Three Paradigms of Deep Learning
for Wide-Ranging AI Applications with Big Data

驱动大数据人工智能多种应用的三类深度学习模式

Li Deng

Chief Scientist of AI

Microsoft

(Research and ASG), Redmond, WA, USA

Keynote at CAAI (中国人工智能大会); August 26, 2016

Thanks go to many colleagues at Microsoft and collaborating universities

Deep learning

From Wikipedia, the free encyclopedia

Definition

Deep learning is a class of machine learning algorithms that

- use a cascade of **many layers of nonlinear processing**
- are part of the broader machine learning field of learning representations of data facilitating **end-to-end optimization**
- learn multiple levels of representations that correspond to **different levels of abstraction**
- ..., ...

Three Paradigms of Deep Learning

- Deep Supervised Learning
  – Paired input-output **big** training data for prediction
  – Paired output serves as "teacher" for corresponding input
- Deep Reinforcement Learning
  – Very weak "teacher" in the form of rewards; i.e. feedbacks (often distant) from environments
  – Q-learning computes "teaching signal" for training DNNs
- Deep Unsupervised Learning
  – Unpaired input-output **bigger** training data for prediction
    - but no teacher/label per se for each input token
  – Non-prediction tasks (clustering, dimensionality reduction, interpretation & understanding of data for transfer/multitask learning ...)

Two Types of Big Data for Deep Learning

- Deep Supervised Learning
  - **Paired input-output big training data** (costly)
  - Paired output serves as "teacher" for corresponding input
- Deep Reinforcement Learning
  - Very weak "teacher" in the form of rewards (often distant)
- Deep Unsupervised Learning
  - **Unpaired input-output bigger training data** (almost no cost; everywhere --- 取之不尽用之不竭)
    - but no teacher for each input token
  - Non-prediction tasks (clustering, interpretation & understanding of data…)

AI = 感知+认知

AI = machine perception (speech, image, video, gesture, touch...)

+ machine cognition (natural language, *reasoning, attention, memory*/learning, knowledge, decision making, action/planning/robotics, interaction/conversation/ChatBot,...)

Deep Learning for AI Applications

- Current Successes of Deep Learning
  - Speech (recognition), 2010 (supervised, bruteforce)
  - Image (recognition), 2012 (supervised, bruteforce)
  - NLP (translation, understanding, text QA), 2014 (supervised)
  - Multimodal image-text (image captioning/QA), 2015 (supervised)
  - Games (AlphaGo), 2016 (reinforcement)

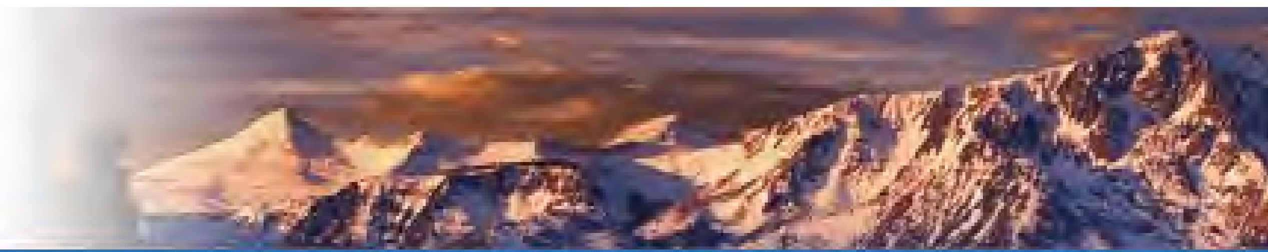
Unsupervised (in the future)

- Deep Learning --- Next waves upon us
  - Information retrieval (2017, predicted by Chris Manning at SIGIR 2016)
  - Mobile UI and dialogues (conversational AI bots, chatbots) --- reinforcement learning
  - Practically useful personal assistants (next-generation Cortana, SIRI, GoogNow, Alexa)
  - Business analytics (e.g. sales and marketing)
  - Finance (hedge funds, excluding “flash boy” type)
  - Robotic/car control, and physical plant control (e.g. energy saving, logistics optimization)
  - Medical and health
  - Security
  - Art and science; video transcription, AI-generated movie scripts
  - Automated journalism and legal assistance
  - Communication networks (e.g. optimal routing)
  - A full range of enterprise scenarios …, …

Deep Supervised Learning for Speech Recognition

- with massive paired input(acoustics)-output(text) big data for training speech recognizers

Neural Information
Processing Systems
Foundation



Li Deng, Dong Yu, Geoffrey Hinton

Microsoft Research; Microsoft Research; University of Toronto

Deep Learning for Speech Recognition and Related Applications

7:30am - 6:30pm Saturday, December 12, 2009

Location: Hilton: Cheakamus

Abstract: Over the past 25 years or so, speech recognition technology has b
dominated by a "shallow" architecture --- hidden Markov models (HMMs). Sig
technological success has been achieved using complex and carefully engine
of HMMs. The next generation of the technology requires solutions to remaini
challenges under diversified deployment environments. These challenges, not
addressed in the past, arise from the many types of variability present in the
generation process. Overcoming these challenges is likely to require "deep" a
with efficient learning algorithms. For speech recognition and related sequent
recognition applications, some attempts have been made in the past to devel
computational architectures that are "deeper" than conventional HMMs, such

The Universal Translator …comes true!

Deep learning technology enabled speech-to-speech translation

The New York Times

Scientists See Promise in Deep-Learning Programs

John Markoff

November 23, 2012

Tianjin, China, October, 25, 2012

A voice recognition program translated a speech given by Richard F. Rashid, Microsoft's top scientist, into Mandarin Chinese.

