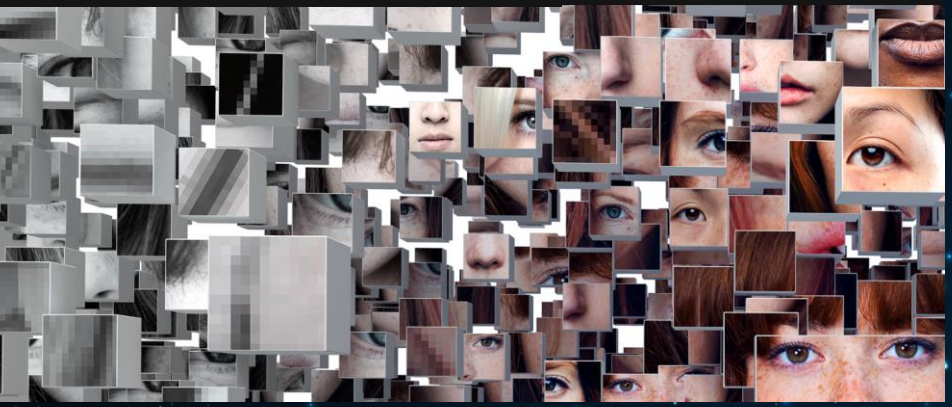
深度学习十年简史和人工智能未来展望

Li Deng

Microsoft AI and Research
Redmond, WA, USA

Keynote at CCL, Yantai,
October 16, 2016

FORTUNE



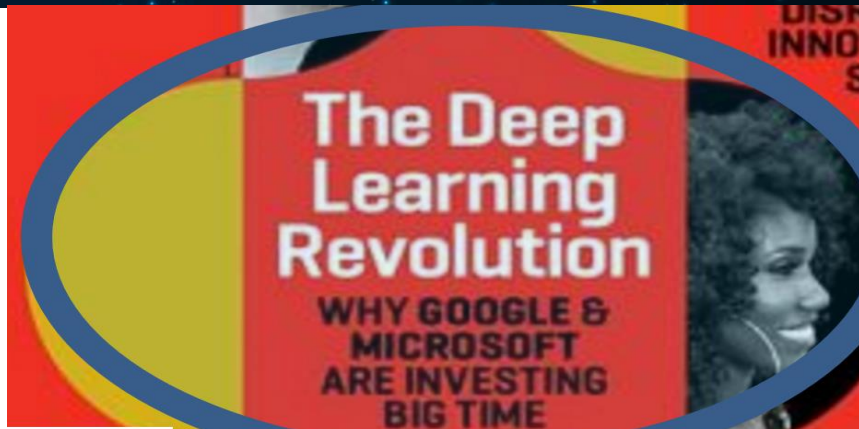
http://fortune.com/ai-artificial-intelligence-deep-machine-learning/

By Roger Parloff

Illustration by Justin Metz

SEPTEMBER 28, 2016, 5:00 PM EDT

WHY DEEP LEARNING IS SUDDENLY CHANGING YOUR LIFE



A GLOSSARY OF ARTIFICIAL-INTELLIGENCE TERMS

- ARTIFICIAL INTELLIGENCE

AI is the broadest term, applying to any technique that enables computers to mimic human intelligence, using logic, if-then rules, decision trees, and machine learning (including deep learning).

- MACHINE LEARNING

The subset of AI that includes abstruse statistical techniques that enable machines to improve at tasks with experience. The category includes deep learning.

- DEEP LEARNING

The subset of machine learning composed of algorithms that permit software to train itself to perform tasks, like speech and image recognition, by exposing multilayered neural networks to vast amounts of data.

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**
- **…, …**

KEY MOMENTS IN DEEP-LEARNING HISTORY

Mid-1990s
Neural nets fall into disfavor again, eclipsed by other machine-learning techniques.

2006 (added)

Deep Belief Networks (DBN)

(Hinton, Salakhudinov, Osindero, Teh)

2007
Fei-Fei Li founds ImageNet and begins assembling a database of 14 million labeled images that can be used for machine-learning research.

2011
Microsoft introduces（DNN） neural nets into its speech-recognition features.

2011
IBM's Watson beats two champions at Jeopardy using traditional AI techniques.

2012
JUNE
Google Brain publishes the "cat experiment." A neural net, shown 10 million unlabeled YouTube images, has trained itself to recognize cats.

AUGUST
Google introduces neural nets into its speech-recognition features.

OCTOBER
A neural net designed by two of Hinton's students wins the annual ImageNet contest by a wide margin.

2013
MAY
Google improves photo search using neural nets.

2014
JANUARY
Google acquires DeepMind, a startup specializing in combining deep learning and reinforcement learning, for $600 million.

2015
DECEMBER
A team from Microsoft, using neural nets, outperforms a human on the ImageNet challenge.

2016
MARCH
DeepMind's AlphaGo, using deep learning, defeats world champion **Lee Sedol** in the Chinese game of go, four games to one.

