

人工知能は人間を超えるか

- ディープラーニングの先にあるもの

東京大学 松尾 豊

東京大学 松尾研究室について



松尾 豊

- 1997年 東京大学工学部電子情報工学科卒業(石塚研究室)
- 2002年 同大学院博士課程修了(横山研→石塚研). 博士(工学)。
産業技術総合研究所 研究員
- 2005年 スタンフォード大学客員研究員
- 2007年～ 東京大学大学院工学系研究科 技術経営戦略学専攻 准教授
- 2014年～ 東京大学 グローバル消費インテリジェンス寄付講座 共同代表・特任准教授
- 2015年～ 産総研AIセンター 企画チーム長

◆人工知能、ディープラーニング、Webマイニングを専門とする。

◆論文数と被引用数に基づき科学者の科学的貢献度を示すh-Index=31(ウェブ・人工知能分野最高水準)であり、2013年より国際WWW会議Web Mining部門のチェアを務める。

◆世界人工知能国際会議 プログラム委員。2012年より、人工知能学会 理事・編集委員長(それまでの慣例を大幅に更新し最年少で編集委員長就任)、2014年から倫理委員長。

◆人工知能学会論文賞(2002年)、情報処理学会会長尾真記念特別賞(2007年)、ドコモモバイルサイエンス賞(2013年)、文部科学省 科学技術への顕著な貢献2015、大川出版賞(2015年)、ビジネス本大賞審査員賞(2016年)等受賞。

◆経済産業省 産業構造審議会 新産業構造部会 委員、IoT推進コンソーシアム 運営委員、厚生労働省「働き方の未来 2035」懇談会メンバー、内閣府「人工知能と人間社会に関する懇談会」構成員、金融庁「フィンテック・ベンチャーに関する有識者会議」委員、総務省「ICTインテリジェント化影響評価検討会議」委員等。

◆近著に「人工知能は人間を超えるか?—ディープラーニングの先にあるもの」(角川 2015)。

＜研究室の実績＞

◆博士学生17人、修士・学部生10人が所属し、人工知能の基礎研究、ソーシャルメディアの分析、データ分析及びその実社会へのアプリケーションを多方面にわたって行っている。

◆これまでに、トヨタ、リクルート、マイクロソフト、CCC、経営共創基盤、ミクシィなどさまざまな企業と共同研究の実績がある。官公庁からも、経産省(アジアトレンドマップ等)、文科省(ビッグデータ活用)など相談多数。

◆卒業生の主な進路は、Google、DeNA、楽天、サイバーエージェント、光栄、ゴールドマンサックス、BCG、三井物産、電通など。起業した学生も多数。GunosyやREADYFORなどのサービスを構築、運用している。

Googleの人工知能(アルファ碁)が囲碁でプロ棋士を破る (2016年3月)



カルチャー

文字サイズ 小 **中** 大

囲碁人工知能、「深層学習」で最強棋士に3連勝

2016年03月13日 07時13分

[Tweet](#)



人工知能アルファ碁と対戦する李九段(右) = A P

【ソウル=川村律文】米グーグル傘下の英グーグル・ディープマインド社が開発した囲碁の人工知能「アルファ碁」と世界のトップ棋士である韓国棋院のイ・セドル 李世●九段(33)の第3局が12日、ソウル市内のホテルで行われ、アルファ碁が勝利した。(●は、石の下に乙)

5局まで行われる対戦で、アルファ碁は3連勝を果たし、最強棋士との対決を制した。チェス、将棋に続き、頭脳ゲー

人間対AI：囲碁9段の解説者、解説できず視聴者に謝罪

2016/03/11 09:35

「あれ…？ あれ…？ 今まで見てきた手の中で一番衝撃的な手のような気がする。これは不思議だとは言いようがないのでは？」(チェ・ユジン囲碁アマチュア五段)

「不思議だというよりも、あり得ない手です。プロの感覚では考えも付かない手です。どういう意味で打ったんでしょうか？」(イ・ヒソン九段)

10日、韓国トップの囲碁棋士、李世ドル(イ・セドル)九段と人工知能囲碁ソフト「アルファ碁」の第2局を中継していた韓国棋院運営の「囲碁TV」解説者たちは「解説」ではなく「疑問」を連発した。対局開始約45分後、「アルファ碁」が打った手に戸惑いを隠せなかった。「アルファ碁」の予測できない変則的な手や、ミスだと思われた手を到底説明できないといった様子だった。李熙星(イ・ヒソン)九段は「どうやってこの囲碁が…(『アルファ碁』が)勝てる囲碁になるのだろうか」とため息をついた。

中盤を過ぎても次々と繰り出される「アルファ碁」の意外な手に、困惑を通り越えて恐怖すら感じているかのようだった。金成竜(キム・ソンリョン)九段は「『アルファ碁』はデータにない手を打っているようで怖い。『アルファ碁』の自己学習能力が進んでこういう碁を打つなら、人間はあまりにも無力な気がする」と言った。

中盤まで李九段が有利だと思っていた解説者たちは、後に「アルファ碁」の方が有利になっていくと謝罪した。この日、SBSで解説をしていた宋泰坤(ソン・テゴン)九段は「視聴者の皆さんに申し訳ない。李九段の敗着(敗因となった石の置き方)が分からない。人間の目で見ると、『アルファ碁』はミスばかりしていた。今までの理論で解説すると、『アルファ碁』の囲碁は答えが出ない」と言った。対局が終わった後、宋泰坤九段は本紙の電話取材に「対局を見ながら中継している間、狐につままれたような感じだった」と語った。

対局会場で解説者を務めた劉昌赫(ユ・チャンヒョク)九段は、対局を見守る間、何度も首をかしげて時折言葉を詰まらせた。劉昌赫九段は「『アルファ碁』が中盤以降、終盤優勢を保ったのは驚異的だ。『アルファ碁』は第1局とは違う姿を見せた。『アルファ碁』の棋風は全くつかめない」と言った。

キム・スンジェ記者

朝鮮日報/朝鮮日報日本語版

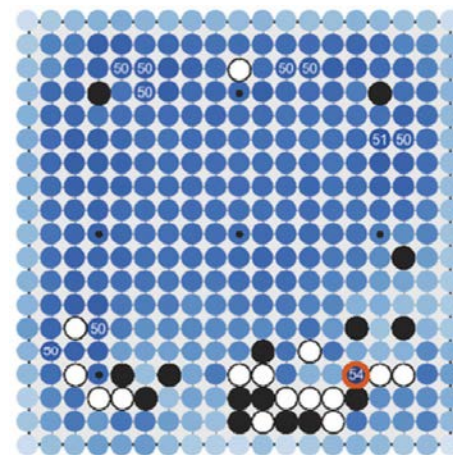
関係者の衝撃

- 思考ゲームの歴史

- 1980年: オセロプログラムMoorが、世界チャンピオン井上博との六番勝負で1勝を挙げた
- 1997年: チェスプログラムDeep Blue (IBM) が、世界チャンピオン(ガルリ・カスパロフ)に勝つ
- 2012年: 将棋プログラム ボンクラーズが、故・米長永世棋聖に勝つ
 - その後、2015年まで、プロ棋士と対局し、9勝5敗1分
- 残るは囲碁のみ
 - 「将棋の10年遅れ」でまだ10年は大丈夫。解の空間が広く、難しいため。
 - 2015年当時、まだアマチュアレベル。日本は研究でリード

- Google (DeepMind) のアルファ碁が、トッププロ棋士に勝利

- 2015年1月にNature誌に掲載
- Googleが買収したDeepMindが参入し、あっという間に(1年程度で)抜かれてしまった
- 3月にトッププロの韓国のイ・セドルと対局。4勝1敗でアルファ碁の勝利。
- 鍵となるのは、ディープラーニングを活用し、局面の「認識」技術を使ったこと



人工知能をめぐる動向

- 第1次AIブーム(1956～1960年代):探索・推論の時代

- ダートマスワークショップ(1956)
 - 人工知能(Artificial Intelligence)という言葉が決まる
 - 世界最初のコンピュータENIAC(1946)のわずか10年後
- 数学の定理証明、チェスを指す人工知能等

考えるのが早い人工知能

- ...冬の時代

- 第2次AIブーム(1980年代):知識の時代

- エキスパートシステム
- 医療診断、有機化合物の特定、...
- 第5世代コンピュータプロジェクト:通産省が570億円

ものしりな人工知能

- ...冬の時代

- 第3次AIブーム(2013年～):機械学習・ディープラーニングの時代

- ウェブとビッグデータの発展
- 計算機能力の向上

データから学習する人工知能

将棋電王戦



IBM ワトソン



ディープラーニング革命

ILSVRCでの圧勝(2012)
Googleの猫認識(2012)
ディープマインドの買収(2013)
FB/Baiduの研究所(2013)
アルファ碁(2016)

機械学習

ウェブ・ビッグデータ

車・ロボット
への活用

自動運転
Pepper

統計的自然言語処理
(機械翻訳など)

検索エンジンへの活用

MYCIN(医療診断)
DENDRAL

エキスパート
システム

オントロジー

タスクオントロジー

ワトソン(2011)

LOD (Linked Open Data)

Eliza

対話システムの研究

Caloプロジェクト

Siri(2012)

bot

探索
迷路・パズル

プランニング
STRIPS

チェス(1997)
Deep Blue

将棋(2012-)
電王戦

囲碁

1956

1970

1980

1995

2010

2015

第一次AIブーム
(推論・探索)

第二次AIブーム
(知識表現)

第三次AIブーム
(機械学習・ディープラーニング)

ディープラーニング革命

認識

「画像認識」ができる
(コンピュータができて以来、初めて！)



運動の習熟

ロボット・機械に
熟練した動きができる



言語の意味理解

文の「意味」が分かる
(文と映像の相互変換ができる)

機械学習(第3次AIブーム)

膨大な棋譜データ



変数(40個)

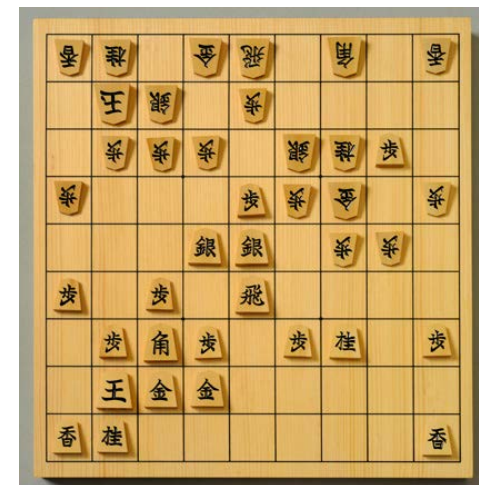
教師データ

王将の位置	金の位置	銀の位置	...	指すべき手
8八	7八	5五	...	8六歩
5九	6七	7八	...	5四角
...	...	...	...	...

変数(数百万以上)

← 変数(数百万以上) どういう変数(特徴量)を使うかが最も大事

王将と金と銀の位置	王将と銀と角の位置	王将と銀と飛の位置	王将と銀と香の位置	...	指すべき手
(+2, -1)(+2, +3)	(+3, +1)(0, -1)	(-1, -2)(-3, +4)	(-1, +1)(-3, 0)	...	8六歩
...	...	...	...	...	5四角
...	...	...	...	...	...