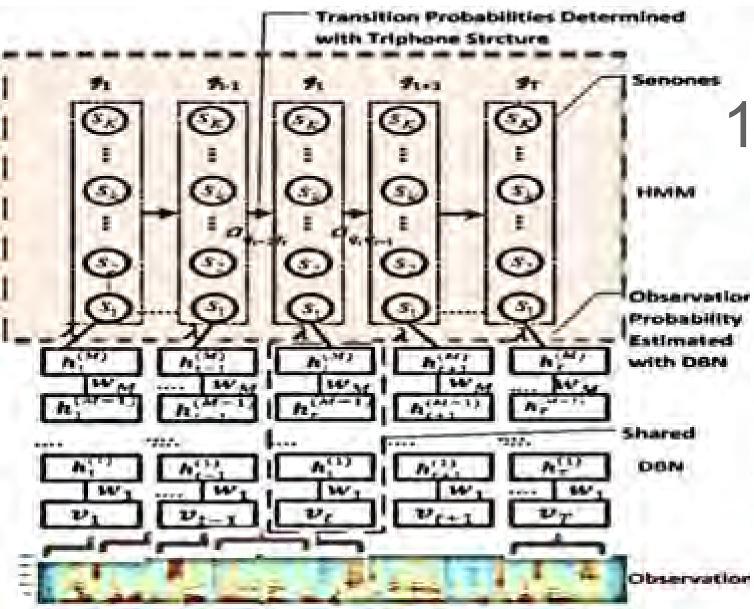
Microsoft Research

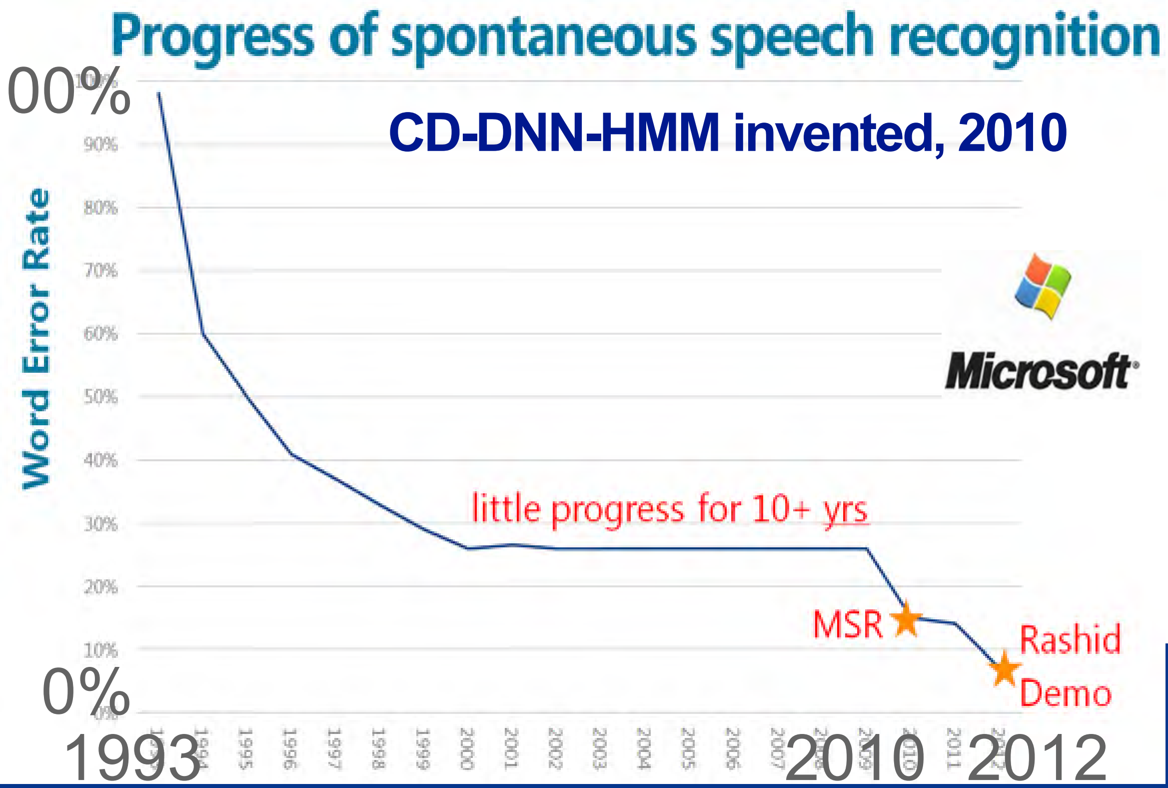
Deep belief networks for phone recognition, NIPS, December 2009; 2012

Investigation of full-sequence training of DBNs for speech recognition., Interspeech, Sept 20
Binary coding of speech spectrograms using a deep auto-encoder, Interspeech, Sept 2(
Roles of Pre-Training & Fine-Tuning in CD-DBN-HMMs for Real-World ASR, NIPS, Dec. 20
Large Vocabulary Continuous Speech Recognition With CD-DNN-HMMS, ICASSP, April 20
Conversational Speech Transcription Using Contxt-Dependent DNN, Interspeech, *Aug. 2011*

Making deep belief networks effective for LVCSR, ASRU, Dec. 2011
Application of Pretrained DNNs to Large Vocabulary Speech Recognition., ICASSP, 2012
【胡郁】讯飞超脑 2.0 是怎样炼成的？2011, 2015

Later years with rapid progress,

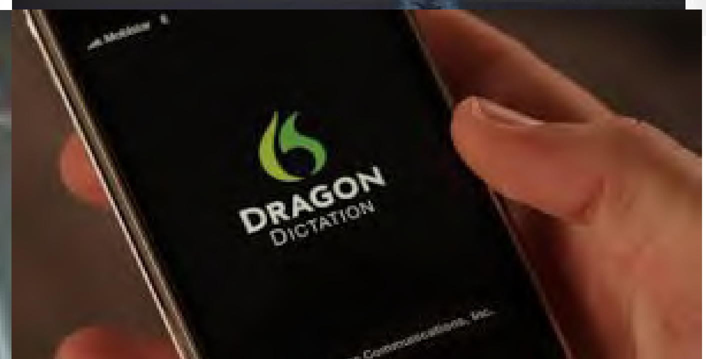




Microsoft Research

Across-the-Board Deployment of DNN in Speech Industry

(+ in university labs & DARPA programs) **(2012-2014)**



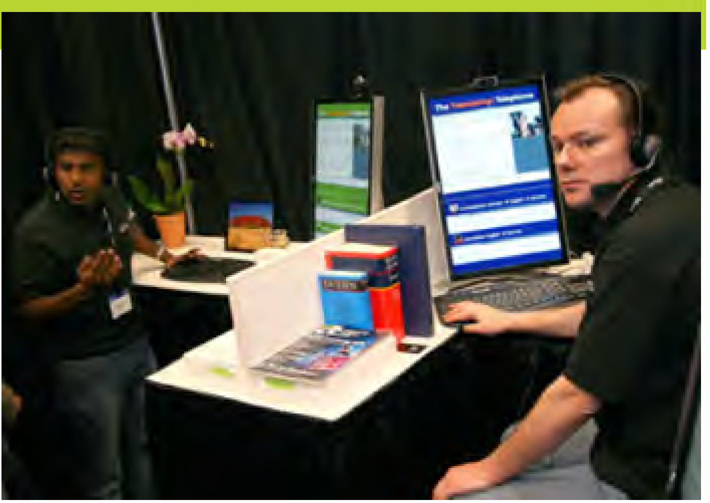
Enabling Cross-Lingual Conversations in Real Time

Microsoft Research

May 27, 2014 5:58 PM PT

View milestones on the path to Skype Translator

#speech2speech

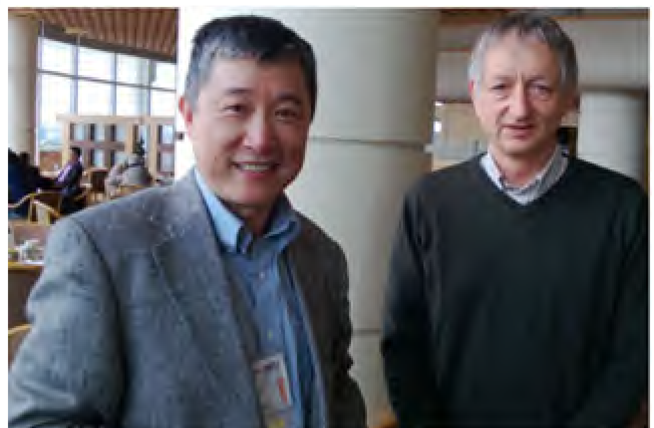


Microsoft Research

WIRED

ROBERT MCMILLAN BUSINESS 12.17.14 1:19 PM

HOW SKYPE USED AI TO BUILD ITS AMAZING NEW LANGUAGE TRANSLATOR



CNN

Taking a cue from science fiction, Microsoft demos 'universal translator'

In the academic world

nature International weekly journal of science

Deep learning

Yann LeCun, Yoshua Bengio & Geoffrey Hinton

Affiliations | Corresponding author

Nature **521**, 436–444 (28 May 2015) | doi:10.1038/nature14539

Received 25 February 2015 | Accepted 01 May 2015 | Published online

Geoffrey Hinton, Li Deng, Dong Yu, George E. Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N. Sainath, and Brian Kingsbury

"This joint paper (2012) from the major speech recognition laboratories details the first major industrial application of deep learning."

