Three Generations of Spoken Dialogue Systems (Bots)

Li Deng

Chief Scientist of AI

Microsoft AI & Research Group, Redmond, WA, USA

Invited presentation on January 11, 2017

Microsoft

Conversational UIs, Bots, Dialogue Systems…

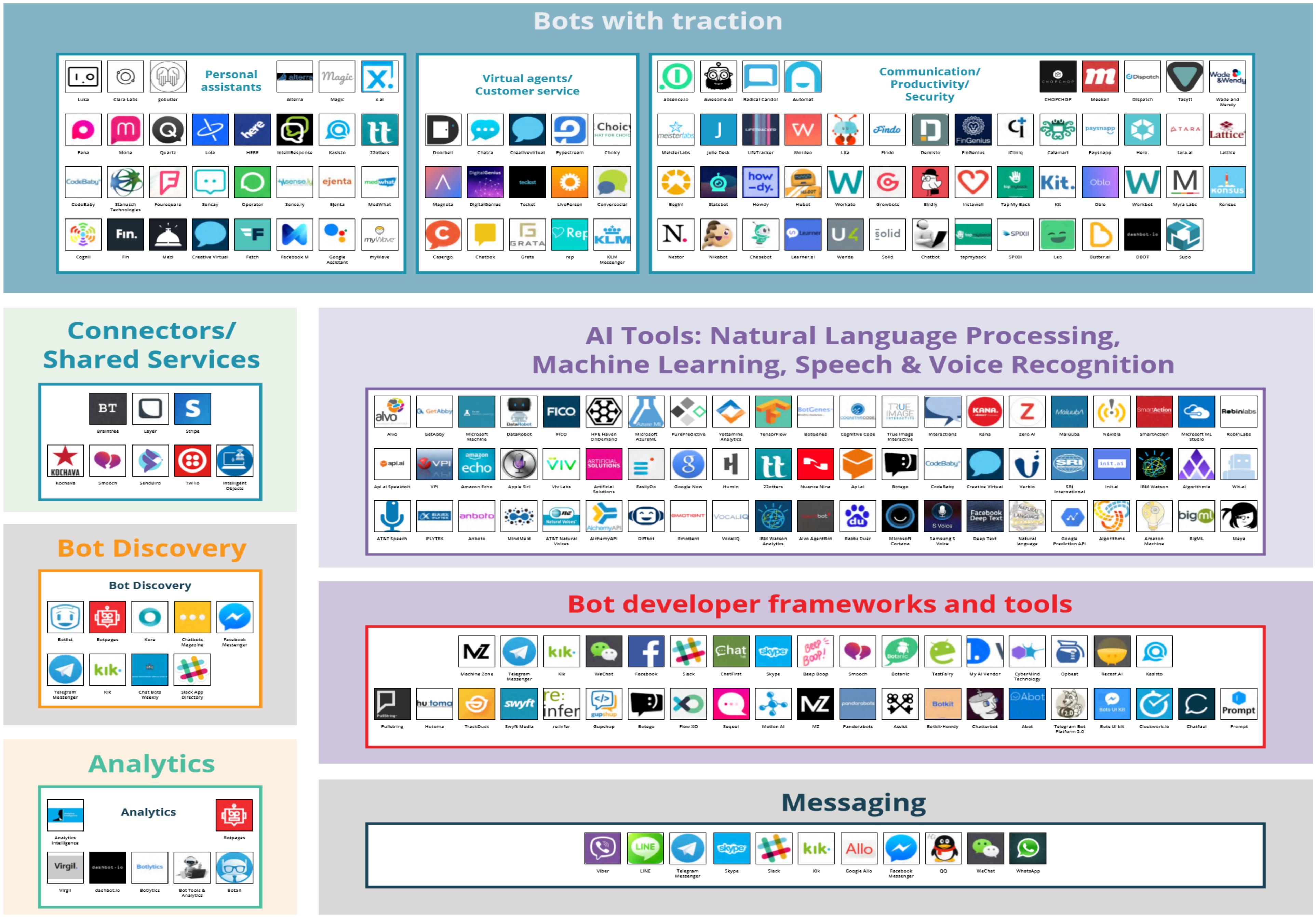
- Speech-based vs text-based
  - Errors in speech recognition treated as "noise" in text as input to text-based bots
  - Spoken dialogue system =|≠ speech recognition + text-based dialogue system
    1) integrated learning vs. simply pipelined
    2) para-linguistic cues in speech signal (prosody, emotion, speaker, etc.)
    3) depending on users, speech input may be simpler or more complex than text input
- Narrow domain vs wide domains
- Speech recognition maturing with deep learning