Science AAAS

1994-1995

SHARE

REPORTS

The "wake-sleep" algorithm for unsupervised neural networks

GE Hinton, P Dayan, BJ Frey, RM Neal

+ Author Affiliations

Science 26 May 1995:
Vol. 268, Issue 5214, pp. 1158-1161
DOI: 10.1126/science.7761831

Neural Networks

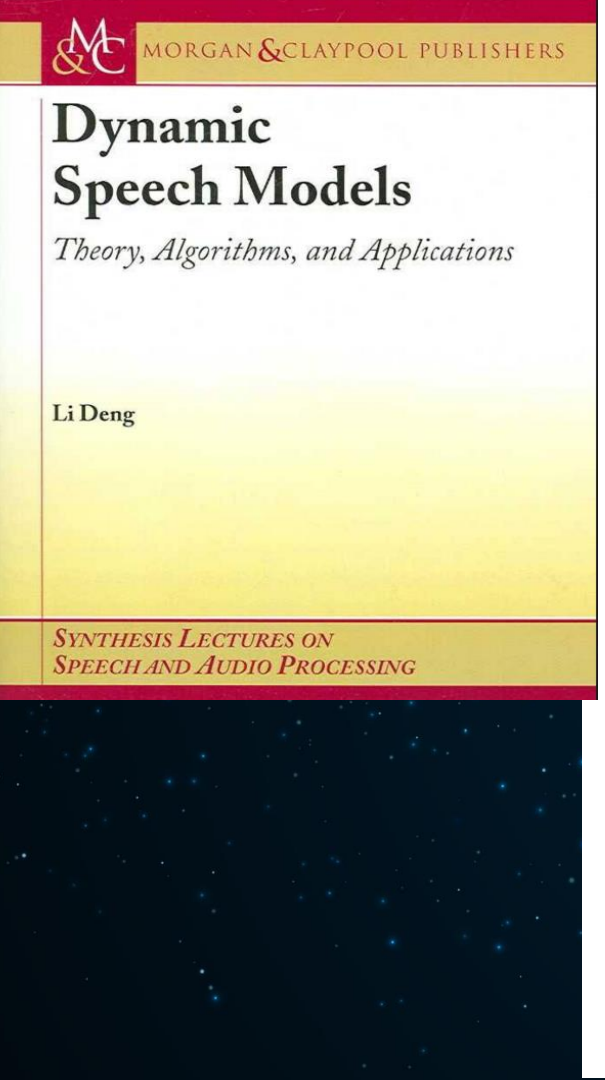
Volume 7, Issue 2, 1994, Pages 331–339

Contributed article

Analysis of the correlation structure for a neural predictive model with application to speech recognition ☆