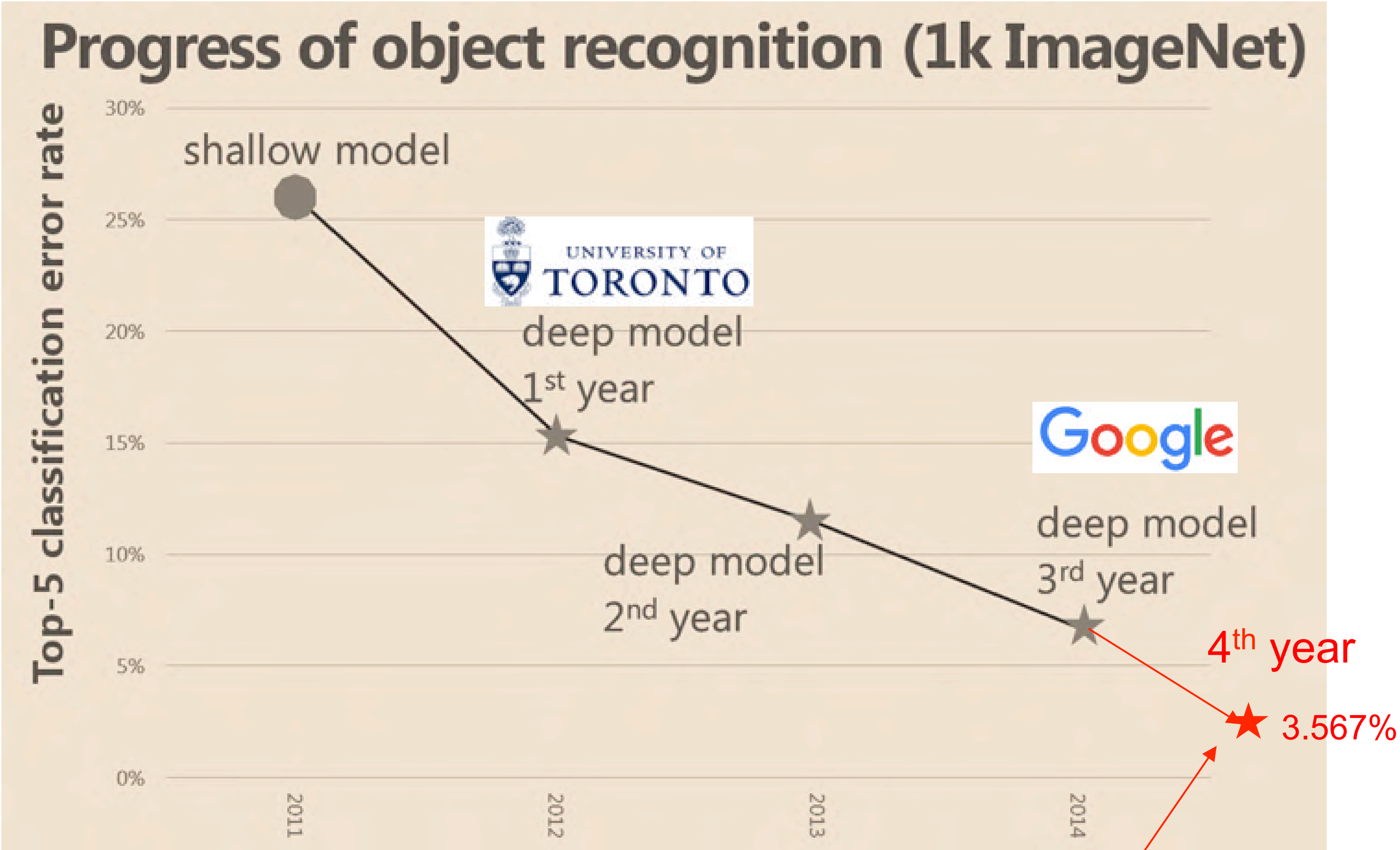
Deep Neural Networks for Acoustic Modeling in Speech Recognition

[The shared views of four research groups]

Deep Supervised Learning also Shattered Image Recognition (in the same way)

(since 2012)

IMAGENET Competition



Announced Dec. 2015

2012 - 2015

Super-deep: 152 layers

Microsoft Research

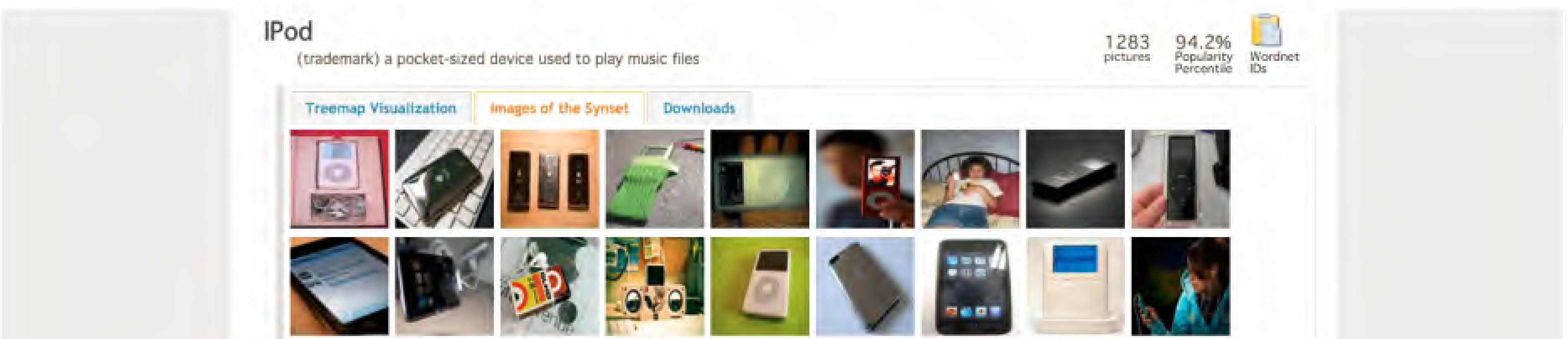
WIRED

MICROSOFT NEURAL NET SHOWS DEEP LEARNING CAN GET WAY DEEPER

Microsoft beats Google, Intel, Tencent, and Qualcomm in image recognition competition

JORDAN NOVET DECEMBER 10, 2015 4:14 PM

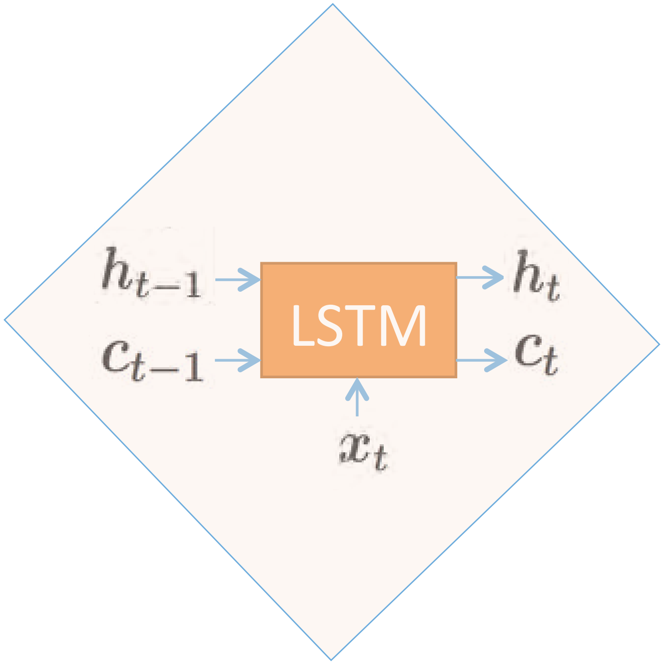
TAGS: ARTIFICIAL INTELLIGENCE, DEEP LEARNING, IBM, IMAGE RECOGNITION, IMAGENET, MICROSOFT, MICROSOFT RESEARCH, NVIDIA, SOFTLAYER