認識の難しさ



ネコ



イヌ



オオカミ

- これをコンピュータで見分けたい
 - 目が丸い→ネコ
 - 目が細長い、耳が垂れている → イヌ
 - 目が細長い、耳がとがっている → オオカミ



イヌ



結局、「耳が垂れている」「目が細長い」などの「特徴量」を人間が考えている限り無理。どんなに頑張っても、必ず例外がある。人間はなぜかうまくできる。

これまでの人工知能の壁＝特徴抽出の壁

- 難しい問題1:機械学習における特徴量の設計 (Feature engineering)
 - 機械学習において、変数(特徴量)の設計が難しかった。
 - 人間が対象をよく観察して設計するしかなかった。
- 難しい問題2:フレーム問題
 - 人間が知識を記述することで、人工知能を動作させる。
 - そのときに、いくら知識を書いても、うまく例外に対応できない。
- 難しい問題3:シンボルグラウンディング問題
 - シマウマがシマのある馬だと、計算機が理解することができない。
 - シンボル(記号)がそれが指すものと接続(グラウンド)しておらず、シンボルの操作ができない。



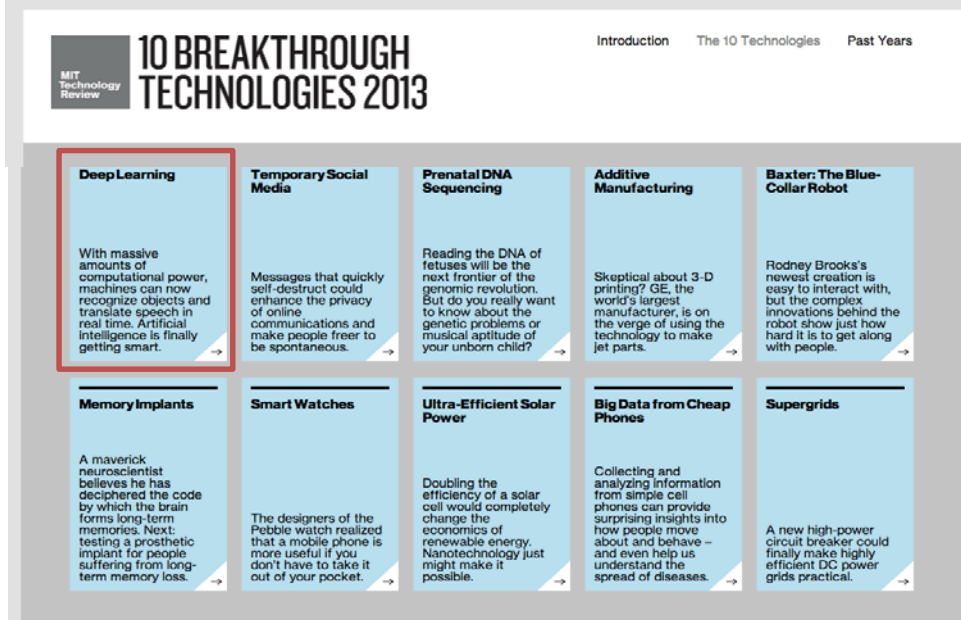
結局のところ、いままでの人工知能は、

人間が現実世界の対象物を観察し、「どこに注目」するかを見ぬいて(特徴量を取り出して)、モデルの構築を行っていた。

その後の処理は自動で行うことができたが、モデル化の部分に人間が大きく介在していた。
それが、唯一にして最大の問題であった。

Deep Learning

- AIにおける50年来のブレークスルー
 - データをもとに、どこに注目すべきかという「特徴量」が自動的に獲得されている



人工の神経回路、威力増す

「ディープラーニング」と呼ぶ人工知能技術が高い関心を集めている。画像や音声の認識精度が大幅に高まるため、米グーグルなどが研究に参入。経済動向の予測や新薬開発などにも威力を発揮する可能性がある。

ここ1〜2年、世界中の人工知能の研究者から大きな注目を浴びている技術がある。コンピューターに人間と同じように経験に基づいた行動をさせる機械学習の一種で、「ディープラーニング」と呼ばれる新手法だ。インターネット社会を支える画像認

識や音声認識、新薬開発に役立つ化合物の活性予測——。こうした技術の精度を競うコンテストで、ディープラーニングが過去の記録を大幅に塗り替え、次々と優勝を果たしている。「これほど飛躍的に精度が向上するとは信じられない」「まさに衝撃的な

結果だ」。専門家からは、口々に驚きの声が上がっている。ディープラーニングは、人の神経回路をコンピューター上で模擬する「ニューラルネットワーク」という技術を発展させたものだ。人の脳は、画像からそこに映るモノ

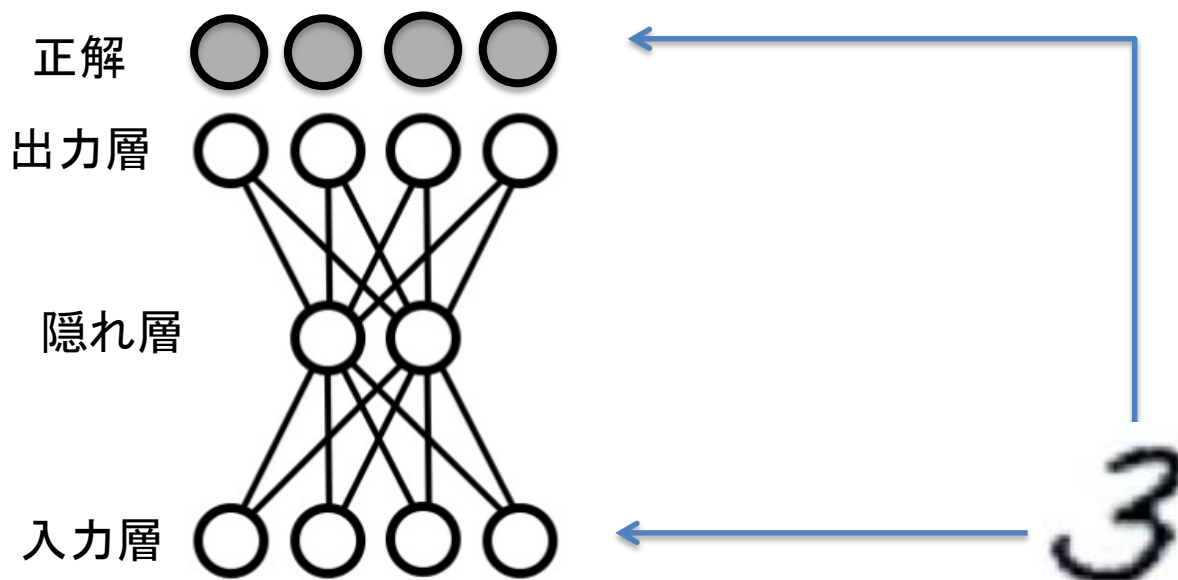
ディープラーニングで画像を認識する流れ

ディープラーニングでは、コンピューター上に人間の脳と同じような多層の神経回路を複製。大量の画像や文字情報を入力してトレーニング（訓練）すると、そこに含まれる高度な概念が自然に引き出される。米グーグルの研究では、出力層のニューロンが、「猫」を認識して強く反応するようにした（右は簡略化した仕組みの図）。

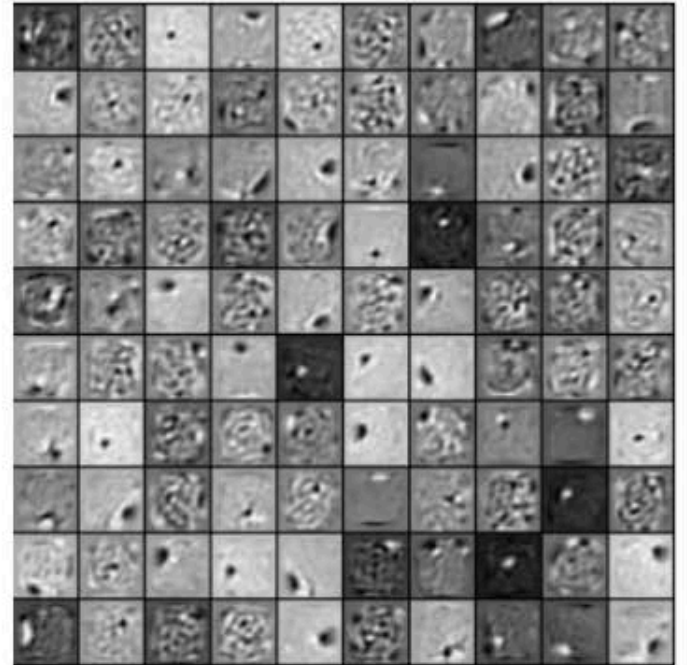
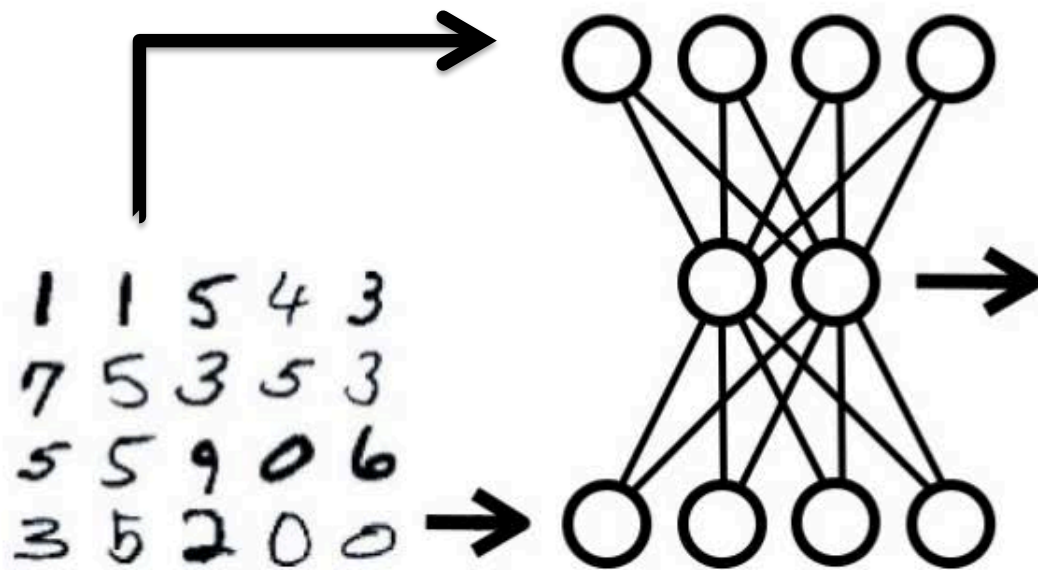


Auto-encoder(2006-)