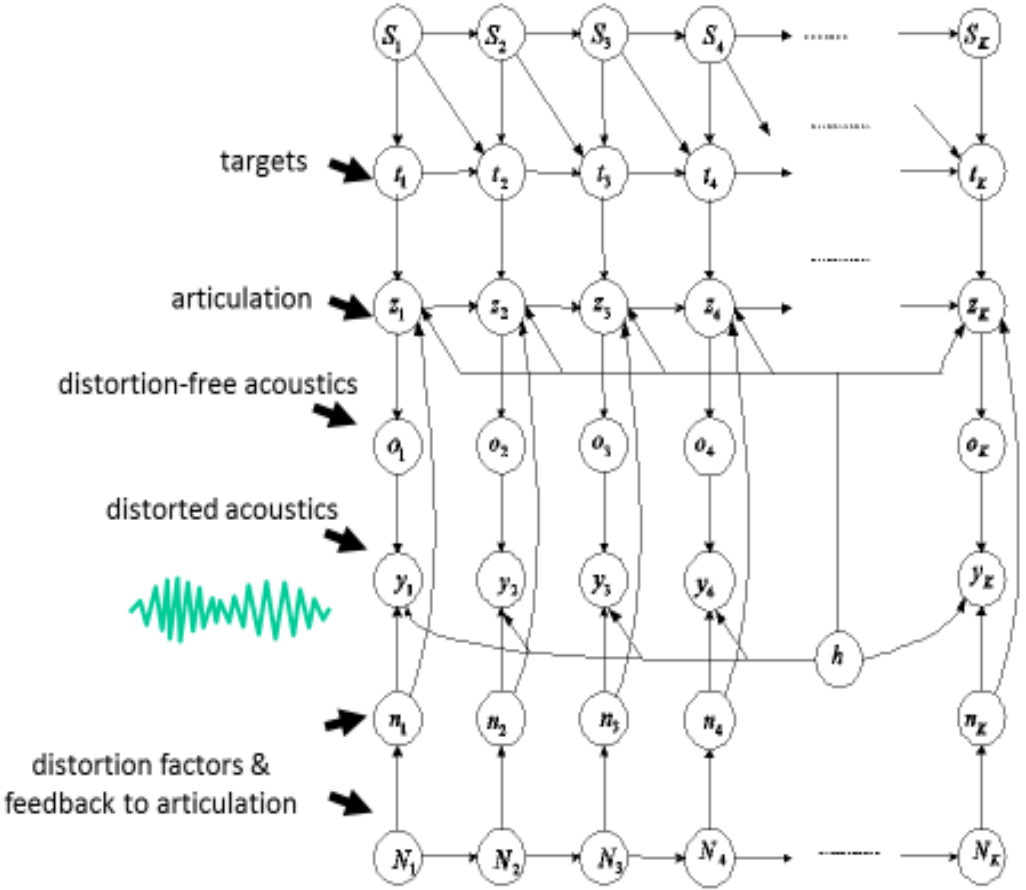
L. Deng, K. Hassanein, M. Elmasry

University of Waterloo Canada



Deep Dynamic Statistical Generative Model

(2006)



LETTER Communicated by Yann Le Cun

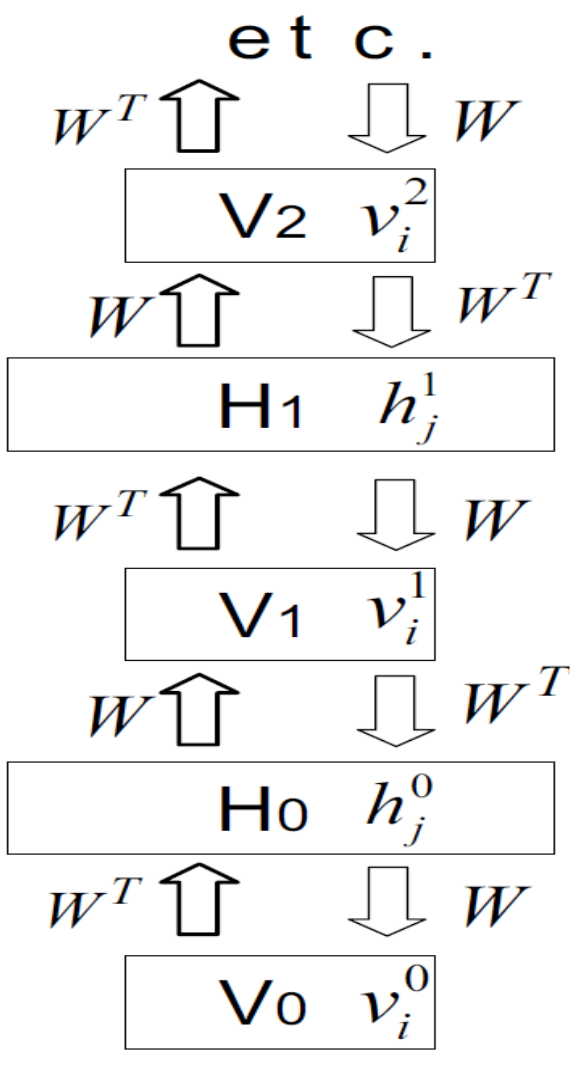
A Fast Learning Algorithm for Deep Belief Nets

(2006)

Geoffrey E. Hinton
hinton@cs.toronto.edu
Simon Osindero
osindero@cs.toronto.edu
Department of Computer Science, University of Toronto, Toronto, Canada M5S 3G4

Yee-Whye Teh
tehyw@comp.nus.edu.sg
Department of Computer Science, National University of Singapore, Singapore 117543

We show how to use "complementary priors" to eliminate the explaining-away effects that make inference difficult in densely connected belief nets that have many hidden layers. Using complementary priors, we derive a fast, greedy algorithm that can learn deep, directed belief networks one layer at a time, provided the top two layers form an undirected associative memory. The fast, greedy algorithm is used to initialize a slower learning procedure that fine-tunes the weights using a contrastive version of the wake-sleep algorithm. After fine-tuning, a network with three hidden layers forms a very good generative model of the joint distribution of handwritten digit images and their labels. This generative model gives better digit classification than the best discriminative learning algorithms. The low-dimensional manifolds on which the digits lie are modeled by long ravines in the free-energy landscape of the top-level associative memory, and it is easy to explore these ravines by using the