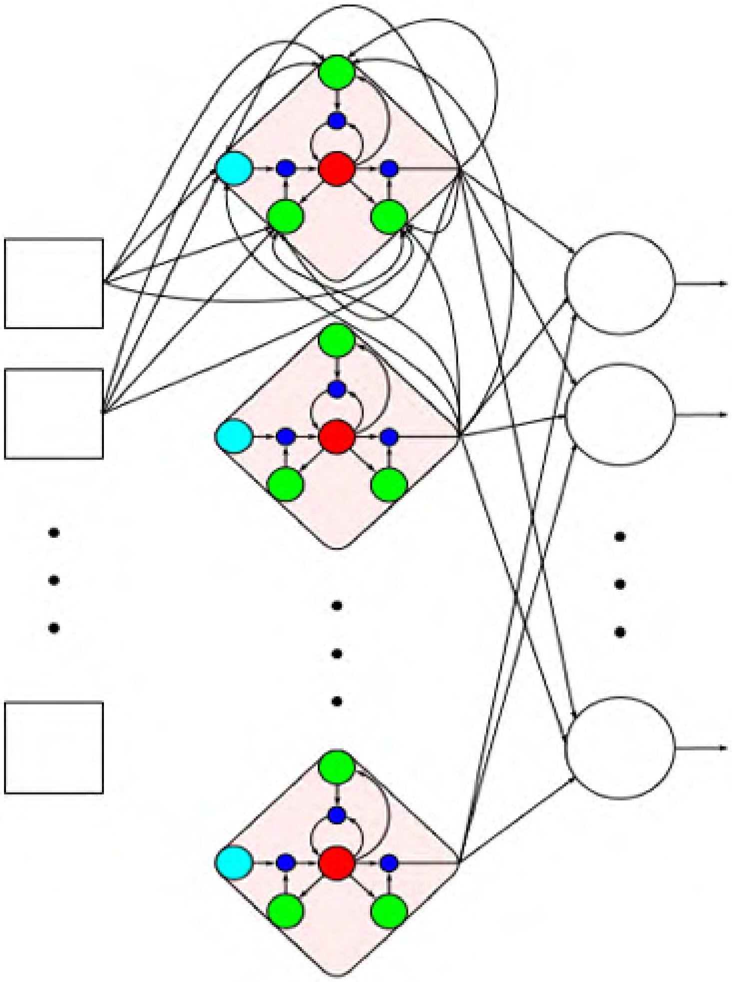


Supervised Deep Learning for Machine Cognition

--- Memory & attention applied to machine translation

Long Short-Term Memory RNN

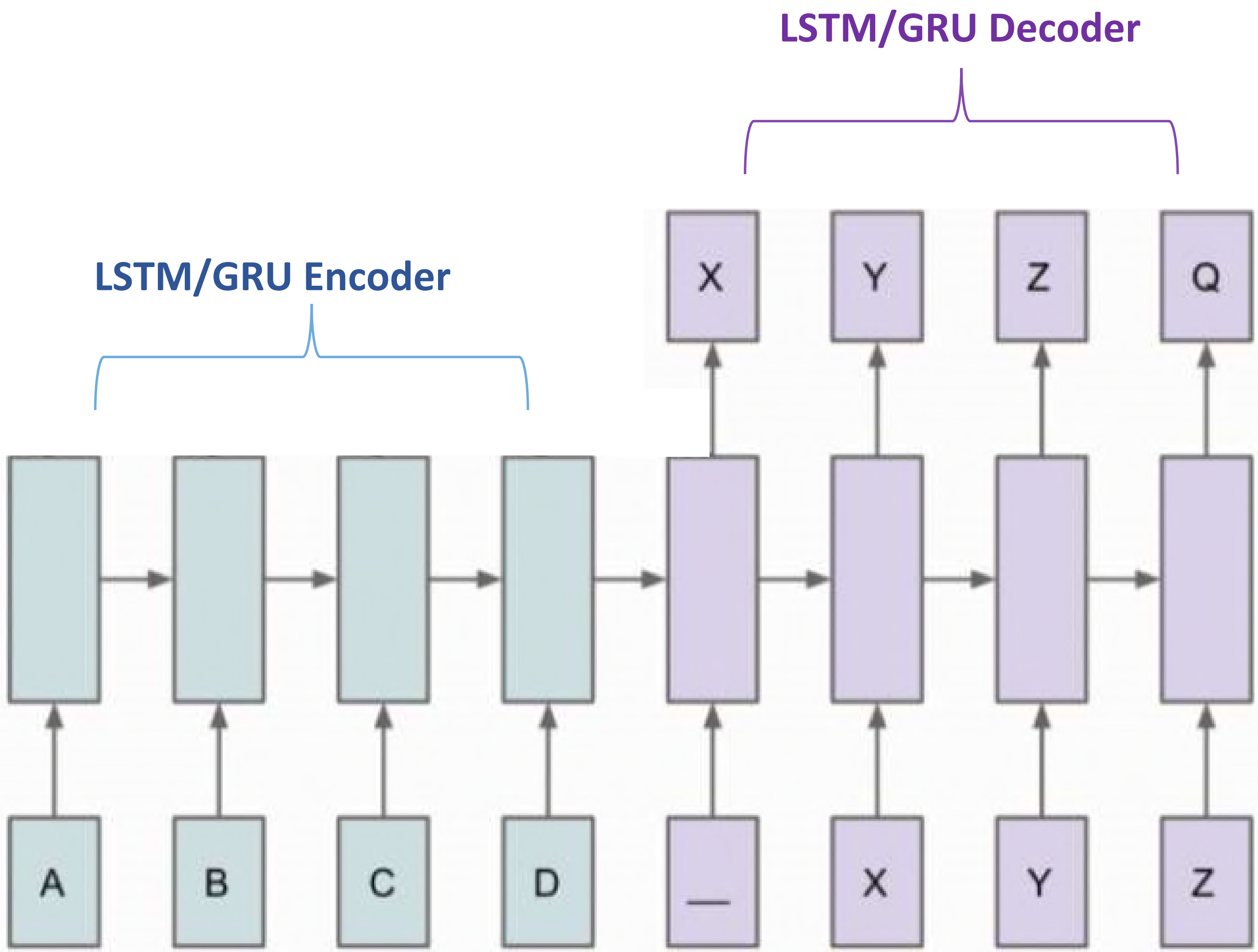




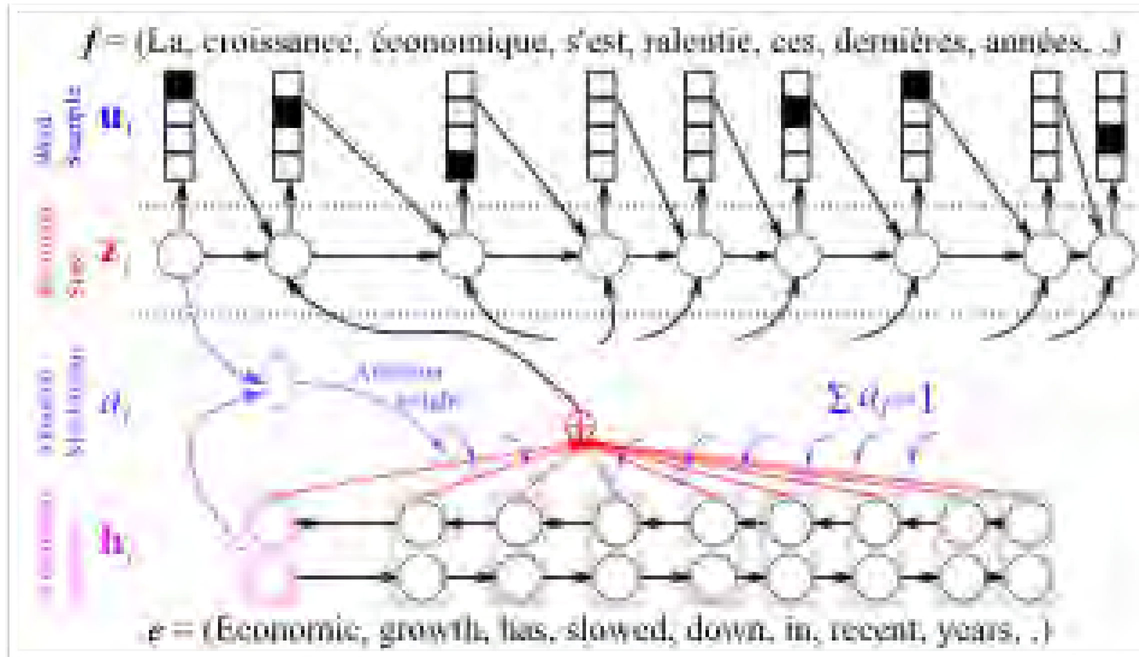
(Hochreiter & Schmidhuber, 1997)

Seq-2-Seq Learning (Neural Machine Translation)

[Sutskever, Vinyals, Le, NIPS, 2014]