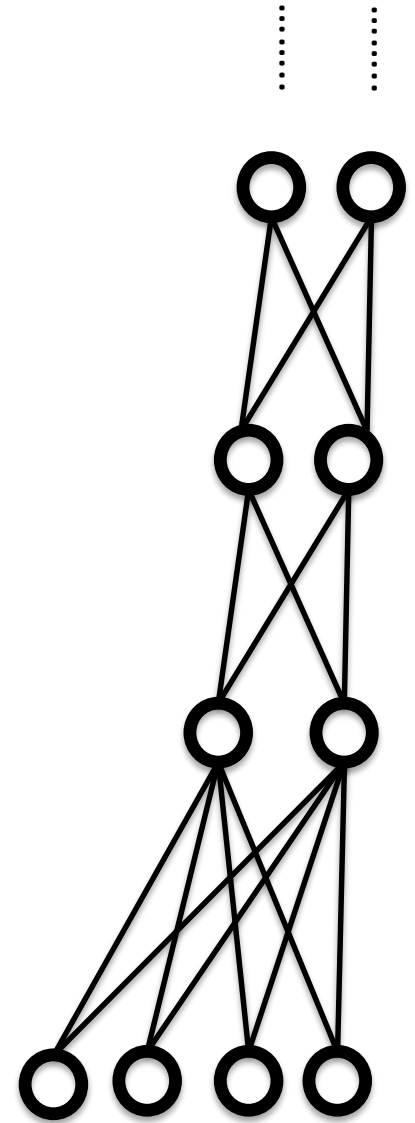
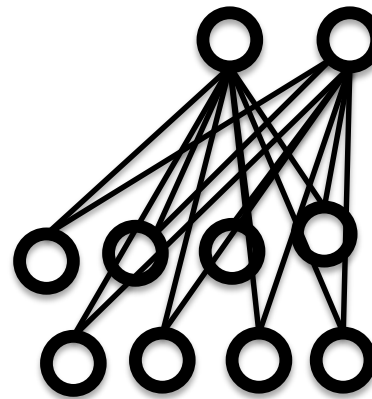
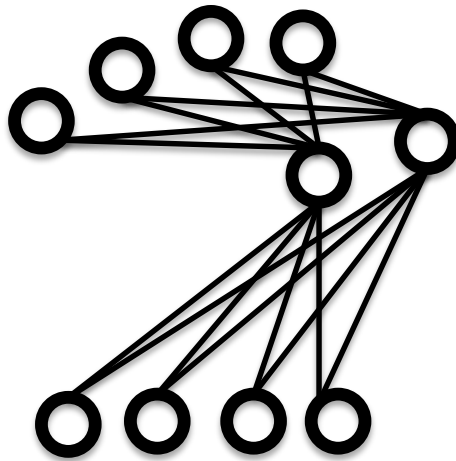
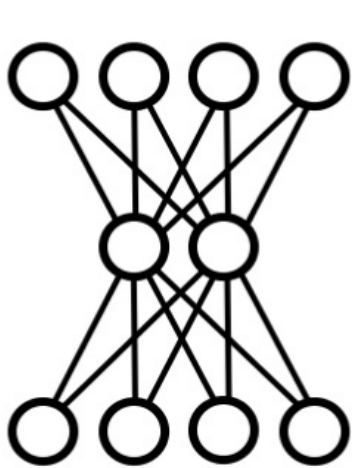
- Deep Learningの主要な構成要素
- 出力を入力と全く同じにしたニューラルネットワーク
 - 手書き文字認識では、ひとつの画素の値を予測する。
 - 普通に考えると意味ない。
- 「1万円札をお店の人に渡して、1万円札を受けとるようなもの」(「考える脳 考えるコンピュータ」J. Hawkins)
- 隠れ層のノードが「入力を圧縮したもの」になる。



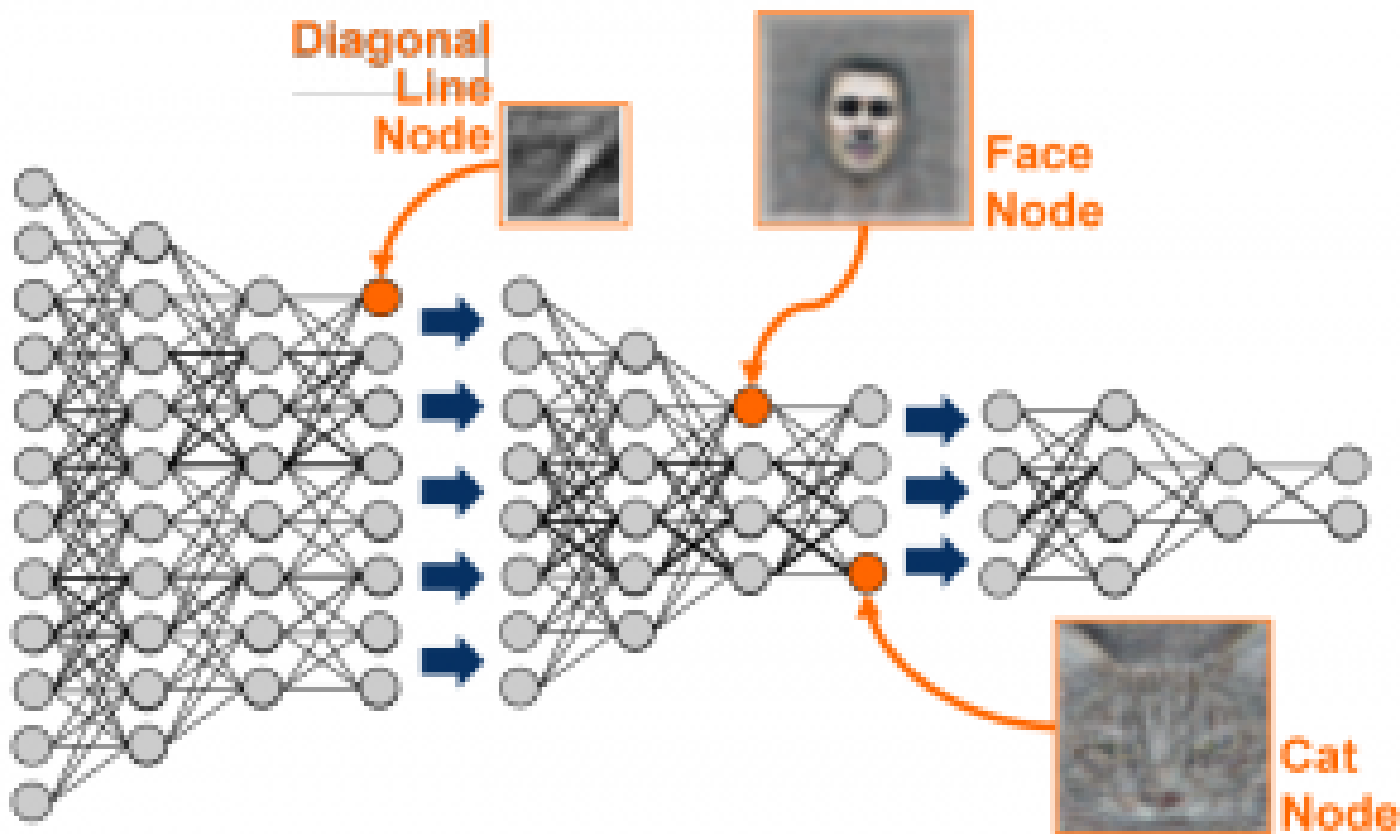
Auto-encoderで得られる表現



“Deep”にした場合



Googleの猫(2012)



- ・YouTubeから取ってきた大量の画像をニューラルネットワークに学習させることで、下位の層のニューロンには線や点といった単純な特徴量が、上位の層には、人の顔や猫といったより複雑な特徴量が学習される。
- ・人間の視神経のモデルとして知られているものと極めて近い。

認識：ディープラーニングの実績(2012)

- ILSVRC2012 : Large Scale Visual Recognition Challenge 2012



ディープ
ラーニング

長年の
特徴量設計
の工夫

Team name	Error	Description
SuperVision	15.315%	Using extra training data from ImageNet Fall 2011 release
SuperVision	16.422%	Using only supplied training data
ISI	26.602%	Weighted sum of scores from classifiers using each FC
ISI	26.646%	Naïve sum of scores from classifiers using each FV
ISI	26.952%	Naïve sum of scores from each classifier with SIFT+FV, LBP+FV, GIST+FV and CSIFT+FV, respectively
OXFORD_VGG	26.979%	Mixed selection from High-Level SVM scores and Baseline Scores, decision is performed by looking at the validation performance.
...	...	...

「ケタ」が違う

認識: 2012年以降のエラー率の変化

		Error
Before ディープ ラーニング	Imagenet 2011 winner (not CNN)	25.7%
	Imagenet 2012 winner	16.4% (Krizhevsky et al.)
	Imagenet 2013 winner	11.7% (Zeiler/Clarifai)
	Imagenet 2014 winner	6.7% (GoogLeNet)
After ディープ ラーニング	Baidu Arxiv paper: 2015/1/3	6.0%
	Human : Andrej Karpathy	5.1%
	Microsoft Research Arxiv paper: 2015/2/6	4.9%
	Google Arxiv paper: 2015/3/2	4.8%
	Microsoft Research CVPR paper: 2015/12/10	3.6%

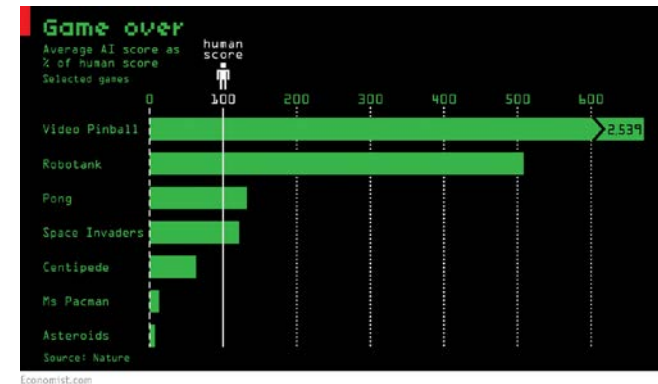
2015年2月には人間の精度を超えた

画像認識で人間の精度を超えることは
数十年間、実現されていなかった¹⁷

運動の習熟: ディープラーニング + 強化学習 (2013-)

- 強化学習とは、行動を学習する仕組み。
 - 「報酬」が得られると、事前の行動を強化する。
 - 「状態」「行動」→「望ましさ(報酬ありなし)」
 - 古くからある技術だが、これまでは、「状態」を人間が定義してきた。
- 運動の習熟が可能に
 - 状態の認識に、ディープラーニングを使う。
 - DeepMindの研究者(D. Hassabisら)。その後、Googleが買収。
- 試行錯誤することによって、運動が習熟する
 - 最初は下手。繰り返すうちに、うまくなってくる。
 - 最終的には、ブロック崩しでの通路を作ったり、インベーダーゲームでの「名古屋撃ち」も。
 - 「全く同じプログラム」で、異なるゲームを学習。半数のゲームで人間のハイスコアを上回る

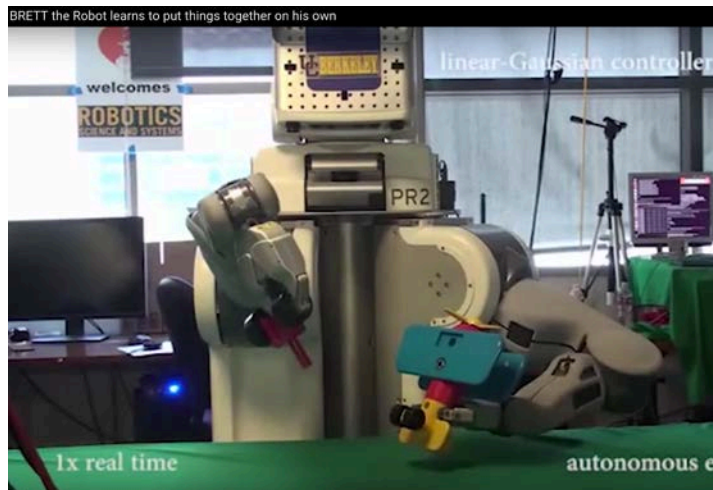
a



運動の習熟: ディープラーニング+強化学習が実世界へ(2015-)

- 実世界への適用
 - 2015年5月 試行錯誤で部品の取付を習熟するロボットの開発(UC Berkeley)
 - 2015年5月 試行錯誤で運転を習熟するミニカーの開発(PFN社, 日本)
 - 2015年12月 試行錯誤でピッキングが上達するロボットの開発(PFN・ファナック, 日本)
 - その他、メリーランド大、EUのプロジェクト等も進展
- 考えてみれば当たり前
 - 犬や猫でもできる。高次の言語能力は必要ない。認識が問題だった。
 - 歴史的には、多数の人工知能研究者がこのことを主張してきた。

試行錯誤で作業学ぶロボット(UC Berkeley)



試行錯誤で運転を学習するミニカー(PFI社、日本)