Li Deng, Dong Yu, Geoffrey Hinton

Microsoft Research; Microsoft Research; University of Toronto

Deep Learning for Speech Recognition and Related Applications

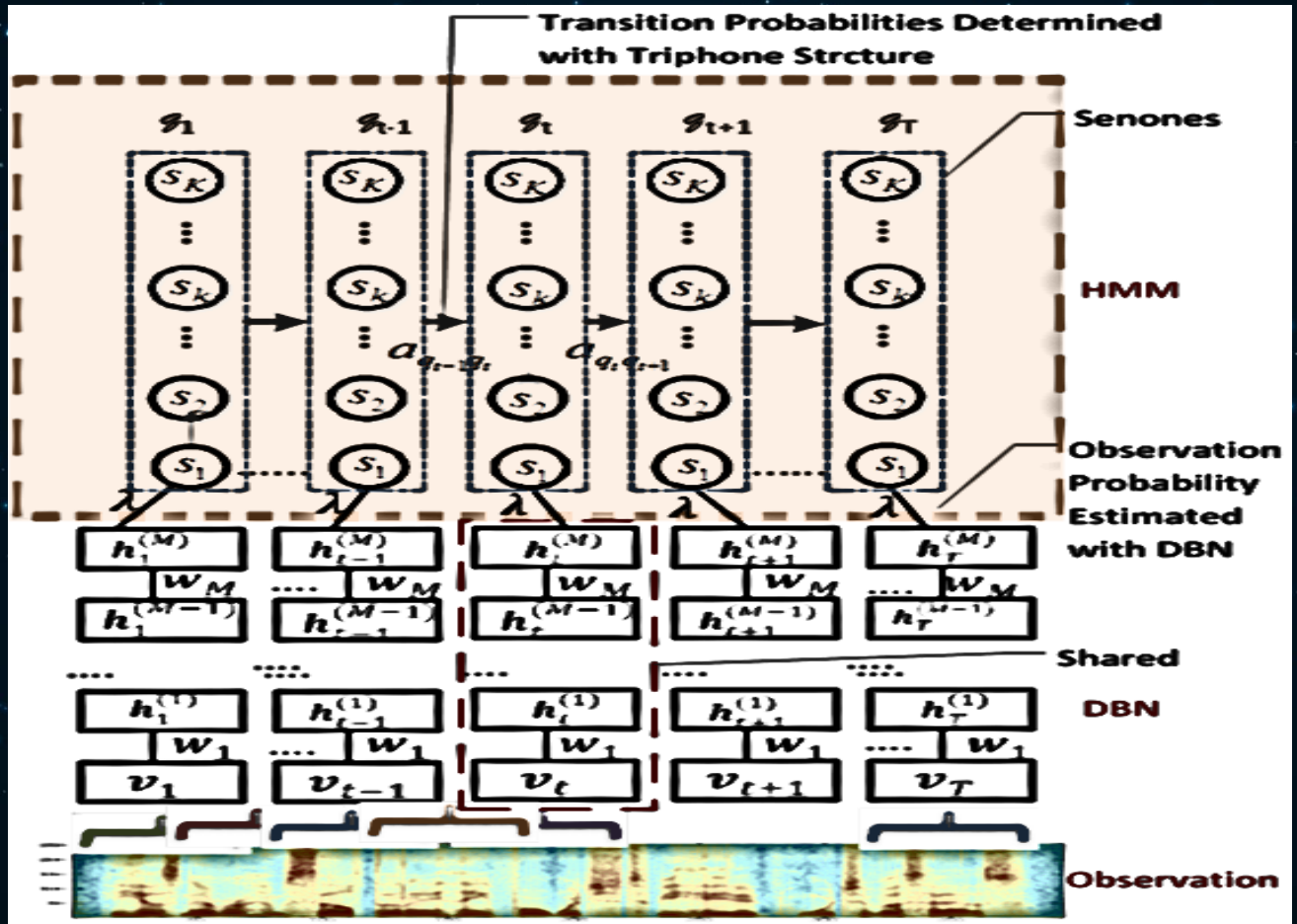
7:30am - 6:30pm Saturday December 12, 2009

Location: Hilton: Cheakamus

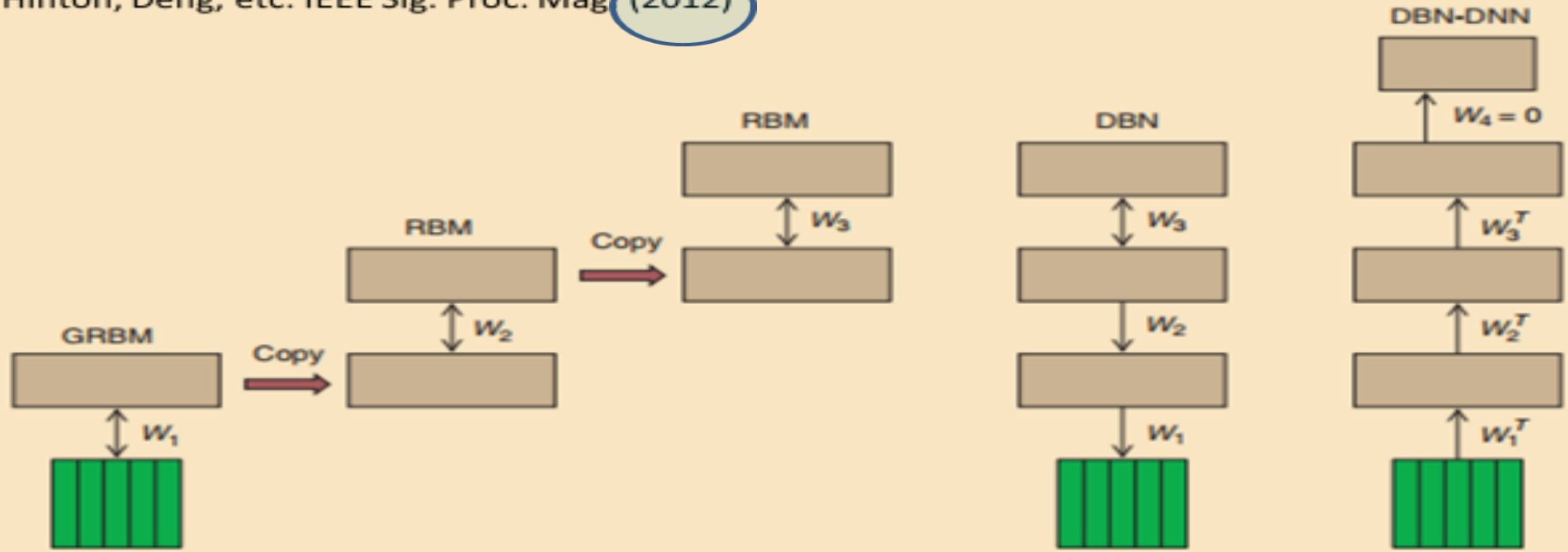
Abstract: Over the past 25 years or so, speech recognition technology has b
dominated by a "shallow" architecture --- hidden Markov models (HMMs). Sig
technological success has been achieved using complex and carefully engin
of HMMs. The next generation of the technology requires solutions to remain
challenges under diversified deployment environments. These challenges, not
addressed in the past, arise from the many types of variability present in the