Neural Machine Translation with Attention



Attention-based Model

- Encoder: Bidirectional RNN
  - A set of *annotation* vectors

    $\{h_1, h_2, \ldots, h_T\}$

- Attention-based Decoder

  (1) Compute attention weights

  $\alpha_{t',t} \propto \exp(e(z_{t'-1}, u_{t'-1}, h_t))$

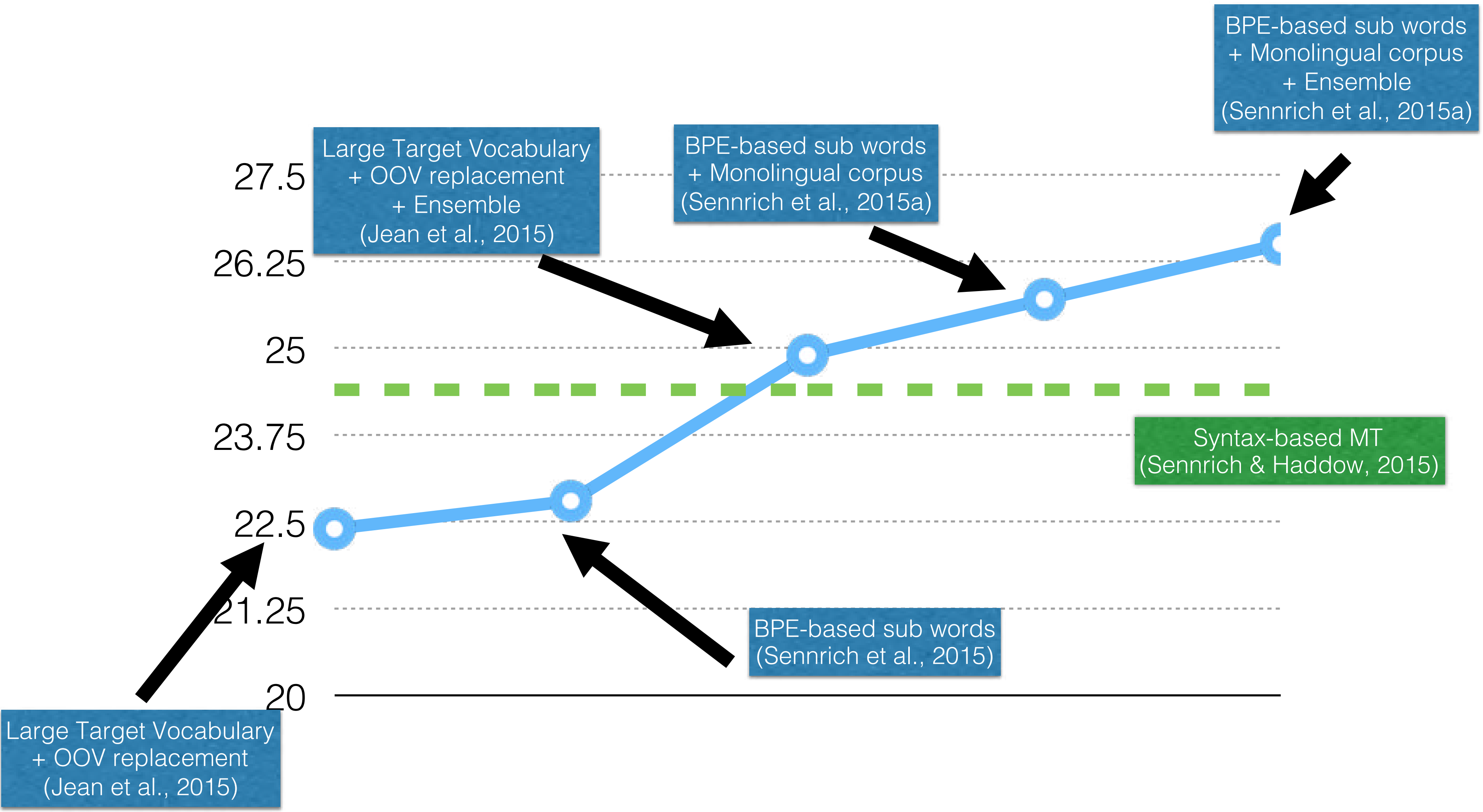
  (2) Weighted-sum of the annotation vectors

  $c_{t'} = \sum_{t=1}^{T} \alpha_{t',t} h_t$

  (3) Use $c_{t'}$ to replace "though vector" h_T

(modified from: Kyunghyun Cho, 2016)

Benchmark: WMT'15 En-De

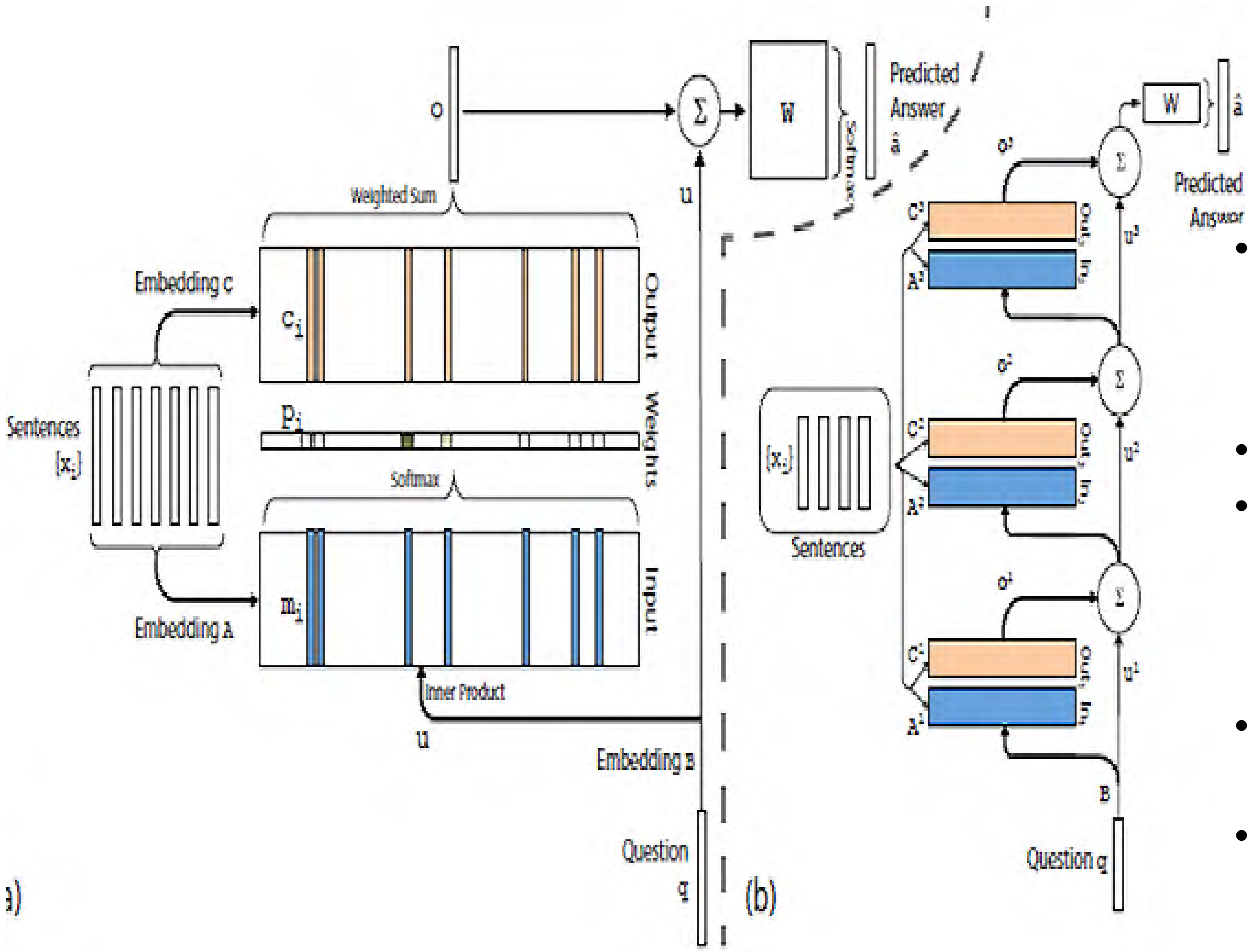


(modified from: Kyunghyun Cho)

Supervised Deep Learning for Machine Cognition

--- Neural reasoning: memory networks

Memory Networks for Reasoning



- Rather than placing "attention" to part of a sentence, it can be placed to cognitive space with many sentences
- This allows "reasoning"
- Embedding input