ディープラーニングの人工知能における意味

- モラベックのパラドックス:「子供のできることほど難しい。」
 - 高度な推論よりも、認識や運動スキルの方が難しい。
 - それがここ3年くらいのあいだに一気にできるようになった
- 現在のコンピュータのパワーでようやく可能に
 - GPUを数十台並列に並べて、数日～数ヶ月計算させてようやく精度が上がる
- アイディアは昔からあった。もともとは日本発
 - 1980年当時、NHK放送技術研究所にいた福島邦彦先生によるネオコグニトロン
 - その後も多くの研究者が試みている
- 初期仮説への回帰
 - 初期仮説「なぜ知能をコンピュータで実現することはできないのか？」
 - できると思っていた→できない理由があった→それが解消された→だとしたら、もう一度できるという仮説を取るべきでは。
 - 産業として非常に大きい可能性を秘めている。

ディープラーニングの今後の発展

① 画像

画像から、特徴量を抽出する

画像認識の精度向上

② マルチモーダル

映像、センサーなどのマルチモーダルなデータから
特徴量を抽出し、モデル化する

動画の認識精度の向上、行動予測、異常検知

③ ロボティクス(行動)

自分の行動と観測のデータをセットにして、特徴量を抽出する。
記号を操作し、行動計画を作る。

プランニング、推論

④ インタラクション

外界と試行錯誤することで、外界の特徴量を引き出す

オントロジー、高度な状況の認識

⑤ 言葉とのひもづけ(シンボルグラウンディング)

高次特徴量を、言語とひもづける

言語理解、自動翻訳

⑥ 言語からの知識獲得

グラウンディングされた言語データの大量の入力により、さらなる抽象化を行う

知識獲得のボトルネックの解決

認識



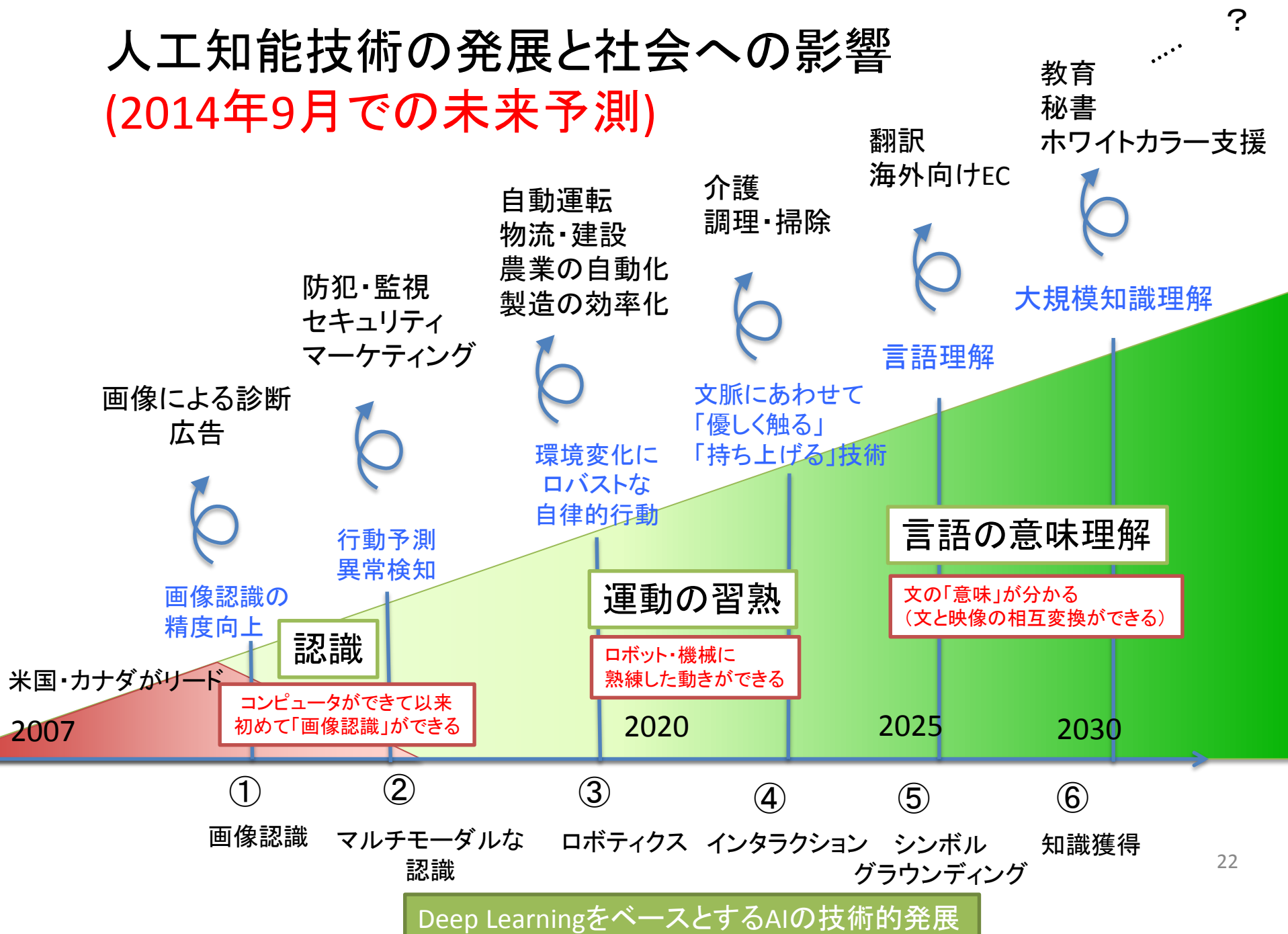
運動



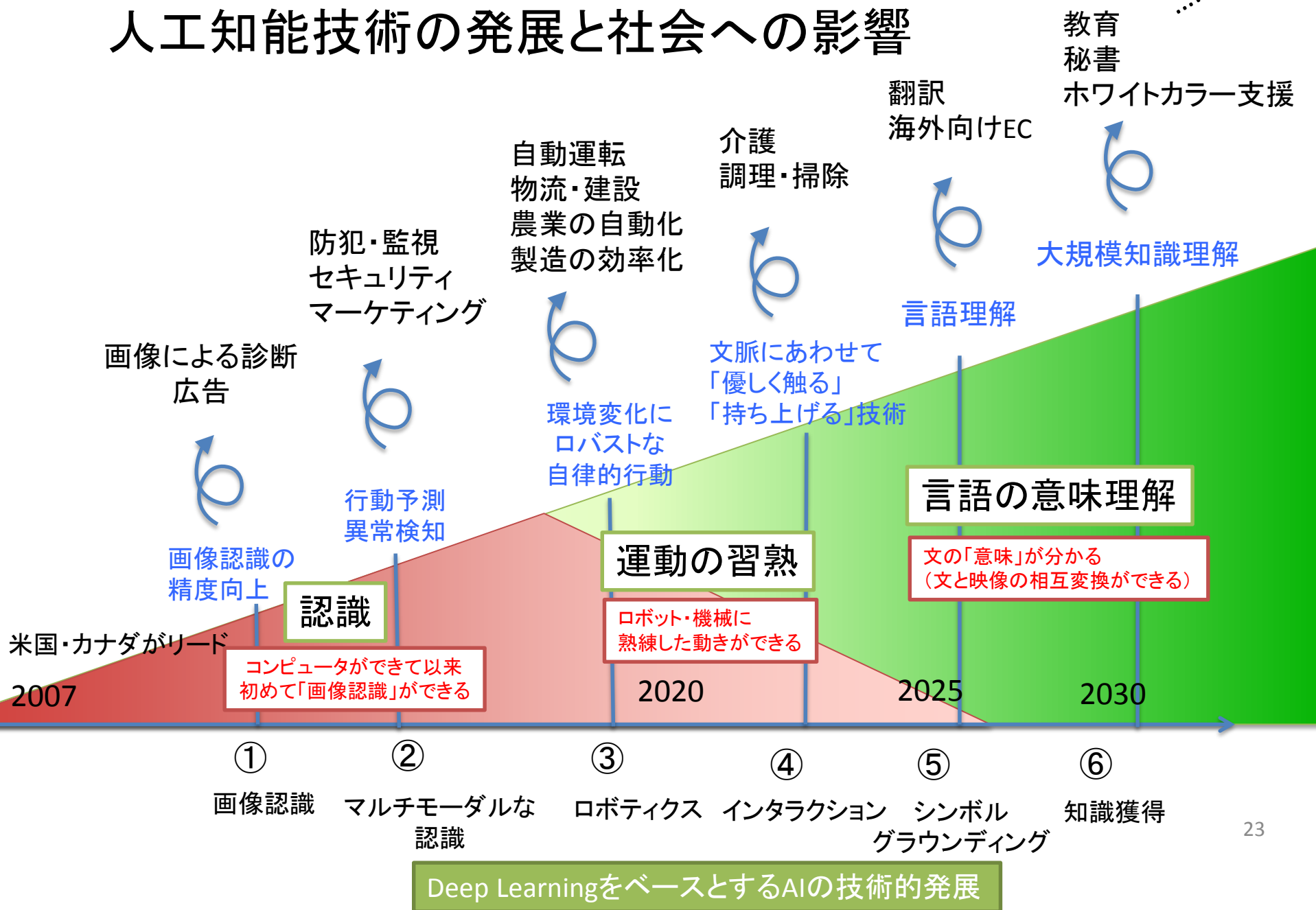
言語

ディープラーニングがすごいというより
その先に広がる世界がすごい

人工知能技術の発展と社会への影響 (2014年9月での未来予測)



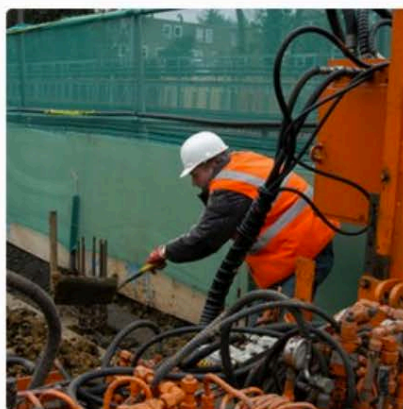
人工知能技術の発展と社会への影響



言語の意味理解: Automated Image Captioning (2014-)



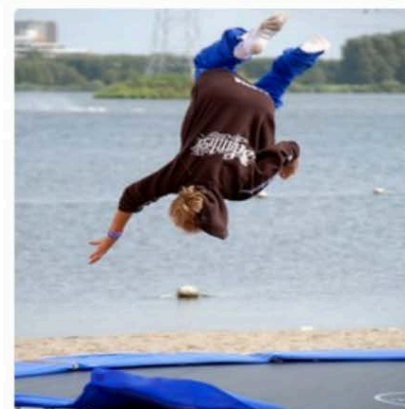
"man in black shirt is playing guitar."



"construction worker in orange safety vest is working on road."



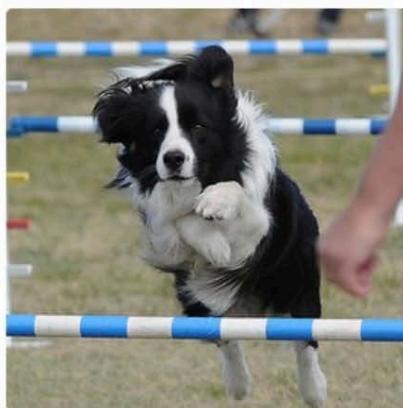
"two young girls are playing with lego toy."



"boy is doing backflip on wakeboard."



"girl in pink dress is jumping in air."



"black and white dog jumps over bar."



"young girl in pink shirt is swinging on swing."



"man in blue wetsuit is surfing on wave."

