資料WG3-2

ディープラーニングの先にあるもの

東京大学 松尾 豊





ディープラーニング(2007-)

車・ロボット への活用

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)

● 自動運転 Pepper

ウェブ・ビッグデータ機械学習 ----

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

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)

タスクオントロジー

→ bot

探索 迷路・パズル

プランニング STRIPS チェス (1997) Deep Blue 将棋(2012-)

囲碁

電王戦

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)

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、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 識や音声認識、新薬開発に役立つ化合 結果だ」。専門家からは、口々に驚き 研究者から大きな注目を浴びている技 物の活性予測――。こうした技術の 術がある。コンピューターに人間と同 精度を競うコンテストで、ディープラ ディープラーニングは、人の神経回 じように経験に基づいた行動をさせる ーニングが過去の記録を大幅に塗り替 路をコンピューター上で模擬する「二 機械学習の一種で、「ディープラーニ え、次々と優勝を果たしている。 ューラルネットワーク」という技術を ング」と呼ばれる新手法だ。 「これほど飛躍的に精度が向上する 発展させたものだ。 インターネット社会を支える画像認 とは信じられない」「まさに衝撃的な 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 コンピューター上で神経回路を

入力層

各画素の色を

人の神経回路

シミュレーションする

画像中のエッジを HUSE

隠れ層

目や耳、鼻を

頭の全体像を

画像を入力するたびに、各ニューロンのつな

がり方が少しずつ変わる

奥にある層ほど高レベルの

相念を認識するようになる

出力層

筋の延暫を

民間

猫に反応する ニューロンが できる

画像認識におけるエラー率の変化(ILSVRC)

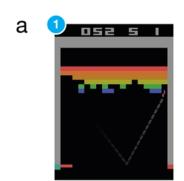
			Error
Before ディープ ラーニング After ディープ ラーニング		Imagenet 2011 winner (not CNN)	25.7%
		Imagenet 2012 winner	16.4% (Krizhesvky et al.)
		Imagenet 2013 winner	11.7% (Zeiler/Clarifai)
		Imagenet 2014 winner	6.7% (GoogLeNet)
		Baidu Arxiv paper:2015/1/3	6.0%
		Human: Andrej Karpathy	5.1%
		MS Research Arxiv paper: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

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とは数年前には考えられなかった。

運動の習熟: ディープラーニング+強化学習(2013-)

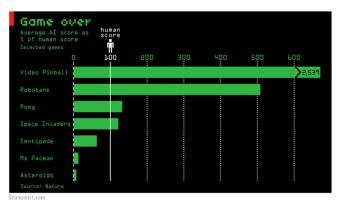
- 強化学習とは、行動を学習する仕組み。
 - 「報酬」が得られると、事前の行動を強化する。
 - 「状態」「行動」→「望ましさ(報酬ありなし)」
 - 古くからある技術だが、これまでは、「状態」を人間が定義してきた。
- 運動の習熟が可能に
 - 状態の認識に、ディープラーニングを使う。
 - DeepMindの研究者(D. Hassabisら)。その後、Googleが買収。
- 試行錯誤することによって、運動が習熟する
 - 最初は下手。繰り返すうちに、うまくなってくる。
 - 最終的には、ブロック崩しでの通路を作ったり、インベーダーゲームでの「名古屋撃ち」も。
 - 「全く同じプログラム」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





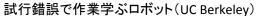






運動の習熟:ディープラーニング+強化学習が実世界へ(2015-)

- 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。高次な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

- AIにおける50年来のブレークスルー
 - モラベックのパラドックス:「子どものできることほど難しい」
 - 認識や運動の習熟は、何十年もずっとできなかった。
 - 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