  → spoken systems no longer limited to narrow domains
- 3 generations of dialogue technology since early 90's

DESIGNED BY
JON CIFUENTES

Bots Landscape

POWERED BY
VB | Profiles



Sense.ly | Ejenta | MedWhat
Facebook M | Google Assistant | myWave

Magneta | DigitalGenius | Teckst | LivePerson | Conversocial
Casengo | Chatbox | Grata | rep | KLM Messenger

Begin! | Statsbot | Howdy | Hubot | Workato | Growbots | Birdly | Instawell | Tap My Back | Kit | Obie | Workbot | Myra Labs | Konsus
Nestor | Nikabot | Chasebot | Learner.ai | Wanda | Solid | Chatbot | tapmyback | SPIXII | Leo | Butter.ai | DBOT | Sudo

AI Tools: Natural Language Processing, Machine Learning, Speech & Voice Recognition

Alvo | GetAbby | Microsoft Machi... | DataRobot | FICO | HPE Haven OnDemand | Microsoft AzureML | PurePredictive | Yottamine Analytics | TensorFlow | BotGenes | Cognitive Code | True Image Interactive | Interactions | Kana | Zero AI | Maluuba | Nexidia | SmartAction | Microsoft ML Studio | RobinLabs

Api.ai Speaktoit | VPI | Amazon Echo | Apple Siri | Viv Labs | Artificial Solutions | EasilyDo | Google Now | Humin | 22otters | Nuance Nina | Api.ai | Botego | CodeBaby | Creative Virtual | Verbio | SRI International | Init.ai | IBM Watson | Algorithmia | Wit.ai

AT&T Speech | IFLYTEK | Anboto | MindMeld | AT&T Natural Voices | AlchemyAPI | Diffbot | Emotient | VocalIQ | IBM Watson Analytics | ... AgentBot | Baidu Due... | Microsoft Cortana | ...msung S Voice | Deep Text | Natural language | Google Prediction API | Algorithms | Amazon Machine | ... | Meya

Bot developer frameworks and tools

Machine Zone | Telegram Messenger | Kik | WeChat | Facebook | Slack | ChatFirst | Skype | Beep Boop | Smooch | Botanic | TestFairy | My AI Vendor | CyberMind Technology | Opbeat | Recast.AI | Kasisto