DNN (Dynamics via HMM)

2010-2011
at Microsoft



Hinton, Deng, etc. IEEE Sig. Proc. Mag (2012)



First train a stack of three models each of which has one hidden layer. Each model in the stack treats the hidden variables of the previous model as data.

Then compose them into a single Deep Belief Network.

Then add outputs and train the DNN with backprop

The Universal Translator ..*comes true!*

The New York Times

Scientists See Promise in Deep-Learning Programs

John Markoff

November 23, 2012

Tianjin, China, October, 25, 2012

Deep learning technology enabled speech-to-speech translation

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

Microsoft Research

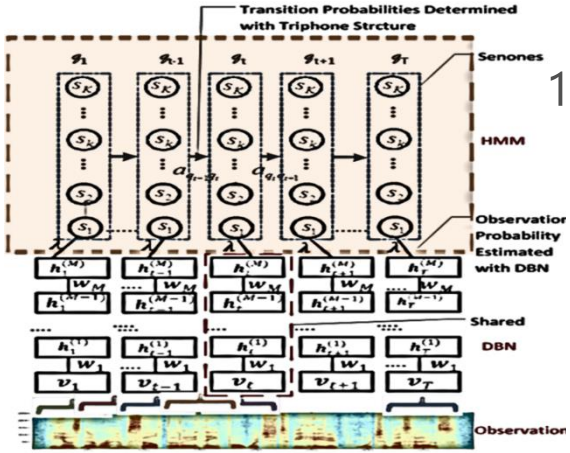
Deep belief networks for phone recognition, NIPS, December 2009; 2012

Investigation of full-sequence training of DBNs for speech recognition., Interspeech, Sept 2010
Binary coding of speech spectrograms using a deep auto-encoder, Interspeech, Sept 2010
Roles of Pre-Training & Fine-Tuning in CD-DBN-HMMs for Real-World ASR, NIPS, Dec. 2010
Large Vocabulary Continuous Speech Recognition With CD-DNN-HMMS, ICASSP, April 2011
Conversational Speech Transcription Using Contxt-Dependent DNN, Interspeech, *Aug. 2011*

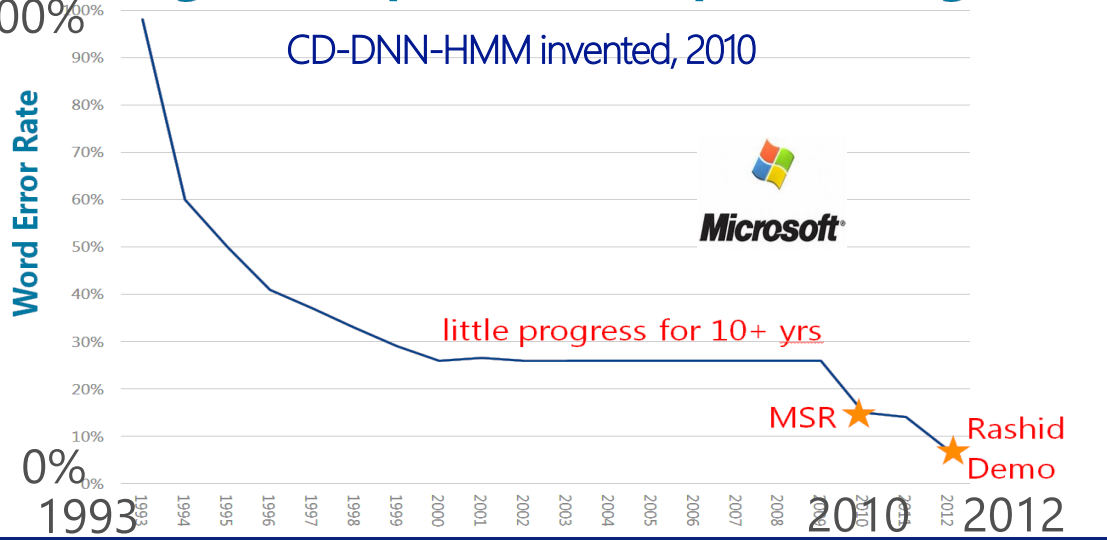
Making deep belief networks effective for LVCSR, ASRU, Dec. 2011
Application of Pretrained DNNs to Large Vocabulary Speech Recognition., ICASSP, 2012