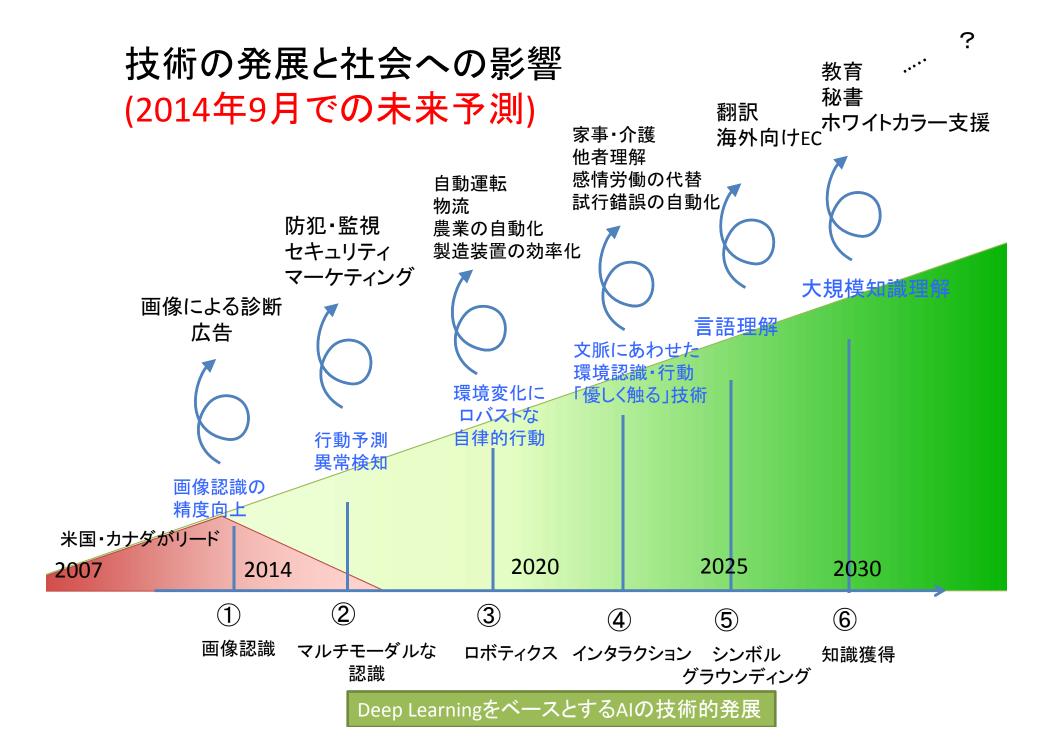
言語理解、白動翻訳

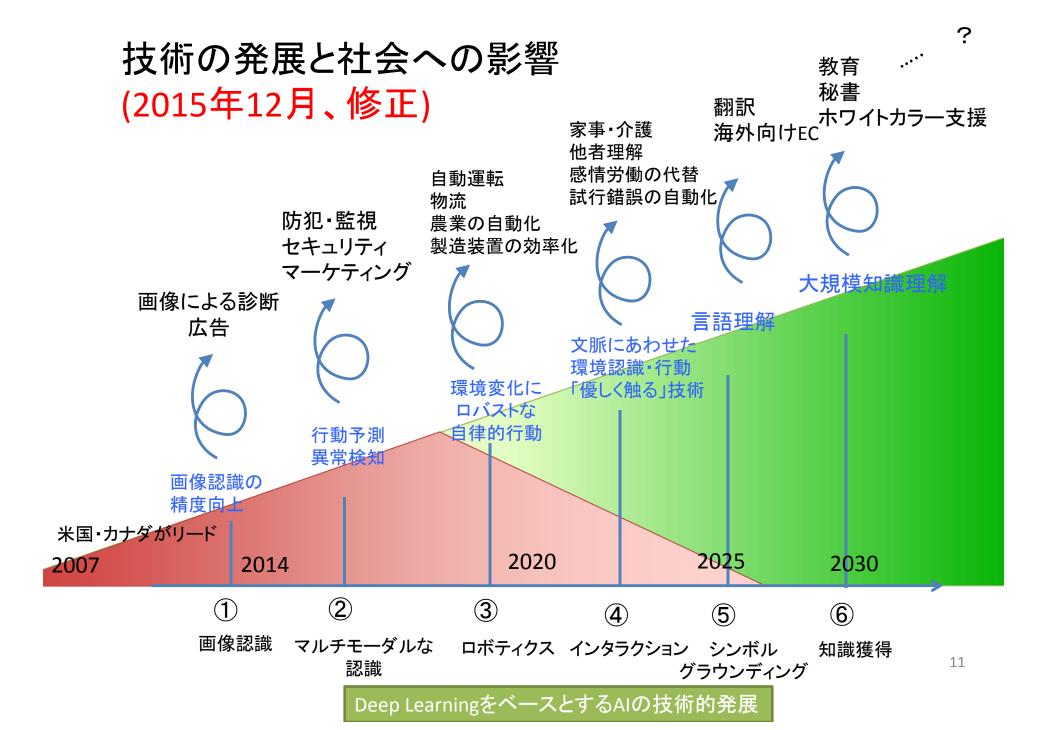
⑥言語からの知識獲得

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

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





強化学習とプランニングの融合(3の技術の難しいはずのところ)

- 動いたらどうなるかを予測して どう動くかを決める。
- アクションをはさんだオートエンコーダ



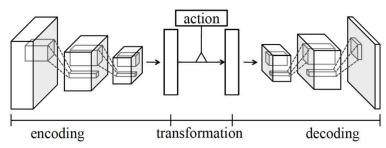




Prev. frame

Next frame

Prediction







Action

Non-Action

J. Oh et. al: Action-Conditional Video Prediction using Deep Networks in Atari Games, NIPS2015, Deep Reinforcement workshop, 2015

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."

Generating Images (2015.12-)



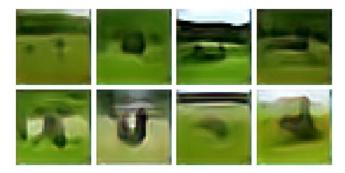