$$m_i = Ax_i$$

$$c_i = Cx_i$$

$$u = Bq$$

- Attention over memories

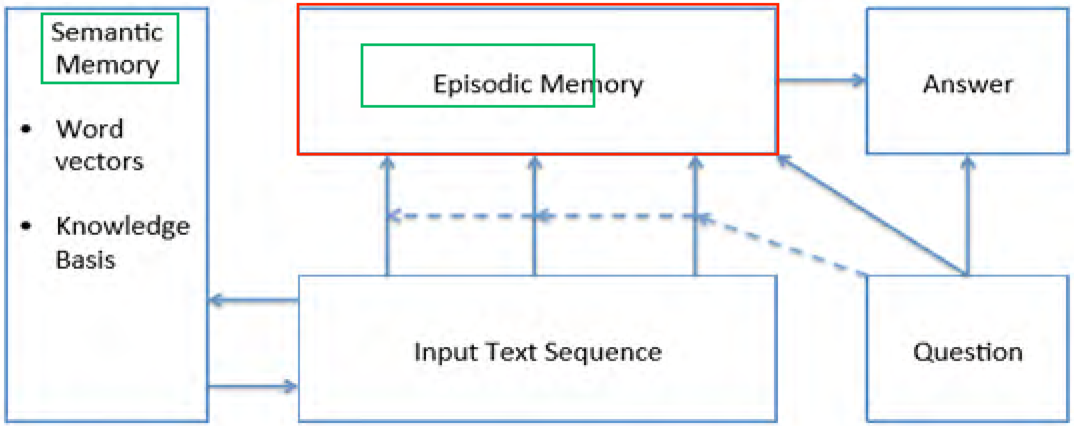
$$p_i = \text{softmax}(u^T m_i)$$

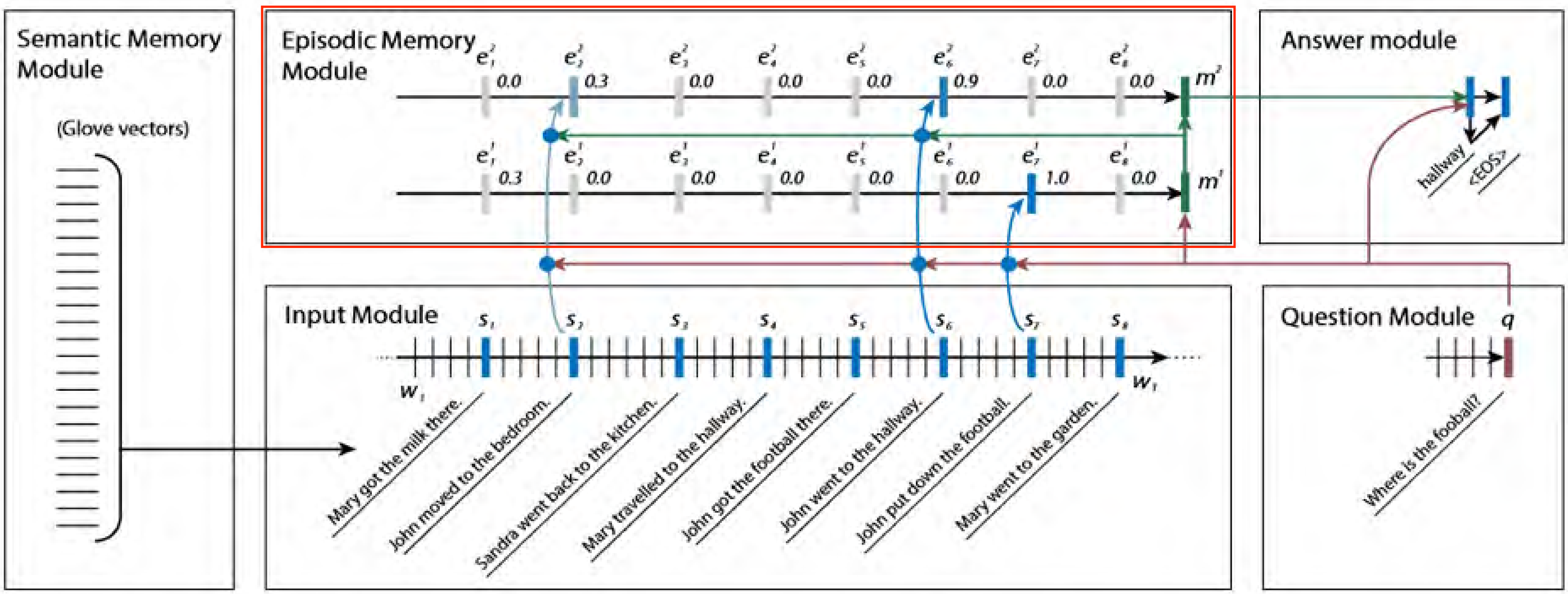
- Generating the final answer

$$o = \sum_i p_i c_i$$

$$a = \text{softmax}(W(o+u))$$

[Sukhbaatar, Szlm, Weston, Fergus: "End-to-end memory networks," NIPS, 2





[
Kumar, Irsoy, …Socher: "Ask me anything: **Dynamic Memory Networks** for NLP," NIPS, 2015]

Deep Reinforcement Learning

== Deep Learning (for representing gigantic "state-space")
+ Reinforcement Learning (optimal decision making)

- Distant, weak teacher via "evaluative" rewards
- Need for exploration (not for supervised learning)
- More important AI applications to come (than supervised learning)

nature International weekly journal of science

LETTER

doi:10.1038/nature14236

Human-level control through deep reinforcement learning

Volodymyr Mnih[1]*, Koray Kavukcuoglu[1]*, David Silver[1]*, Andrei A. Rusu[1], Joel Veness[1], Marc G. Bellemare[1], Alex Graves[1], Martin Riedmiller[1], Andreas K. Fidjeland[1], Georg Ostrovski[1], Stig Petersen[1], Charles Beattie[1], Amir Sadik[1], Ioannis Antonoglou[1], Helen King[1], Dharshan Kumaran[1], Daan Wierstra[1], Shane Legg[1] & Demis Hassabis[1]

The theory of reinforcement learning provides a normative account[1], deeply rooted in psychological[2] and neuroscientific[3] perspectives on animal behaviour, of how agents may optimize their control of an environment. To use reinforcement learning successfully in situations approaching real-world complexity, however, agents are confronted

agent is to select actions in a fashion that maximizes cumulative future reward. More formally, we use a deep convolutional neural network to approximate the optimal action-value function

$$Q^*(s,a) = \max_\pi \mathbb{E}\left[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \ldots \mid s_t = s,\ a_t = a,\ \pi\right],$$

Reinforcement learning from "non-working" to "working", due to Deep Learning (much like DNN for speech)

Deep Reinforcement Learning for Games
--- optimizing long-term values

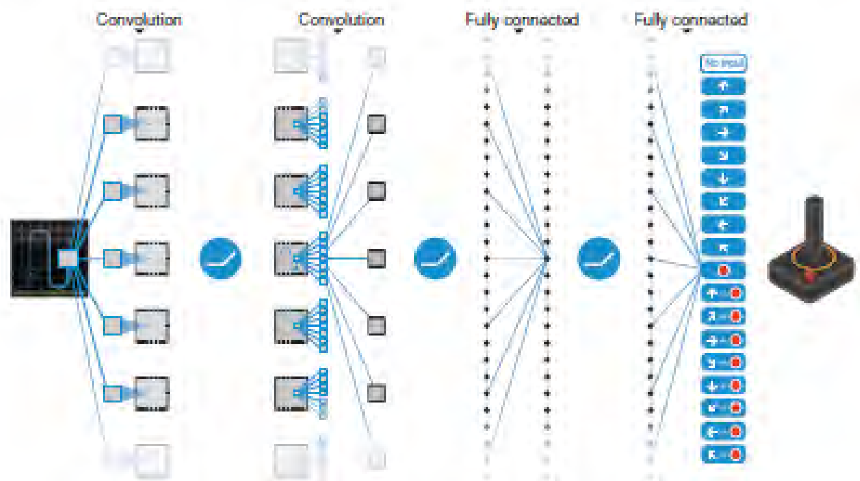
	Short-term	Long-term
Playing the Breakout game		
Optimizing Business Decision Making	*Maximize immediate reward*	*Optimize life-time revenue, service usages, and customer satisfaction*