VB NEWS EVENTS RESEARCH

How deep reinforcement learning can help chatbots

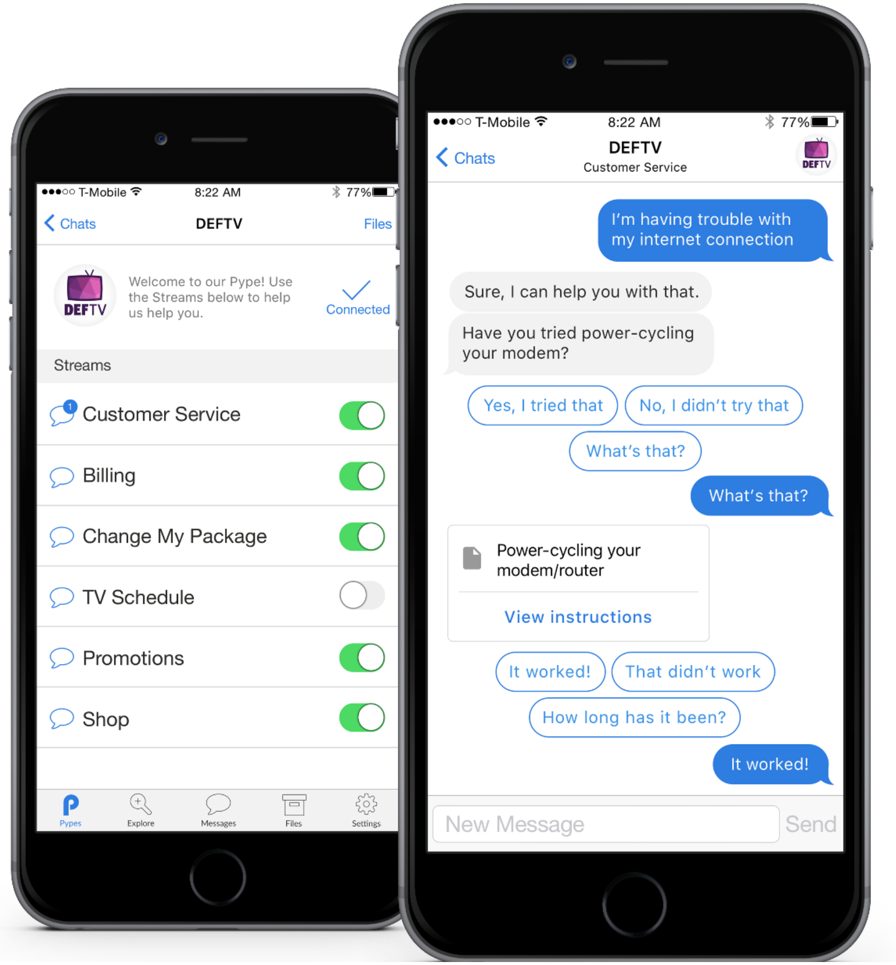
LI DENG, MICROSOFT AUGUST 1, 2016 6:10 PM

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



Major types of conversational bots:

- Social ChatBots (e.g. XiaoIce)
- InfoBots
- TaskCompletion Bots (goal-oriented)
- Personal Assistant Bots (above + recommd.)

Bots Technology Overview: three generations; latest deep RL

http://venturebeat.com/2016/08/01/how-deep-reinforcement-learning-can-help-chatbots/

Generation I: Symbolic Rule/Template Based

- Centered on grammatical rule & ontological design by human experts (early AI approach)
- Easy interpretation, debugging, and system update
- Popular before late 90's
- Still in use in commercial systems and by bots startups
- Limitations:
  - reliance on experts
  - hard to scale over domains
  - data used only to help design rules, not for learning
- Example system next