【胡郁】讯飞超脑 2.0 是怎样炼成的？2011, 2015

Later years with rapid progress,

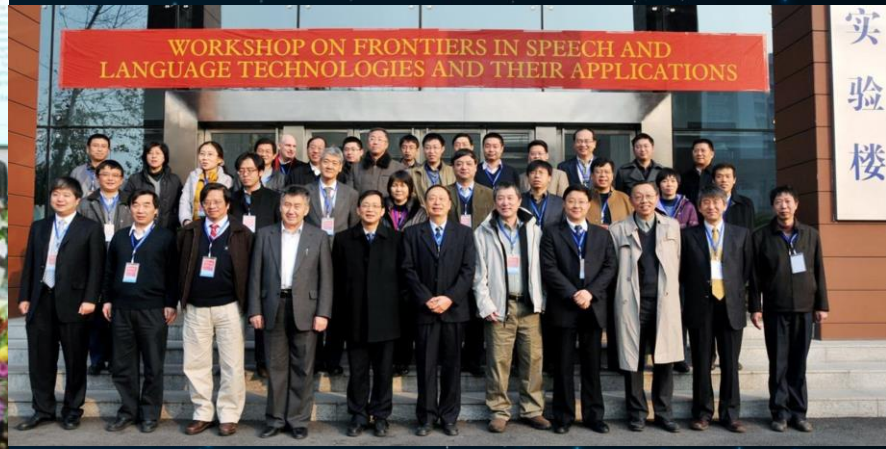


Progress of spontaneous speech recognition



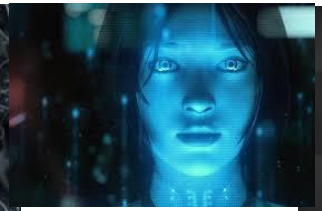
Deep Learning from Canada/USA to China

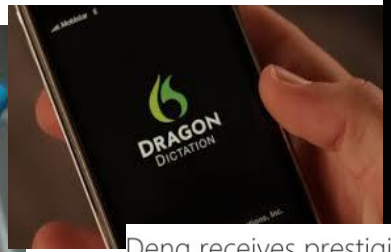
(2010-2012, MSR-Asia, 科大USTC, …)



Across-the-Board Deployment of DNN in Speech Industry

(+ in university labs & DARPA programs) (2011-2014)



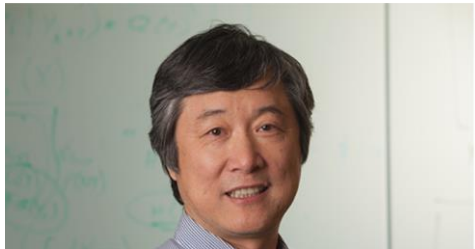


Deng receives prestigious IEEE Technical Achievement Award

December 3, 2015 | Posted by Microsoft Research Blog

Share Tweet Share YouTube

By George Thomas Jr., Writer, Microsoft



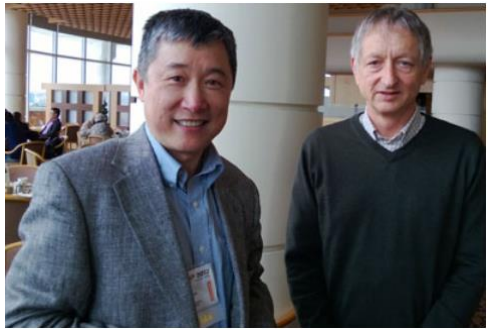
Enabling Cross-Lingual Conversations in Real Time

Microsoft Research
May 27, 2014 5:[illegible] PM PT

View milestones on the path to Skype Translator
#speech2speech

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'

By Jacopo Prisco, for CNN
Updated 12:35 PM ET, Thu October 16, 2014

Microsoft Research

Deep Learning

2014-today

- Speech recognition errors continue to drop rapidly (Microsoft, Google, Baidu, IBM, iFlyTek…)
- Computer Vision: ImageNet errors drop below human level (Microsoft 2015 & Chinese teams 2016)
- Image captioning (看图说话); artistic image generation

Deep Learning

2014-today

- AlphaGo

Deep reinforcement learning integrated with other AI techniques defeats top human GO players (2015-2016)

Deep Learning

2014-today

- AI Bots (智能对话机器人)

  “Many companies are trying to develop more realistic and helpful “chatbots”—automated customer-service representatives.”