AlphaGo

Self play to improve skills

Deep Q-Network (DQN)

mapping raw screen pixels



to predictions of final score for each of 18 joystick actions

- Input layer: image vector of s
- Output layer: a single output Q-value for each action a, $Q(s,a,\theta)$
- DNN parameters: θ

nature International weekly journal of science

ARTICLE PREVIEW

view full access options ▸

NATURE | ARTICLE

日本語要約

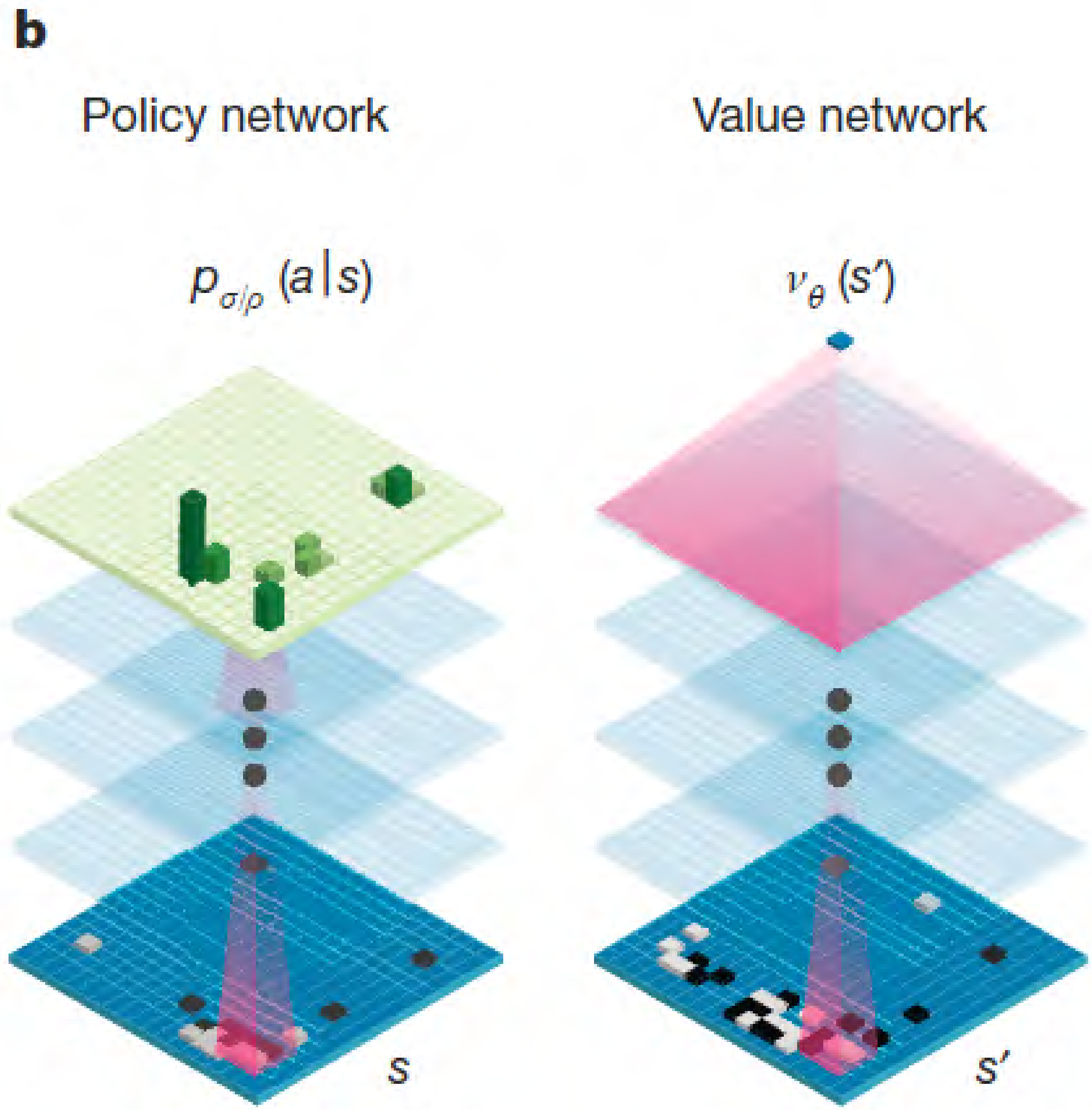
Mastering the game of Go with deep neural networks and tree search

David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel & Demis Hassabis

Affiliations | Contributions | Corresponding authors

Nature **529**, 484–489 (28 January 2016) | doi:10.1038/nature16961
Received 11 November 2015 | Accepted 05 January 2016 | Published online 27 January 2016

DNN architecture use AlphaGo

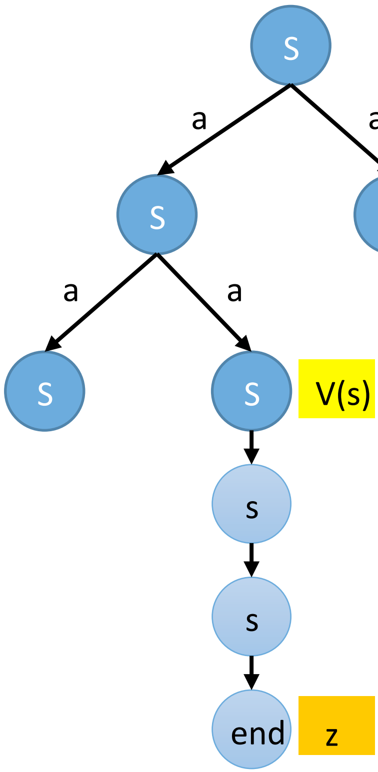


Analysis of four DNN

DNNs	Properties	Architecture	Additional details
$\pi\downarrow SL\ as$	Slow, accurate stochastic supervised learning policy, trained on 30M (s,a) pairs	13 layer network; alternating ConvNets and rectifier non-linearities; output dist. over all legal moves	Evaluation time: 3 ms Accuracy vs. corpus: 57% Train time: 3 weeks
$\pi\ \downarrow SL\ as$	Fast, less accurate stochastic SL policy, trained on 30M (s,a) pairs	Linear softmax of small pattern features	Evaluation time: 2 us Accuracy vs. corpus: 24%
$\pi\downarrow RL\ as$	Stochastic RL policy, trained by self-play	Same as $\pi\downarrow SL$	Win vs. $\pi\downarrow SL$ 80%
$V(s)$	Value function: % chance of $\pi\downarrow RL$ winning by starting in state *s*	Same as $\pi\downarrow SL$ but with one output (% chance of winning)	15K less computation than evaluating $\pi\downarrow RL$ with roll-outs

Monte Carlo Tree Search

AlphaGo



$$\pi(s) = \arg\max_a Q(s,a)$$

$$Q(s,a) = Q'(s,a) + u(s,a)$$

Roll-out estimate

Exploration bonus

Mixture weight

$$Q'(s,a) = \frac{1}{N(s,a)} \sum_i \left[(1-\lambda)V(s_L^i) + \lambda z_L^i\right]$$

# of times action *a* taken in state *s*

Value function computed in advance

Win/loss result of 1 roll-out with π_{SL} as

$$u(s,a) = c \cdot \pi_{SL}(a|s) \frac{\sqrt{\sum_b N(s,b)}}{1+N(s,a)}$$

- Think of this MCTS component as a highly efficient “decoder”, concept familiar to ASR
- -> A* search and fast match in speech recognition literature in 80’s-90’s
- This is tree search (GO-specific), not graph search (A*)
- Speech is a relatively simple signal → sequential beam search sufficient, not need for A* or tree
- Key innovation in AlphaGO: “scores” in MCTS computed by DNNs with RL