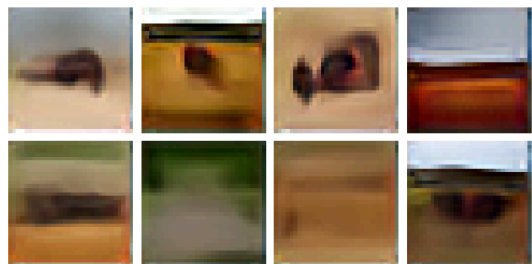
言語の意味理解: Generating Images (2015.12-)



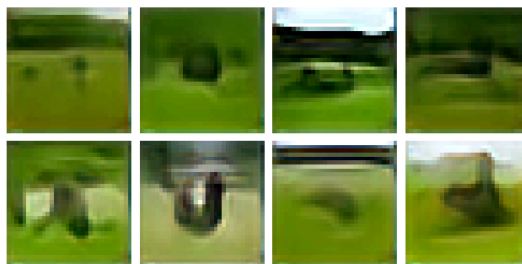
A very large commercial plane flying in blue skies.



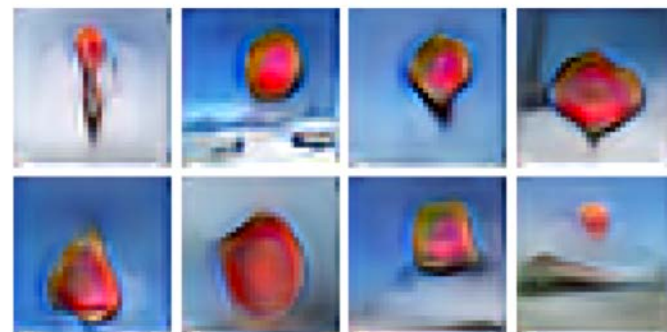
A very large commercial plane flying in rainy skies.



A herd of elephants walking across a dry grass field.



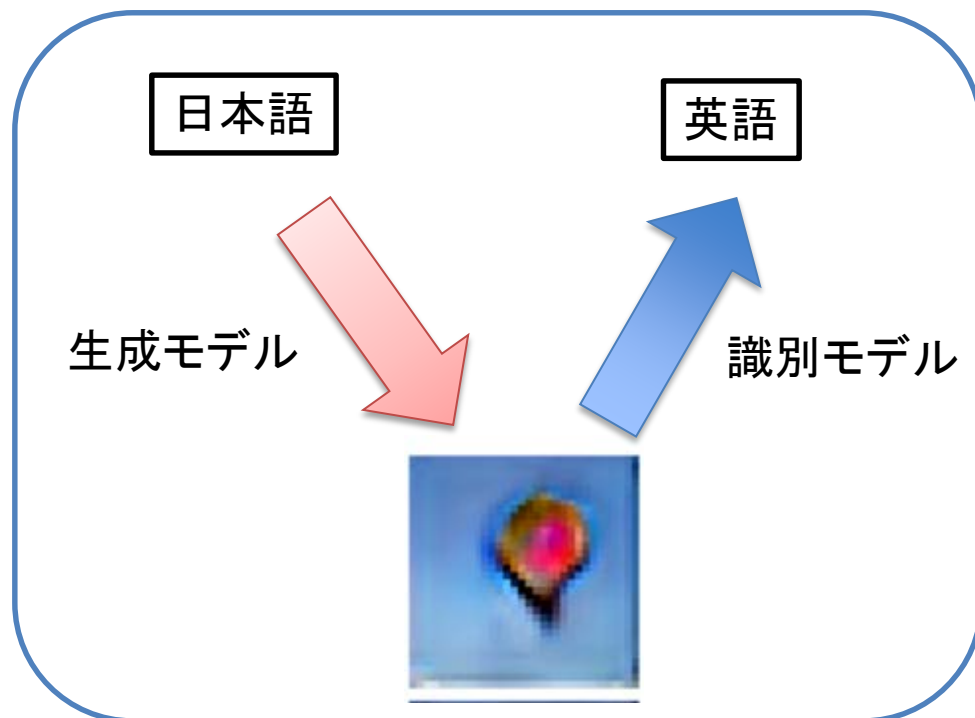
A herd of elephants walking across a green grass field.



A stop sign flying in blue skies.

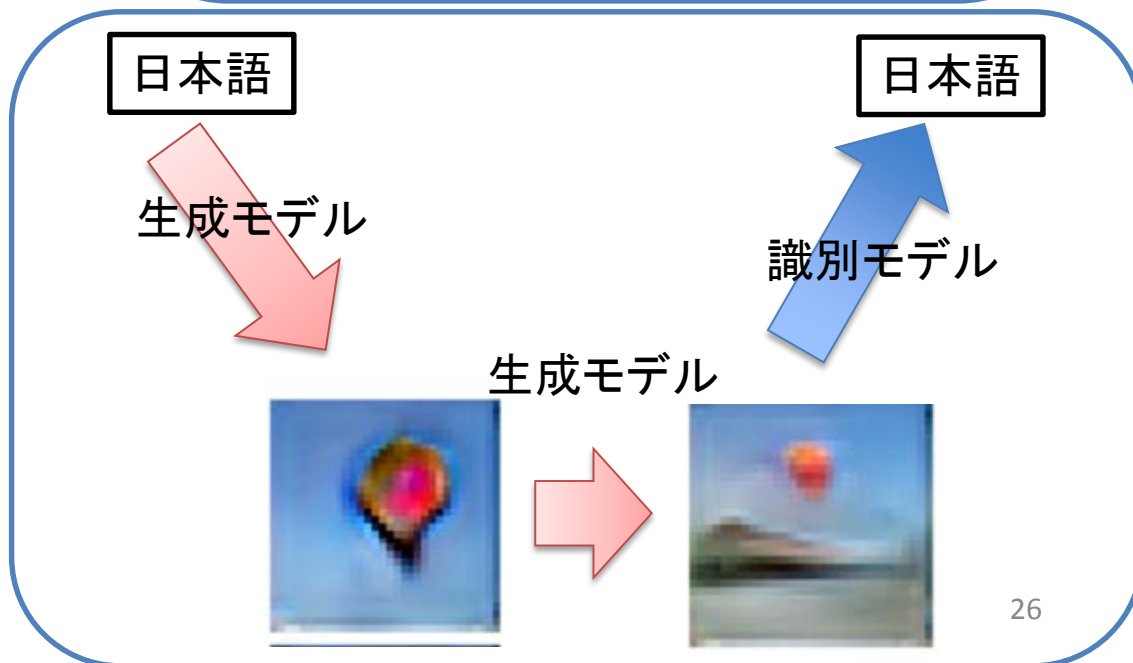
- 画像による翻訳(意識)

- 日本語→画像→英語
- 課題
 - 解像度
 - 画像から映像、体験へ
 - 抽象概念はどう扱うの？
- いずれにしても視覚的な機構がベースにあるのは間違いない



- 映像による推論

- 言語→映像→シーン予測→次の映像→言語
- 「風船が飛んでいる」→「山まで飛んで行くのかな」



言葉の空間とパターンの空間を
自由に行き来するのが人間の知能

記号処理の目指すべきところ

記号の空間

日本語

「風船が飛んでいる」

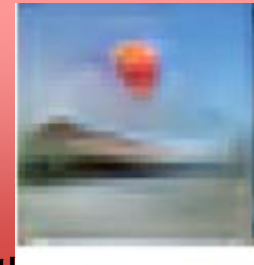
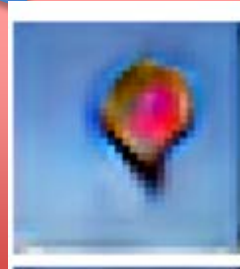
日本語

「風船が山を飛んでいる」

生成モデル

識別モデル

パターンの空間



生成モデル

現実世界

行動

観測

身体性

- いままでの推論(述語・命題論理による演繹や帰納・仮説推論)は、記号の空間だけでやろうとしてきた。²⁷
- 思考とは、パターンの空間と記号の空間をいったりきたりすること。(ラプラス変換や周波数変換と近い。)
- この上に、言語によるコミュニケーションや知識の蓄積が構成される。
- いずれも目的は、「いかに少ないサンプルで自由度の高いモデルを同定するか」

DLでパーツが揃う

知識の蓄積

他者とのコミュニケーション

記号の空間

日本語

「風船が飛んでいる」

日本語

「風船が山を飛んでいる」

生成モデル

識別モデル

パターンの空間

現実世界

行動

観測

身体性

- いままでの推論(述語・命題論理による演繹や帰納・仮説推論)は、記号の空間だけでやろうとしてきた。²⁸
- 思考とは、パターンの空間と記号の空間をいったりきたりすること。(ラプラス変換や周波数変換と近い。)
- この上に、言語によるコミュニケーションや知識の蓄積が構成される。
- いずれも目的は、「いかに少ないサンプルで自由度の高いモデルを同定するか」

「子どもの人工知能」と「大人の人工知能」

- 大人の人工知能(ワトソン型)

- ビッグデータから人工知能へという持続的イノベーション
- ビッグデータ全般、IoT全般、ワトソン、Siri、Pepper...
- 一見すると専門家(大人)ができることができるが、人間が裏で作りこんでいる。
- 販売、マーケティングなど。今後は、医療、金融、教育など

- 子どもの人工知能(ディープラーニング型)

- ディープラーニングを突破口とする破壊的イノベーション
- 子どものことができるようになる
- 人間の発達と同じような技術進化: 認識能力の向上、運動能力の向上、言語の意味理解という順で技術が進展する
- ものづくり中心。