A very large commercial plane flying in <u>blue</u> skies.



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.

「子どもの人工知能」と「大人の人工知能」

- 大人の人工知能: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を突破口とする破壊的イノベーション
 - ディープラーニングを中心とする発展
 - 子どものできることができるようになっている
 - 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

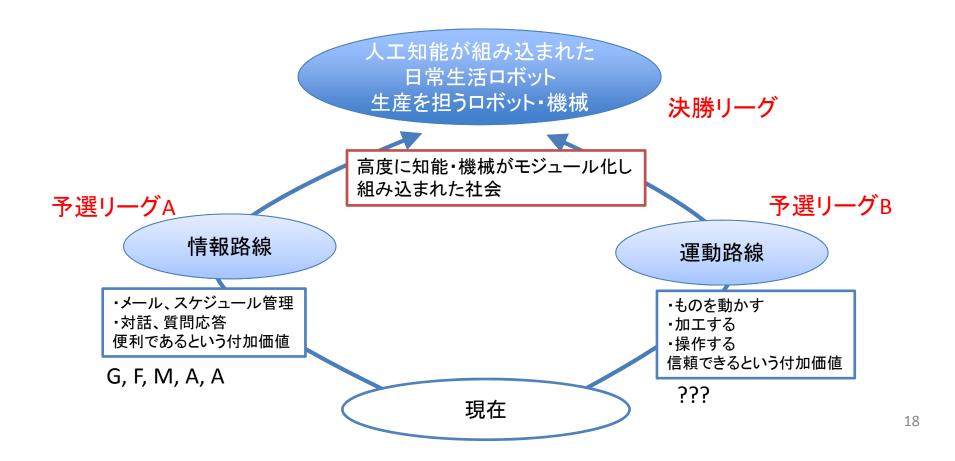
C: 計画立案を伴う運動

変化の本質は何か

- 画像・映像認識
 - 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。
 - 新たな事業が次々と。
- 運動の習熟
 - 我々は、機械は「機械的な動き」しかできない、ロボットは「ロボット的な動き」 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。)
 - 機械も習熟するし、ロボットも上達するようになる。
 - 自然物を相手にしているものは場面場面で状況が異なるので、そもそも自動 化が難しかった。それが自動化される。
 - 例えば、農業、建設、食品加工。
 - − さらには、日常生活のロボット、生産・仕事を担う機械・ロボットが実現される。