Deep Learning

2014-today

- Big data analytics:

  Microsoft “sales teams are using neural nets to recommend which prospects to contact next or what kinds of product offerings to recommend.”

Deep Learning

2014-today

- Enterprise deep learning

“Companies like IBM and Microsoft are also helping business customers adapt deep-learning-powered applications—like speech-recognition interfaces and translation services—for their own businesses, while cloud services like Amazon Web Services provide cheap, GPU-driven deep-learning computation services for those who want to develop their own software.”

Deep Learning

2014-today

- Natural Language Processing

Machine translation, reading comprehension, dialogues, question/answering, email auto-reply …

- New technology

sequence-to-sequence learning, memory networks, attention models, neural Turing machine, differential neural computer (DNC), …

未来展望

What Lies Ahead for Deep Learning & AI Breakthrough

1. Applications
2. Research and Technology

PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE

Executive Office of the President
National Science and Technology Council
Committee on Technology

October 2016

AI holds the potential to be a major driver of economic growth and social progress.

“Deep Learning”

October 2016

“…. The dramatic success of these very large (deep) networks at many machine learning tasks has come as a surprise to some experts, and is the main cause of the current wave of enthusiasm for machine learning among AI researchers and practitioners.”

未来展望

What Lies Ahead for AI Breakthrough

- Medicine and health

"… computers to read X-rays, MRIs, and CT scans more rapidly and accurately than radiologists, to diagnose cancer earlier and less invasively, and to accelerate the search for life-saving pharmaceuticals"

FORTUNE

未来展望

What Lies Ahead for AI Breakthrough

- Robotics, autonomous drones, self-driving cars
- Conversational bots for brand-new mobile UI
- Business operations: inventory management; logistics
- Energy efficiency; environments
- Automated science
- Finance (hedge funds)

未来展望

What Lies Ahead for Deep Learning & AI Breakthrough

1. Applications
2. Research and Technology

Limitations of current deep learning/AI & how to overcome them

AI WORLD 2016 世界人工智能大会

- Blackbox of AI → Need interpretable AI and Deep Networks
- Requiring large labeled data → Need unsupervised learning
- …, …
- Hard to incorporate common-sense knowledge
- Hard to grow knowledge
- "Neural nets are good at recognizing patterns—sometimes as good as or better than we are at it. But they can't reason."