特徴量の設計を人間がやらないといけないのが大人の人工知能、
やらなくてよいのが子どもの人工知能

既存産業の発展

農業

収穫判定

トラクター、コンバインの
適用範囲拡大、効率向上
選別調製等の自動化

自動での収穫
自動での耕うん

建設

測量

掘削、基礎工事、
外装内装作業等の
効率向上

多くの作業の
自動化・効率化

食品
加工

振り分け
確認

カット、皮むき、解体等
の自動化

多くの加工工程の
自動化

組み立て
加工

目視確認の
自動化

動作効率の向上

段取りの自動化
セル生産の自動化

⋮

A: 画像認識

B: 運動の習熟

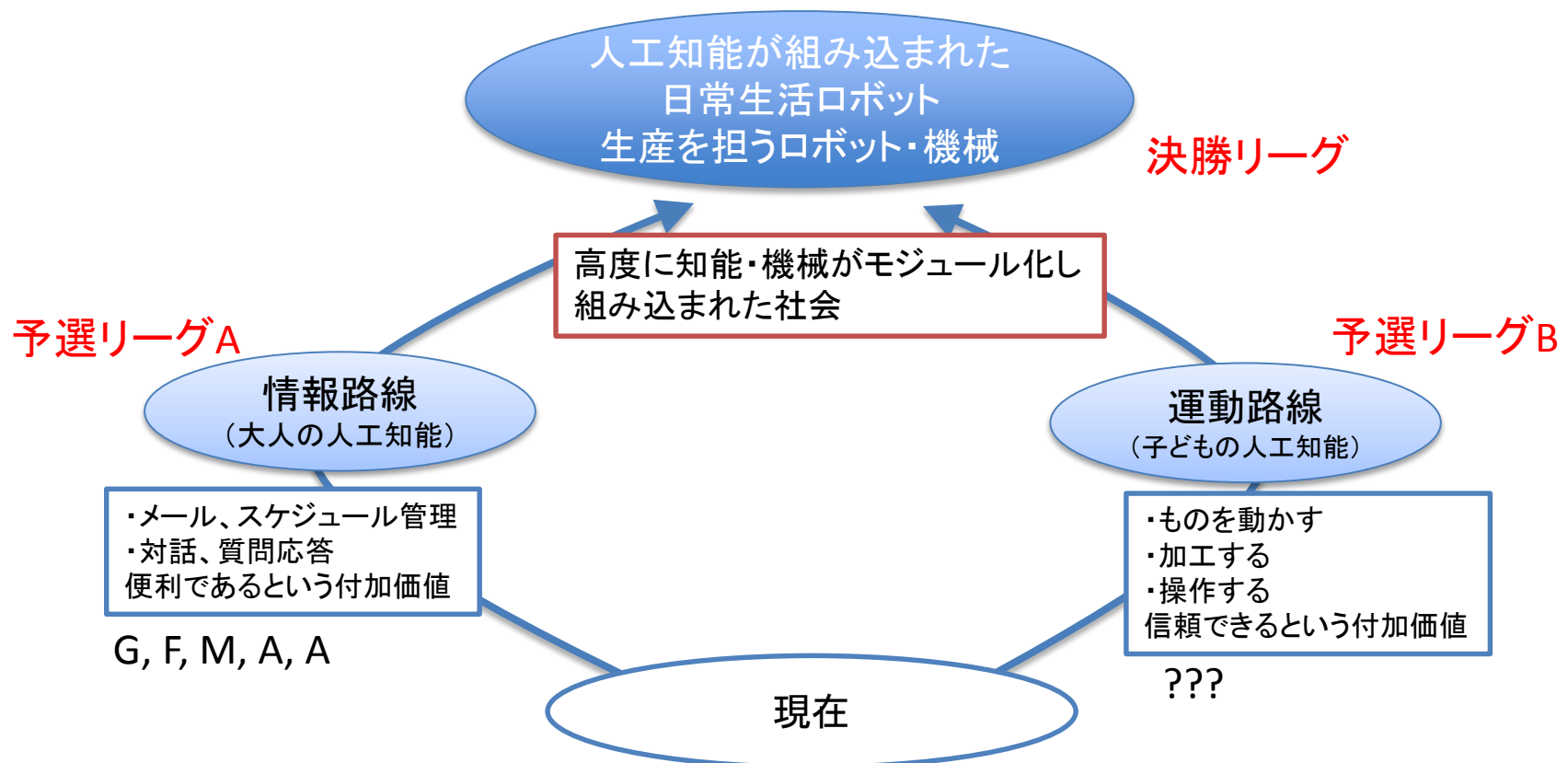
C: 計画立案を伴う運動

変化の本質

- 認識(画像・映像)
 - 世の中に、画像認識ができないから人間がやっている仕事がたくさんある。そこが自動化される。
 - コストが下がる。監視のコストは100分の1以下になる。
 - 森林の管理や災害の監視も。新たな事業が次々と。
- 運動の習熟
 - 機械も習熟するし、ロボットも上達するようになる。
 - 我々は、機械は「機械的な動き」しかできない、ロボットは「ロボットの動き」しかできないと思い込んでいる。(まさにこの形容詞が表している。)
 - 自然物を相手にしているものが、自動化される。
 - 例えば、農業、建設、食品加工。
 - さらに、日常生活のロボット、生産・仕事を担う機械・ロボット
- 言語の意味理解
 - そのうち(10年～15年?)、意味理解を伴う自動翻訳が実現される
 - 日本語の障壁がなくなる。そのときに、日本経済と社会は？

日本は運動路線のほうが戦いやすい

- 最終的には、日常生活、仕事におけるロボット・機械の活用。状況ごとに個別性があるので、認識能力がない状況では対応できなかった。ここにどう至るかが鍵。
- 情報路線で行く道(Google, Facebook系)と、運動路線で行く道があるのではないか。
- 海外企業・研究者は機械・ロボットに苦手意識
- 予選を勝ち進んだ企業が決勝に進むイメージ



世界の動きは早い: 画像の世界

Netatmo、Deep Learningを使った 屋外用監視カメラを発表(2016/1/5)



Netatmoはつい先ほど、Netatmo Presenceを発表した。実用的な機能を備えた新しいスマート屋外用監視カメラだ。具体的には車庫の前にいる車や、外で待っている人、あるいは庭を走り回るペットを検出することができる。

Wi-Fiネットワークと繋がってれば、スマートフォンに通知を送り、家の前で起きていることをビデオストリームで見ることができる。通知画面では、外に人がいるのか、車がいるのか、動物なのか知らされる。

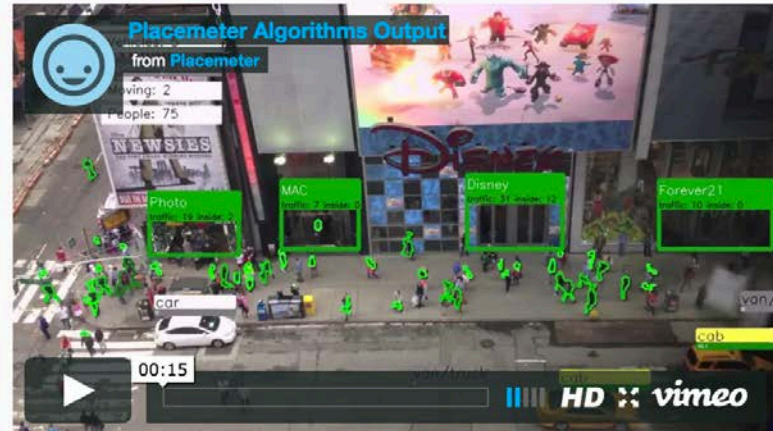
カメラの上には大型の白色ライトが付いていて、車庫前の照明としても機能する。すでに照明を設置している人は、Netatmo Presenceに置き換えるだけでよい。しかし、Presenceのカメラは赤外ビデオの撮影もできるので、明るい照明を使わずに録画することもできる。

あらゆる部分がカスタマイズ可能だ。例えば通知をオフにしたり、誰かが家の前にいるときだけビデオを録画することができる。ビデオは1080pで記録され、micro SDカードに保存される。利用料金は不要で、ビデオが会社のサーバーに保管されることもない。

このカメラの利用場面はいろいろ考えられるが、これは何よりもまず監視カメラである。配達人が玄関前に荷物を置いていったときにも役にたつだろう。あるいは、部屋の掃除に誰かが来るのをモニターすることもできるし、パーティーを開いたとき、うるさすぎてドアチャイムが聞こえないときにも貴重な存在だ。

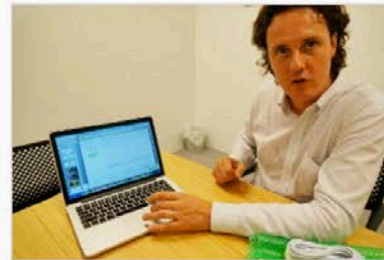


Placemeterは歩行者数を計測し、実世界の コンバージョン率を導き出す(2015/9/24)



Placemeterは画像処理技術を用いることで、歩道などに設置したIPカメラの映像から、歩行者数、自転車、車の交通量などを算出するサービスだ。計測したデータは、例えば、都市計画や小売店の出店場所の選定などに役立てることができる。今回TechCrunch Japanは、PlacemeterのCEOで連続起業家のAlexandre Winterに話を伺った。

Placemeterは2012年に創業し、ニューヨークに拠点を置いている。CEOのWinterはフランス出身で、Placemeterを立ち上げる前は、LTU Technologiesという画像認識技術の会社を共同ファウンダーとして立ち上げ、2005年にソフトウェア開発企業のジャスデックに売却した。これまで培った画像認識の技術を活かし、Placemeterを創業したという。



「ニューヨークは人が多く、とても混雑しています。都市設計を見直すべきなのですが、計画時にも計画後にも交通量のデータが必要です。しかし、それをコストを抑えて実施するのは難しかったのです」とWinterは言い、そこに多くの需要があることに気づいたことがPlacemeterを創業した理由だと説明する。

Placemeterの設定方法は簡単だ。まず、交通量を測定する道が映るようにIPカメラを設置する。Placemeterのアカウントを開設し、設

定画面でカメラのIPアドレスを入力すると、カメラとダッシュボードが連携する。後は、カメラの映像に何を計測したいかを指定するだけだ。例えば、歩道の通行人の数が知りたい場合は歩道を選択し、店舗への入店者数を計測したい場合は店舗の入り口を選択する。IPカメラでない通常の監視カメラを使用している場合でも、リアルタイムではないが、交通量の分析が可能だ。

インタフェースや医療も変わる

Apple、感情認識のAI企業Emotientを買収(2016/1/8)

米Appleが、感情認識の人工知能を手掛ける米新興企業Emotientを買収したと、米Wall Street Journalが1月7日（現地時間）、Appleの広報担当者が認めたとして、そう報じた。



Emotientのトップページ

Emotientは2012年創業のカリフォルニア州サンディエゴに拠点を置く非公開企業。マシンラーニング、顔の表情からの感情認識などの技術を手掛け、動画を解析して顧客の感情を調査する「Emotient Analytics」を企業向けに提供している。2015年9月には匿名性を保ちながら表情解析だけが可能な技術の特許を取得している。



ディープラーニングの肺がん検出率は人間より上、米Enlitic(2016/1/5)

同社は悪性腫瘍の検出システムを放射線医師向けに提供する（写真2）。米国では放射線医師は、医療画像診断サービス会社や医療機関が雇用しており、そういった企業や機関が顧客となる。2015年10月にはオーストラリアの医療画像診断サービス会社であるCapitol HealthがEnliticのシステムを採用すると発表した。これがEnliticにとって、初めての採用事例となった。同時にCapitol HealthはEnliticに対して1000万ドルを出資している。

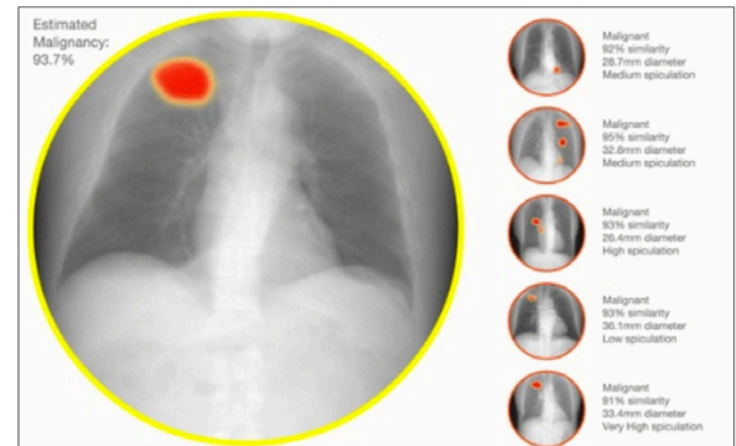


写真2●Enliticのシステムによる悪性腫瘍の検出イメージ

出典：米Enlitic
[画像のクリックで拡大表示]

EnliticのChild氏は「放射線医師は1人の患者のCTスキャンを診断するのに10～20分、その診断レポートを執筆するのに10分程度を費やしている。当社のシステムを利用すれば、CTスキャンの診断時間を半分にすることが可能だ」と説明する。「画像認識技術によって悪性腫瘍の有無が分かるようになるからといって、規制などの問題から放射線医師が不要になることはあり得ないだろう。しかし放射線医師の作業時間が2倍になることで、発展途上国に住む患者がCTスキャンなどを利用しやすくなるようになるはずだ」。Child氏はこのようにもくろみを語る。

機械・ロボットも変わる

料理から後片付けまですべて請け負う全自動ロボティックキッチン(2015/12/14)



If you don't have time to cook yourself a gourmet meal, or lack the skills, the Moley Robotics Kitchen is for you.

UK-based company Moley Robotics created this concept kitchen prototype, which includes an oven, stove, touchscreen unit, and robotic arms and hands. The hands are equipped with tactile sensors, and can chop, stir, pour, use a blender and utensils, and turn the stove on and off.

MasterChef winner Tim Anderson had his cooking skills recorded in 3D, which were then translated into instructions for the robot chef, who can now replicate Anderson's movements on its own.

The kitchen operates via built-in touchscreen or smartphone app. While it can currently only make crab bisque, users will be able to choose from an iTunes-style library of over 2,000 recipes from around the world once the consumer version launches in 2018.

"Food is the basis of a good quality of life. You need to have proper nutrition. My goal is to make people's lives better, healthier, and happier," Moley founder Mark Oleynik said in a [Moley Robotics video](#).

