analytics and Machine Learning (“狭い”AI)

How fast and big
a program can
learn

Computing: Processing and Data Kernel

How fast, big,
and organized a
computer can
process and store

人や物理世界とインターアクションする Autonomous Systems のチャレンジ(再)

- 物理プロセス(センシング、物体の性質、動き、摩擦、、、マニピュレーションなど)のモデル化
- マップなどの状況表現法
- 雑音、矛盾するデータ・情報の扱い(データと地図)
- Quick ReactionからDeliberate Actionまでの統一的なアーキテクチャ
- システムの一部としての“ひと”のモデル化
- 一つ一つ少しずつ違う物体などの性質をどう扱うか
- Online Learning / Adaptation
- 高レベルのInstructionと低レベルの動きの対応(例:料理のレシピと実際のアクション)
- Hard Realtime性

成功するAI研究

—私の観察—

- 具体的問題、さらに誤解を恐れずに言えば
「役に立つ」問題に絞る
- End to End で研究する
- いわゆる人工知能研究者だけでなく、いろいろな領域の専門家とのInterdisciplinary Teamを作る
 - 心理、神経、社会、経済、機構、、、
- Processor込みの研究
 - 低消費電力、Intelligent Storage、精度vs.速度

チャレンジ問題を作ろう(例)

- 自動議事録 — 会議に同席して、会議の議事録（誰が何を主張し、何を決め、誰がなにをするかのアクションプラン）を自動的に作る
 - 音声認識・視覚理解・心理解析・言語解析と生成
 - (More advanced version) 会議に参加し、有用な情報・意見を述べる
- アマチュアのゴルフコーチ — アマチュアのスイングを観察し、欠点を指摘し、その人向きの上達のアドバイスを、スイングを見せながらコーチする
 - 視覚解析・動作解析・動作生成(プランと実行)
 - (More advanced version) 必要ならば多少のうそも交えた激励も言う

人工知能研究の問題点

- 技術の“Implication(意味合い)”を人間のすること(していそうなこと)になぞらえて、解釈・解説・誇張がなされ、それに対して技術的な誇張と感情的な拒絶・恐怖がおきる
- マイルストーン的成果は技術的進歩に裏付けられたものであるが、「出来た」の意味のはっきりしないものが“成果”として混ざって主張される。
 - 私の“失敗談”： Soft AI vs. Hard AI
- 計算機の“単純な”計算能力自体が人よりはるかに低い(特に、これまでは)にもかかわらず、人との安易な(背伸びした比較)が誤解と失望を招く

使って欲しくないスローガン

— 私の好みでない —

- 「X年先を見すえた基礎的」、「X年後に役に立つ」研究 $X > 5 \sim 10$
- 「欧米と違う」、「日本独自の」、「オールジャパンの」研究

人工知能は人を超えるか—Yes!

- 人に出来ることは、同じ程度かそれ以上にできるはず
 - 「知能」は計算機科学のいう「Computation」に過ぎない
 - Computationに必要な生の計算能力は着実に増加し、人のそれ(それほど増加しない)をいずれ超える
 - Computationに必要なデータは着実にデジタル化され、人の記憶しアクセスできる範囲を(多分、既に)超える
 - 計算機の使えるセンサーは、人のそれを、種類と能力において(既に、一部は)人を超える
 - 計算機は遺伝によらず、一から学習することなく、能力を極めて能率よく継承できる
- 人は出来ると思っていることのうち、そもそも「本来的に」出来ない(または、出来なさそうな)ことは、人工知能にも出来ない(だろう)。