FORTUNE

→Need to unify symbolic logic and neural learning

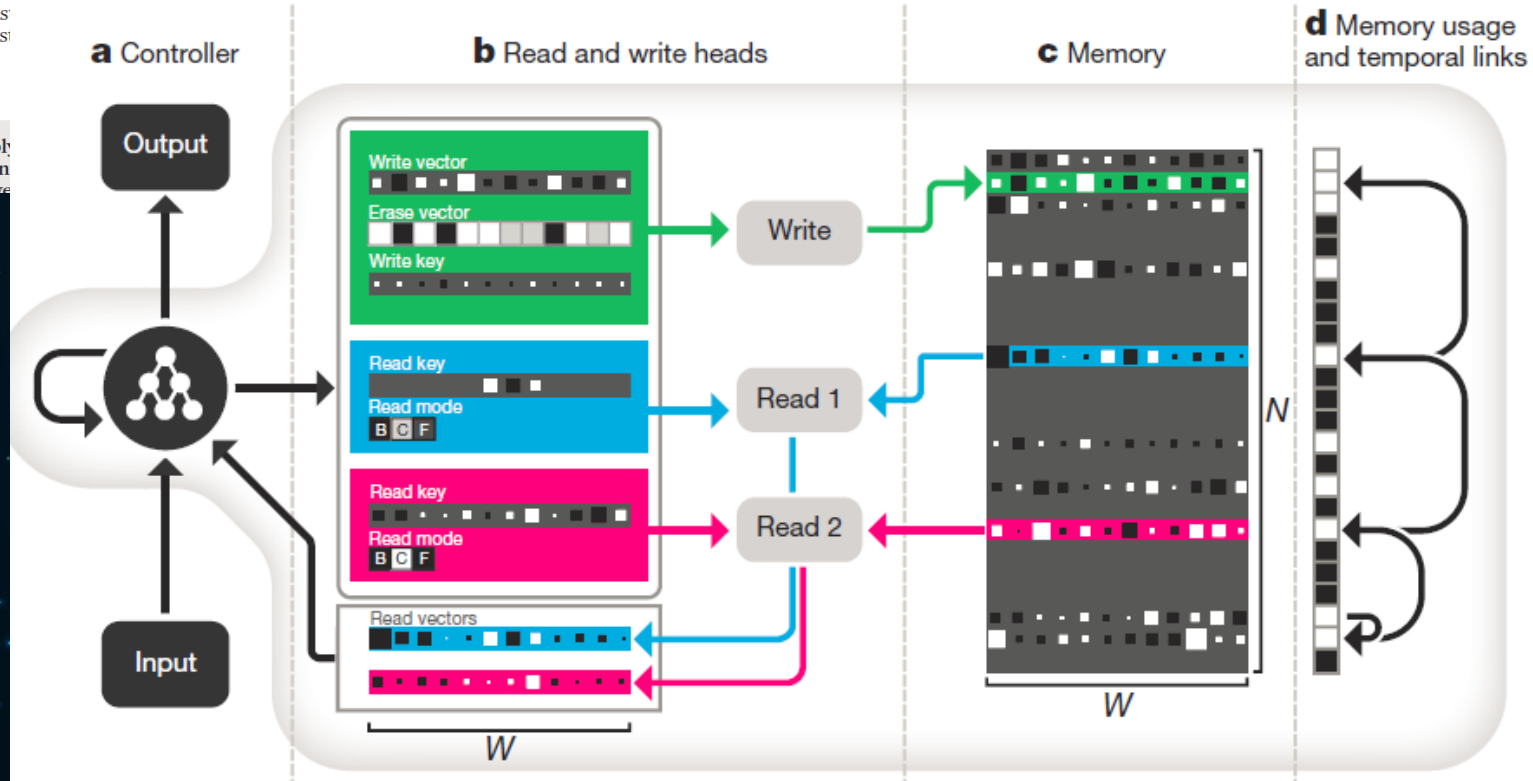
ARTICLE

doi:10.1038/nature20101

Hybrid computing using a neural network with dynamic external memory

Alex Graves[1]*, Greg Wayne[1]*, Malcolm Reyno
Sergio Gómez Colmenarejo[1], Edward Grefens
Karl Moritz Hermann[1], Yori Zwols[1], Georg Os
Koray Kavukcuoglu[1] & Demis Hassabis[1]

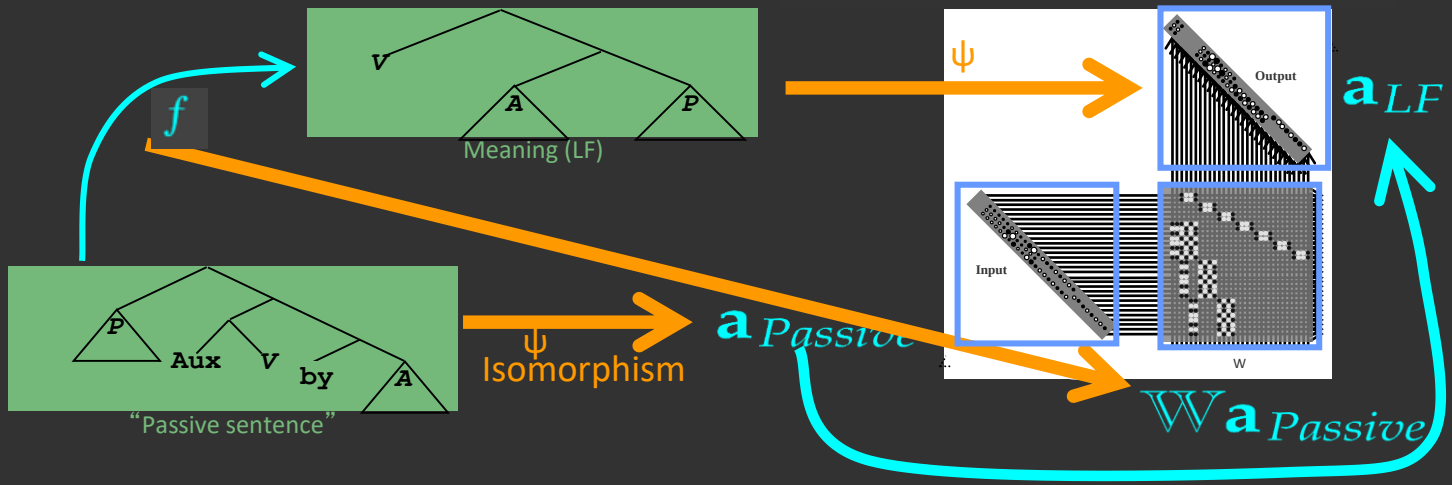
Artificial neural networks are remarkabl
but are limited in their ability to represen
the lack of an external memory. Here we



Example of symbolic tree-to-tree transformation via neural tensor learning
Logic → Tree → Tensor-product

Few leaders are admired by George Bush $\xrightarrow{f}$ admire(George Bush, few leaders)

$f(s) = \mathbf{cons}(\mathbf{ex}_1(\mathbf{ex}_0(\mathbf{ex}_1(s))),$
$\mathbf{cons}(\mathbf{ex}_1(\mathbf{ex}_1(\mathbf{ex}_1(s))), \mathbf{ex}_0(s)))$



Revised from Smolensky/Deng, 2015, at Microsoft Research

Thanks!