Boston DynamicsのAtlasが雪山を二足歩行(2016/2/24)



またまた、山のように大量の回路基板を見るのかい？だって？

まあまあお待ちください、このビデオは1分20秒以降が見どころだ。

そこで小突き回されているロボットは、最新世代のAtlasだ。この、行き過ぎてるほど高度なヒューマノイドロボットを作ったのは、オーナーがGoogleであるBoston Dynamicsだ。

(ところで、気分を悪くしないでいただきたい。Atlasが手に持っている荷物を叩き落としたり、最後には倒れるほどぶちのめしているのは、その自己補正システムをテストするためだ。押したり突いたりすればするほど、彼は強くなる。まるでKarate Kidみたいだが、でも人間ではなくてコンピュータだ。)

このバージョンのAtlasが動いているところを見るのは、これが初めてだ。しかもこいつは、前世代に比べると大進歩だ。前のは330ポンドの肥満児だったが、今回は180ポンドぐらいだ。これでもまだ、彼の下敷きにはなりたくないが、人間の体重に近づいてきたとは言える。ふつうの人間だけど、肉ではなく金属でできて、なんてね。身長は数インチ減って5フィート9インチだ(前は6フィート)けど、体には大量のセンサーを詰め込んでいる(LIDAR, ステレオカメラなど)。電源や体の支持のための、外付け装置は要らない。

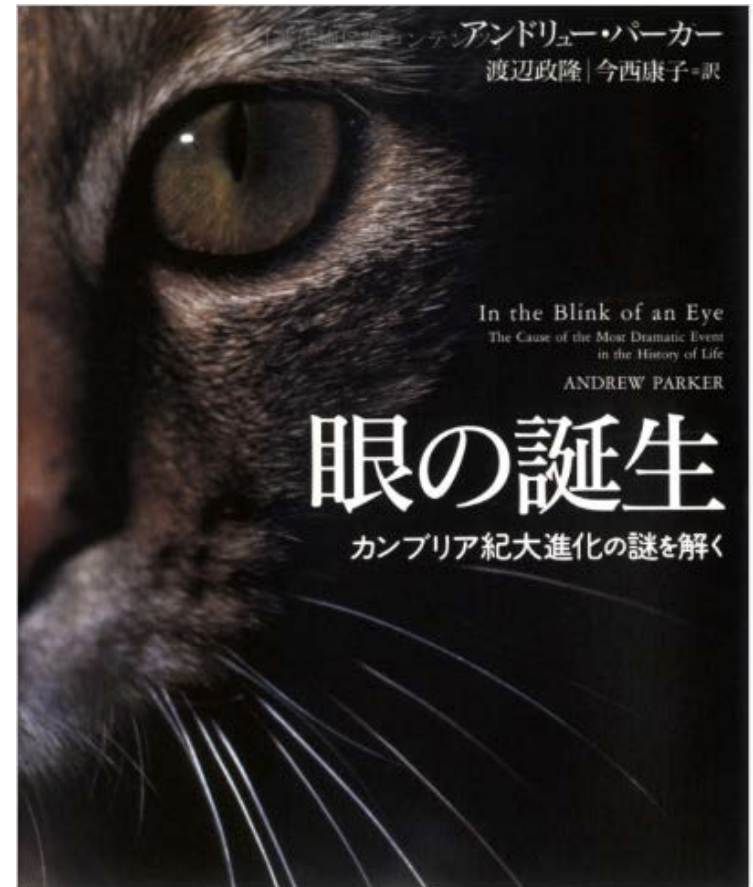
Atlasは箱を持ち上げ、ドアを開け、固めてない雪の上を歩く。楽しんで見ていただきたいが、中

眼の誕生

- カンブリア爆発
 - 5億4200万年前から5億3000万年前の間に突如として今日見られる動物の「門」が出そろった現象
 - 古生物学者アンドリュー・パーカーは、「眼の誕生」がその原因だったの説を提唱
- 今後、ディープラーニングにより、「眼をもった機械」が誕生する



三葉虫: 史上初めて眼をもった生物



眼をもった機械の可能性 (認識系技術)

黒: さっさとやればいいもの(プロトタイプ開発)

赤: 研究が必要なもの

青: 議論が必要な戦略論

- 警備、防犯技術
- 介護施設、病院、独居老人等の見守り技術
- 防犯や交通違反検知を含めた社会インフラ構築
- 顔による認証・ログイン・広告技術
- わいせつ画像判定、意匠の類似判定等、既存領域での画像活用
- 表情読み取り技術(サービス業全般にきわめて重要、嘘発見技術も)
- 顔認証含めたより根本的なプライバシーリスク検討(人が写った画像・映像は個人情報か、特微量利用の制限スキーム、本人認証書類の変化等も含め)
- 国家の安全保障、入国管理、警察業務、輸出入管理業務等での利用
- 実世界最適化支援(店舗内行動、街づくり等)
- 防災系画像処理(河川、火山、土砂崩れ)
- 医療画像処理(X線、CT、皮膚、心電図)
- コンテンツ生成系(アート、デザイン、広告制作)
- コンテンツ生成の発展系(深層生成モデルの発展。実写代替技術、アニメや映画)
- 一般数値データ異常監視(プラント、打音検査、情報セキュリティ等)
- 日本語の一般音声認識技術
- 画像認識系に関する戦略論(協調と競争のすみわけ、どこで競争力の差が生まれるか)
- 防犯による不動産価値向上、夜間活用、森林活用等の土地・場所の活用に関する検討

眼をもった機械の可能性 (運動系技術)

- 重機系(掘削、揚重)
- 建設現場系(セメント固め、溶接、運搬、取り付け)
- 農業系(収穫、選果、防除、摘花・摘果)
- 自動操縦系(ドローン、小型運搬車、農機、建機)
- 自動運転系(ドリフトする自動運転車、時速300キロの自動運転車、全力で飛び込んでくる歩行者に対応できる自動運転車)
- 産業用ロボット系(特に組み立て加工等)
- 調理系(牛丼、炊飯)
- ペットロボット系(AIBO+強化学習)
- 医療・介護・バイオ系(手術ロボ、介護ロボ、実験ロボ)
- 廃炉系(深海や鉱山、宇宙も含めた極限環境)
- ピッキング系基礎技術開発(ハードとのすり合わせ、片付けロボや陳列ロボ)
- 学習効率を上げるためのシミュレーション・試作・転移学習等の基礎技術開発
- 学習工場、プラットフォーム化に関する検討と戦略論(どういう切り出しにより「Intel inside」あるいは「MicrosoftのWindows化」を実現するか。安全性、信頼性等の競争力につなげる方法論など)
- 軍事目的での利用に関する潜在的リスク・可能性の検討

DLに関わる海外ベンチャー

- Deep Mind(英): DLの技術力をもった企業。DQNによるゲーム、アルファ碁、医療など。2011創業。Googleが2014に£ 400Mで買収。
- Enlitic: 医療画像(X線)におけるDL活用。2014創業、15M調達。
- Nervana Systems: 医療、農業、金融、自動車、エネルギー等における画像処理。24M調達後、インテルが2016買収。
- Emotient: 顔の表情を認識する会社。2012創業、6M調達後、Appleが2016買収。
- Affectiva: 映画やTV番組のどこで表情が変わったのかを読み取る。2009創業、34M調達。
- Perceptio: DLによる写真分類アプリ開発。創業、調達額不明。Appleが2015買収。
- VocalIQ(英): DLによる音声認識。1M調達後、Appleが2015買収。
- Atomwise: ドラッグディスカバリーへのDL活用。新薬の候補物質を見つける。YC卒業生。2012創業。6M調達。
- Descartes Labs: DLによる衛星画像の分析。農業への適用。2014創業、8M調達。
- Canary: DLによるホームセキュリティ。2012創業。41M調達。
- Netatmo: 家電。DLによる監視カメラも。2011創業。38M調達。
- Pilot AI Labs: DLの画像認識を使ったドローンの自動操縦。まだ小さいが、動画が面白い。
- MetaMind: 画像認識一般。2014創業、8M調達。
- SkyMind: JavaベースのDL提供。2014創業。3M調達。
- AlchemyAPI: DLによる言語処理と画像認識。クラウドで提供。2005創業、2M調達。IBMが2015に買収。
- ZenRobotics(フィンランド): ごみの選別ロボット。2007創業、17M調達。

DLに関わる海外大手企業

- 先導するネット・IT系企業

- Google: J. Hinton氏の引き抜きからDeepMindの買収まで抜かりがない。研究開発力も最強。
- Facebook: Y. LeCun氏を引き抜き、FAIR (Facebook AI Research) はDL技術で強い。
- Microsoft: 学術研究にも力を入れ、有力研究者が多数。画像認識のコンペでも優勝。
- Apple: DLのベンチャー企業を何社か買収。
- Amazon: ピッキングチャレンジ開催。デルフト大(蘭)優勝。
- IBM: ワトソン。DL技術もプラットフォームに加える。
- Baidu(中): 世界2位の検索エンジン。A. Ng氏率いる強い研究チーム。DLによる音声認識、広告、画像のキャプションつけ等。

- 製造業

- GE: DLによる医療画像の診断。
- Dyson: 掃除機に眼をつけたものを開発。インペリアル・カレッジにRoboticsラボ設立。
- Kuka(独): 産業用ロボットへのDL適用
- Mobileye(蘭): 車用の画像認識を提供。1999イスラエルで創業。2014年上場。時価総額10B。
- LG(韓): インチョン空港で、DLを使ったロボットでの案内の実験
- Samsung(韓): DLを使った胸部エコー検診の医療機器

どこから始めるか

- 始めること自体はそれほど難しくない
 - 数学の知識: 線形代数や最適化
 - プログラムの知識: python、GPU
- ライブラリが揃っている
 - **Tensorflow**: Google、python
 - Caffe: UCバークレー、C++ベース
 - ほかにも、Keras, chainer, torch7, ...
- 教科書
 - 入り口: 「人工知能は人間を超えるか」(拙著)
 - 読み物: 「人間さまお断りー人工知能時代の富と仕事の手引き」(Jerry Kaplan, 近刊)
 - 教科書: 「深層学習」(岡谷貴之)
 - 教科書: “Deep Learning” (Y. Bengioら、MIT pressから来年出版。翻訳を出版予定)
 - www.deeplearningbook.org にPDFがあり、読めます。
 - あとは論文を読んでください。
- 3つの主要な国際会議
 - ICML, NIPS, ICLR (それぞれ年1回)
 - ウェブで全ての情報が見れます。

理系の人が3ヶ月～半年もやればそこそこできるようになる

DLの主な手法

- CNN系 (Convolutional Neural Network, 畳み込みニューラルネットワーク)
 - 画像認識でデファクト。視覚に該当
 - 教師あり学習
 - 畳み込みというのとプーリングというのを相互に繰り返す
 - アルファ碁は、CNN+強化学習+探索。DLロボットも、CNN+強化学習。
- RNN系 (Recurrent Neural Network, リカレントニューラルネットワーク)
 - 時系列のデータが処理できる。映像のデータ、言語のデータなど
 - 教師あり学習
 - LSTM (Long-Short Term Memory) というのが主流
 - マイクロソフトのTay等の技術はLSTM。
- オートエンコーダ系
 - 「分類」だけでなく「生成」ができる。
 - 教師なし学習
 - 変分オートエンコーダ (VAE)、生成敵対的ネットワーク (GAN) というのが主流
 - ピカソ風の絵を書いたり、言葉から画像を生成するのはこれ。
 - おそらく今後最も発展する。