産業としてみたときの方向感

- 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、 特徴量生成の能力がない状況では対応できなかった。ここの自動化にどう至るかが鍵。
- 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: cf) 2015 ICML deep learning workshop panel
- 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ベ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いれば、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、動物なのかも知らされる。

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CFO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テックに売却した。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、カメラとダッシュポードが連携する。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<u>報じた</u>。



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。

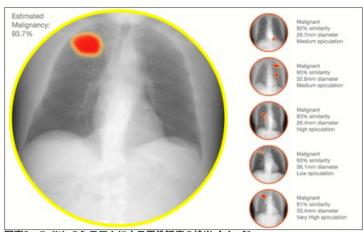


写真2●Enliticのシステムによる悪性腫瘍の検出イメージ

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れ、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る、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

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、いいところだけを見せる権利、悪いことをする権利、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)で議論。 社会全体で議論していきたい。

日本の社会課題を人工知能で解決する

- 農業分野に「習熟したロボット」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。

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

「ものづくり」を強みに、何とかすがりつくべき

- ウェブの分野で起こったこと
 - 1999年くらいは普通に研究していた。しかし、だれも検索エンジンの研究ができなくなった
 - 巨大な産業になり、多くの研究者が吸い込まれていった
 - この分野は、巨大産業を生み出すと、一気に科学技術も進んでしまう
 - GoogleやFacebookを前にした圧倒的な敗北感
- 「昔からやっていた」とか言っていても仕方がない
 - ウェブの世界ではそんなことは当たり前
 - 例えば、インターネットの動画配信はみんなやっていた。Youtubeが出るまで。
 - 検索エンジンもSNSもソーシャルゲームもLINEも。条件が揃って初めて大きな変革が起こる。
 - 子ども人工知能も同じ。これまでの研究は何も間違っていない。ただ、タイミング。いまは データも計算能力もアルゴリズムも揃っている。
- 「子どもの人工知能」は、広い意味でものづくりと相性がよい
 - 日本の強みを活かせる。素材や駆動系も強い。
 - 知能の研究において、認識の次に運動が重要である以上、学術的な進展としても正しい。
 - 何とかものづくりの強さを活かして、少しでも差を縮めたい。
 - 「ディープラーニングの先」とか言うのではなく、現実を見て、素直にキャッチアップすべし。
 - 例えば、NVIDIAの主催するGTCが主要イベントになりつつある。
 - アメリカの大学は空っぽになりつつある。

想定質問

- DL以外も重要では?
 - その通りです。ただ、難しかった特徴抽出のところが一気に伸びてるので、そこを基盤に変化が起こると思います。
- DLは、画像はいいけど、Xはできないですよね。
 - はい、その通りです。ですから、身体性や記号処理が重要なわけです。(昔からの議論の通り)
- DLが注目されすぎでは?
 - 者在的な可能性はおいておきましょう。すでに達成されていることだけ見ても、相当すごいです。
- 昔からやっている。
 - はい、その通りです。タイミングの問題です。
 - そうは言っても、細かい工夫もたくさん積み重なっています。
- 理論がない。黒魔術。
 - 急速に進展する領域は必ずそうで、あとから理論化されます。理論化している人もいっぱいいます。
- 三度目のブームもどうせ終わるのでは?
 - 評論家ですか?
 - プレイヤーとしては、大きな壁はないと思っていますし、ないと思ってやるべきと思います。
- 単に計算機の能力があがっただけでしょ。
 - はい。でもそれがいろいろ可能にしています。
- 流行にのっても仕方ない。
 - 評論家ですか?
 - 一研究者として、本質的な変化でないと信じる理由があれば、静観してればいいんじゃないでしょうか。
- 日本が勝てるの?
 - 厳しいと思います。それを認めた上での戦略は、あると思います。