NEWS EVENTS RESEARCH

How deep reinforcement learning can help chatbots

LI DENG, MICROSOFT AUGUST 1, 2016 6:10 PM

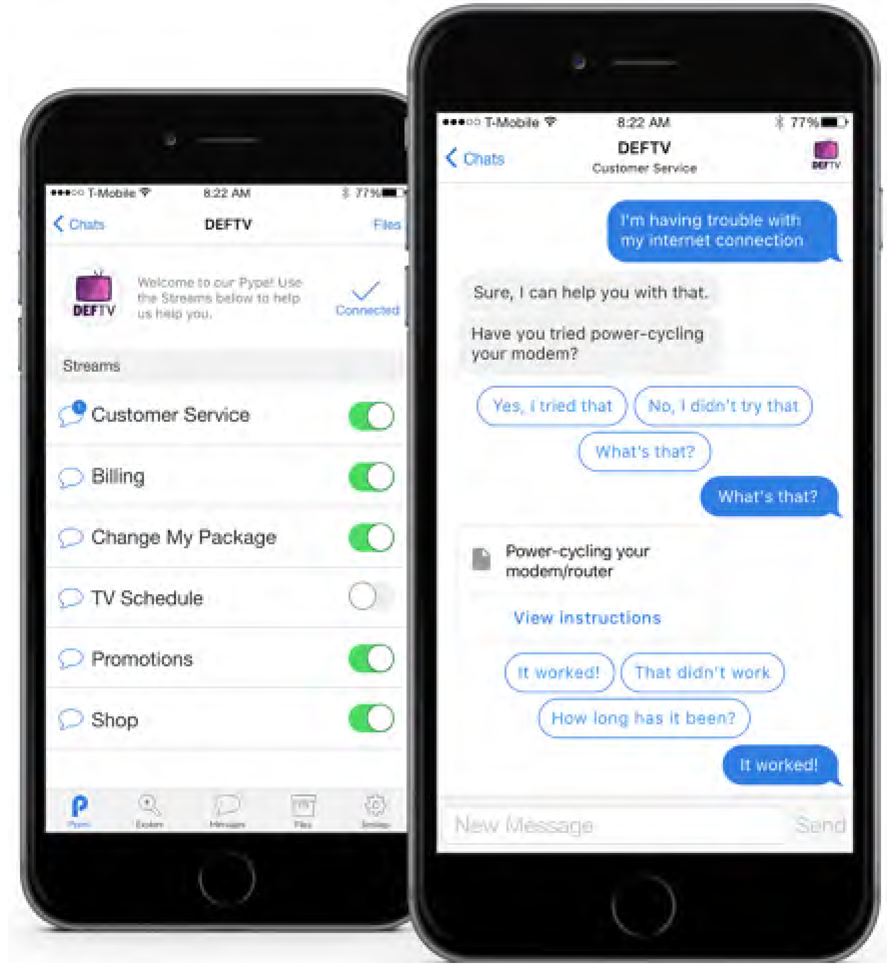
TAGS: BOT PLATFORM, BOTS, CHATBOTS, DEEPMIND, GOOGLE, MICROSOFT

What is wrong with apps and web models?

Conversation as an emerging paradigm for mobile UI

Bots as intelligent conversational interface agents

Three types of A.I. conversational bots



Deep Reinforcement vs. Supervised Learning

- Ostensible similarity to structured supervised learning
- Use of dynamic programming (DP) by both
- DP in Reinforcement Learning: Q-learning for credit assignment over long distance
- DP for Structured SL: efficient alignment for matching sequence targets (teachers)
- Different ways of DP approximation
- Need for exploration or otherwise

Deep UNsupervised Learning

- Input-output samples do not need to match for training
- Huge practical benefit: No cost to create training data
- Both input and output data are found naturally
- E.g. use of millions of hrs of speech for speech recognition
- Millions of hrs of video for video story telling
- All images in the web for image captioning
- Etc, etc.

"No Free Lunch" for Deep Unsupervised Learning

- Intensive research is needed --- How the human mind works
- Major frontier for deep learning and AI
- Our approach: integrate diverse sources of world knowledge and application-domain knowledge into a single, consistent framework

- To effectively exploit and integrate knowledge of:
  - Strong output statistics and structure (motivated by cryptography research)
  - Strong input statistics and structure
    - traditional unsupervised learning
  - Relationship/Mapping from output to input
    - generative model world; Bayesian learning, etc.
  - Relationship/Mapping from input to output
    - discriminative model world; deep neural nets
  - Relationship between input/output from another domain
    - transfer learning; compositionality of in much of real world data
  - Distributional properties of input and output sequences
    - smart sequence prediction exploiting motion/temporal cues (video, word embedding)

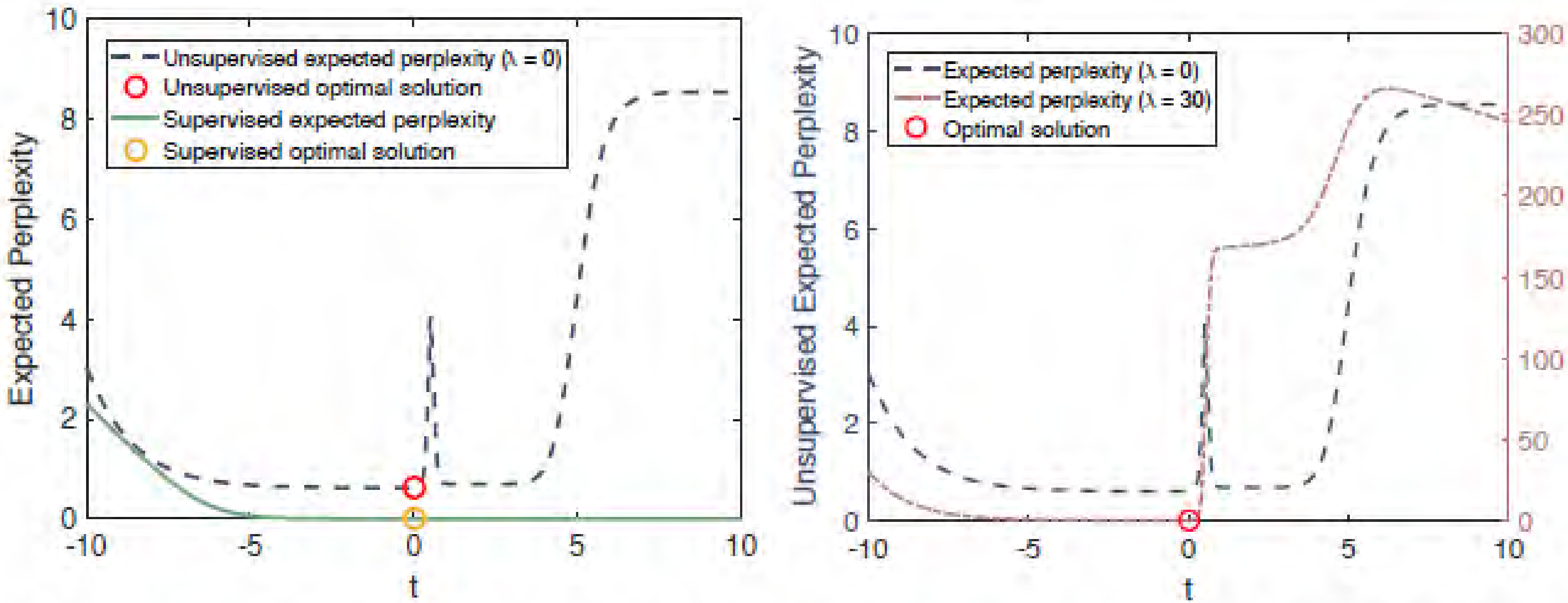
Exploiting output statistics and structure

$$
\begin{aligned}
\mathbb{E}\left[\ln p(y_1,\ldots,y_T)\middle|x_1,\ldots,x_T\right] &= \mathbb{E}\left[\sum_{t=1}^{T}\ln p(y_t|y_{t-1},\ldots,y_1)\middle|x_t,\ldots,x_1\right] \\
&= \sum_{(y_t,y_{t-1},\ldots,y_1)}\prod_{t=1}^{T}p(y_t|x_t,W_d)\sum_{t=1}^{T}\ln p(y_t|y_{t-1},\ldots,y_1) \\
&= \sum_{t=1}^{T}\prod_{\tau=1}^{t-1}p(y_\tau|x_\tau)\sum_{y_t}p(y_t|x_t)\ln p(y_t|y_{t-1},\ldots,y_1) \\
&= \sum_{t=1}^{T}\mathbb{E}\left[\sum_{y_t}p(y_t|x_t)\ln p(y_t|y_{t-1},\ldots,y_1)\middle|x_{t-1},\ldots,x_1\right]
\end{aligned}
$$

$$
\max_{W_d,W_g}\sum_{t=1}^{T}\left\{\mathbb{E}\left[\sum_{y_t}p(y_t|x_t)\ln p(y_t|y_{t-1},\ldots,y_1)\middle|x_{t-1},\ldots,x_1\right]+\lambda\sum_{y_t}p(y_t|x_t,W_d)\ln p(x_t|y_t,W_g)\right\}
$$

(Chen et al., ArXiv 2016)

Looking into the nature of the difficulty of UL



(a) Unsupervised vs supervised costs (b) The importance of regularization

- Much more difficult optimization problem in UL than SU
- Huge barrier that prevents from reaching the global optimum
- Quantum computing comes to rescue? (tunneling effect?)

(Chen et al., ArXiv 2016)

The Future of AI

- One AI engine that solves a wide range of AI problems with a unified first principle (general AI)
- Based on deep unsupervised, reinforcement, and transfer learning with little or no human labels
- Harmonize symbolic and neural representations & computation
- Automatically acquire, create, and grow knowledge through interactions with the world
- Better than human intelligence in many ways (due to much stronger computing power)
- Many successful applications along the way

Thank you!

Q/A

新书签售活动

时间：13:15—13:30

地点：B14，博文视点展位