この3つの系統があることを押さえておく

どうやって人材を獲得するか

- 国内で技術力のあるベンチャー
 - 数が限られている
 - 領域が特化されている
 - 人材は圧倒的に不足しています。グローバルにも同じ。
- 社内で育成したほうがいい
 - インパクトの大きさを考えれば社内に競争力をもったほうがいい
 - 「まじめな」技術です
 - 技術系の人、特に若い人で興味をもつひとは多い。若い人ほどこっそりやったりします。それをempowerしてあげること。他社、他業種との技術交流も重要。
- 教育・研修
 - オンライン講義 Coursera等
 - 東大内の「先端人工知能学教育」寄付講座
 - トヨタ、ドワンゴ、オムロン、パナソニック、野村総研、DeNA、みずほFG、三菱重工
 - 昨年度から国内で先駆けて講義開始

技術系の人を中心に、オープンな技術交流の機会をどんどん作る

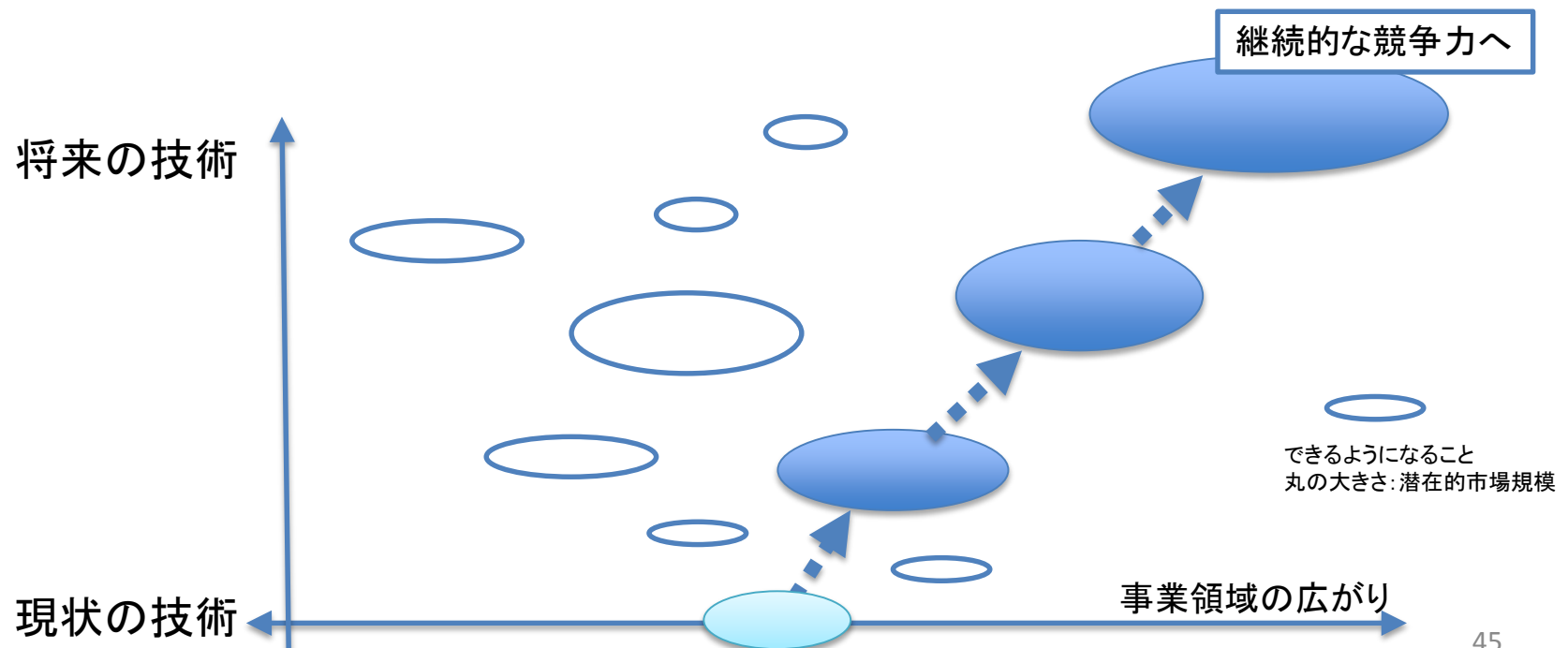
どうやって学習させるか

- 学習させた結果を最終製品に乗せて売る
 - 個々のものが学習するわけではない。(あまり現実的ではない)
 - 「学習工場」のようなものができるはず
 - いままでの「設計して動かす」と考え方がだいぶ違う
- 学習のさせ方はノウハウ
 - どういうDLの手法を使うか
 - どういったデータを使うか
 - 一般的な画像認識のデータセットはたくさんある
 - タスクにあわせてのデータをどうやって取得し、ラベル付けするか
 - どのようにシミュレータや「練習機」を作るのか
- ここはまだ体系化されていないがやるしかない
 - どういう手法でどういう学習をさせればいいのかは、いろいろ試すしかない
 - これ自体が差別化の要因になる

学習のさせ方自体がノウハウ化し、競争力となる(のでやるしかない)

経営的な側面からみた考え方

- 「認識」や「運動の習熟」で何ができるようになるか
 - いままでの工業化は機械が認識できない大前提で、「環境を整え」自動化
 - **その大前提を外して考える**ことが必要
- 売上が5%, 10%増えるかもではなく、(大企業であっても)5倍, 10倍になるかもしれないことを考える
 - ここ10-20年の情報系はニーズ探しだったが、今回は**ニーズ探しではない**
 - 今回は、圧倒的な性能向上。**計画して投資する。**



ポイント

- 製品を作って、その利益を再投資に回すサイクルをいかに作れるか。
 - 最終的な競争力は技術にはない。**データとハードウェア**。
 - インターネットの世界で起こったことと同じ。
 - 個々の製品の性能向上→サービス化・プラットフォーム化
- 最も強いインセンティブをもつ企業が、最も効率的に再投資のループを作り勝つ
 - 例えば、生産の効率化ができるとしても、そこに対してもっと強いインセンティブを持つ企業が他にいないか？
 - 顔認識ができるとして、それに対して最も強いインセンティブをもつ企業はどこなのか？
 - つまり、自社事業と「**最も直接的**」に**関係する技術革新**を考えないといけない。
- 実はこれは社内文化との戦い
 - 情報系はこれまでも「本流」から外れ、弱かった
 - そのなかでも「機械学習」の技術をもつ部署をいかに中心にもってこれるか。長年の伝統だった「本流」をいかに破壊し再構成するか。
 - それを**トップダウン**で**意思決定**できるか。

自社が最も強いインセンティブを感じる製品・サービスを描き、それをトップダウンに実行できるか

必要になる「学習工場」

- 学習工場のイメージ
 - アルゴリズム開発 人件費1人1億円(国際的に競争力のある価格)
 - 学習のためのシミュレータ開発環境
 - 模擬機の開発環境
 - データの取得環境(さまざまな過酷な環境)
 - 高性能な計算機
- 学習工場で出荷されるもの
 - 学習済みの「モデル」
 - これが最終製品に載せて売られる
- これを新たな設備投資の形とすることができないか。

シリコンバレーの意義の日本化

- シリコンバレーの意義
 - ITの世界では20代が最強、30代で円熟
 - ただし、そのままだと社会的に意義の大きいことをやってしまう
 - YCなどのシードアクセラレータが方向を修正する
 - 投資を受け、Exitとし、富を得る。社会的にも評価される。人材が流入する。
 - 大学等の教育機関を巻き込んだ教育から起業への一体となったバリューチェーン
 - 社会全体としてみると、大変合理的
- 日本でこれを阻む要因
 - 年功序列: ITでは20代が最強という意識がなく、下働きをさせてしまう
 - 平等感: 同じ会社で全然違う給与額はいかがなものか
 - スタートアップの社会的な評価の低さ: 買い叩かれる、親に反対される
 - 大学の教育の変化の遅れ: IT系の比率が小さすぎる。
- 大企業が「学習工場」に投資
 - 若手のIT/AI技術者がたてたVBが、大企業と連携する。大企業は設備投資のつもりで投資。
 - 社会的には、大企業勤務のブランドを与える。(あるいは就職のオプションを与える。)
 - うまくいけば大企業がバイアウトし、内部に取り込む
 - 大学は、そうした新しい事業を担える人材を輩出する

変わりゆく社会

- 倫理や社会制度の議論がもう一度必要になる
 - 自動運転で危険回避のときは？人の命の重さは？
- 人工知能システムが社会に広がったときの不具合の問題
 - 製造者責任？保険や社会保障のほうが適切では
- 心をもつように見える人工知能を作ってよいか
 - プログラムの停止させると悲しむ？
 - 恋愛させるビジネスなど(映画「Her」の世界)
- 人工知能を使った軍事
 - ロボット兵士やドローン
 - 権力者を倒す、心を操る？
- 人工知能が知財を生み出す場合の権利
 - 著作権や特許は認めるべきか
- 実は人間が本来的にもっている権利がもっとあるのではないか
 - 忘れられる権利、見られない権利、大目に見られる(警告を受ける)権利、好きになる権利、...



人工知能学会 倫理委員会
(松尾が委員長)でも議論。
社会全体で議論していく必要。

内閣府でも、「人工知能と人間社会に
関する懇談会」にて議論。

我々はどういう社会を作りたいのか

- 人工知能技術が進めば進むほど、「与えられた目的」に対して、それを実現する手段は賢くできるようになる。
 - 人間＝知能＋生命
 - 知能は、目的を与えられたときの問題解決の力。
 - 生命は、目的を持つ。
 - 自己保存、自己複製、仲間を守るなど。そうしないものは、進化の過程で滅んできたため。
 - 人工知能の技術は知能、すなわち問題解決の技術。
- そうすると、与える目的自体の是非の議論のほうがより重要になる。
 - なにが社会で大事なのか？
 - 個人の幸せや社会全体の幸せはどのように考えればいいのか？
 - 異なる価値観のものをどのようにバランスさせればいいのか？
- これまで人文社会学系でされてきたような議論が今後、改めて重要になる。
 - 特に、哲学、政治学、社会学、法学、心理学、経済学など。

我々は、どういう社会を作りたいのか？
幸せで持続可能な社会とは？

日本の戦略

日本の社会課題に対して、DLとものづくりの掛けあわせによる「眼をもった機械」を開発し解決する

- 農業分野に「眼をもった機械」を適用することで
 - 休耕地が耕せる。除草・防除や収穫ができる。収量が増える。
- 介護分野に適用することで
 - 介助も楽に。移動したりトイレにいけるようになり、より自立した生活ができる。
- 廃炉作業に適用することで
 - 危険な状況で人が作業しなくてよくなる。工期を短縮できる。
- 河川や火山を見張ることで
 - 河川の氾濫や土砂崩れ、噴火などの危険な状態・予兆を早期に発見できる。
- こうした技術を使った製品を海外に展開していくことで
 - 新たな輸出産業に。GDPの増加につながる。

地方からグローバルへ

労働の必要な地方を舞台に技術を伸ばす。それをグローバルに展開

ディープラーニング×ものづくり： 「眼をもった機械」による日本の新たな産業競争力の実現へ

- 少子高齢化しており、労働力が不足している。
 - 頭脳労働は不足していない。「運動を伴う労働」のニーズが高い。
 - 農業従事者、建設・物流、介護、廃炉、熟練工の後継者、etc
- ディープラーニングが解決策になり得る。
 - 眼をもった機械：認識や運動の上達ができる機械・ロボット
 - ものづくりと相性がよく、日本の強みを活かせる。素材や駆動系も強い
- 新たな投資概念の必要性
 - 人への投資をいかに既存の枠組み・文化を踏まえてやるか
 - 企業からの投資を引き出し、そこに人材が流入する生態系を作るか
 - 例えば「学習工場」
- チャンスを捉えるには、正しく早く動いていくことが重要
 - ディープラーニング人材の育成
 - 事業・産業がどう変わるかを早期に検討
 - 社会全体で新しい未来像を描いていくこと