PUBLISHED IN PROCEEDINGS OF THE IEEE, VOL. 88, NO. 8, 1166-

Zue & Glass (MIT), August 2000

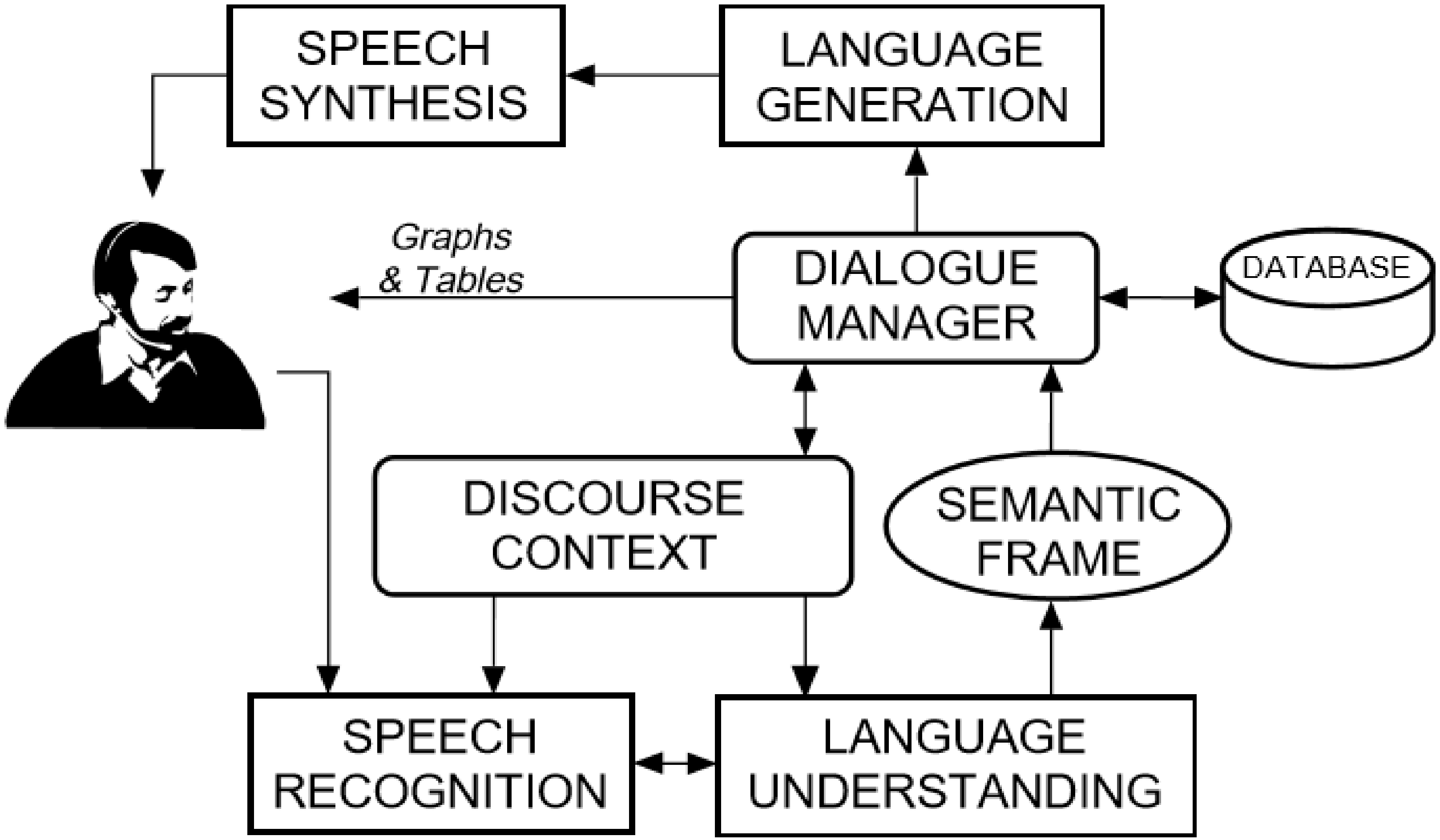


Fig. 3. A generic block diagram for a typical conversational interface.

Domains:

- Air travel info service(ATIS)
- Automobile classifieds
- Berkeley Restaurants
- Automatic railway information systems
- Train schedule planning
- Weather info (JUPITER)
- Flight browsing/booking (MERCURY)
- Navigation assistance & traffic status in Boston (VOYAGER, VOYAGER-II)

Generation II: Data Driven, (shallow) Learning

- Data used not to design rules for NLU and action, but to learn statistical parameters in dialogue systems
  - Reduce cost of hand-crafting complex dialogue manager
  - Robustness against speech recog errors in noisy environment
- MDP/POMDP & reinforcement learning for dialogue policy
- Discriminative (CRF) & generative (HMM) methods for NLU
- Popular in academic research until 2014 (before deep learning arrived at the dialogue world); in parallel with Generation I (BBN, AT&T, CMU, SRI, CU …)
- Limitations:
  - Not easy to interpret, debug, and update systems
  - Still hard to scale over domains
  - Models & representations not powerful enough; no end-2-end, hard to scale up
  - Remained academic until deep learning arrived
- Example system next

PROC IEEE, VOL. 101, NO. 5, 1160-1179, 2013

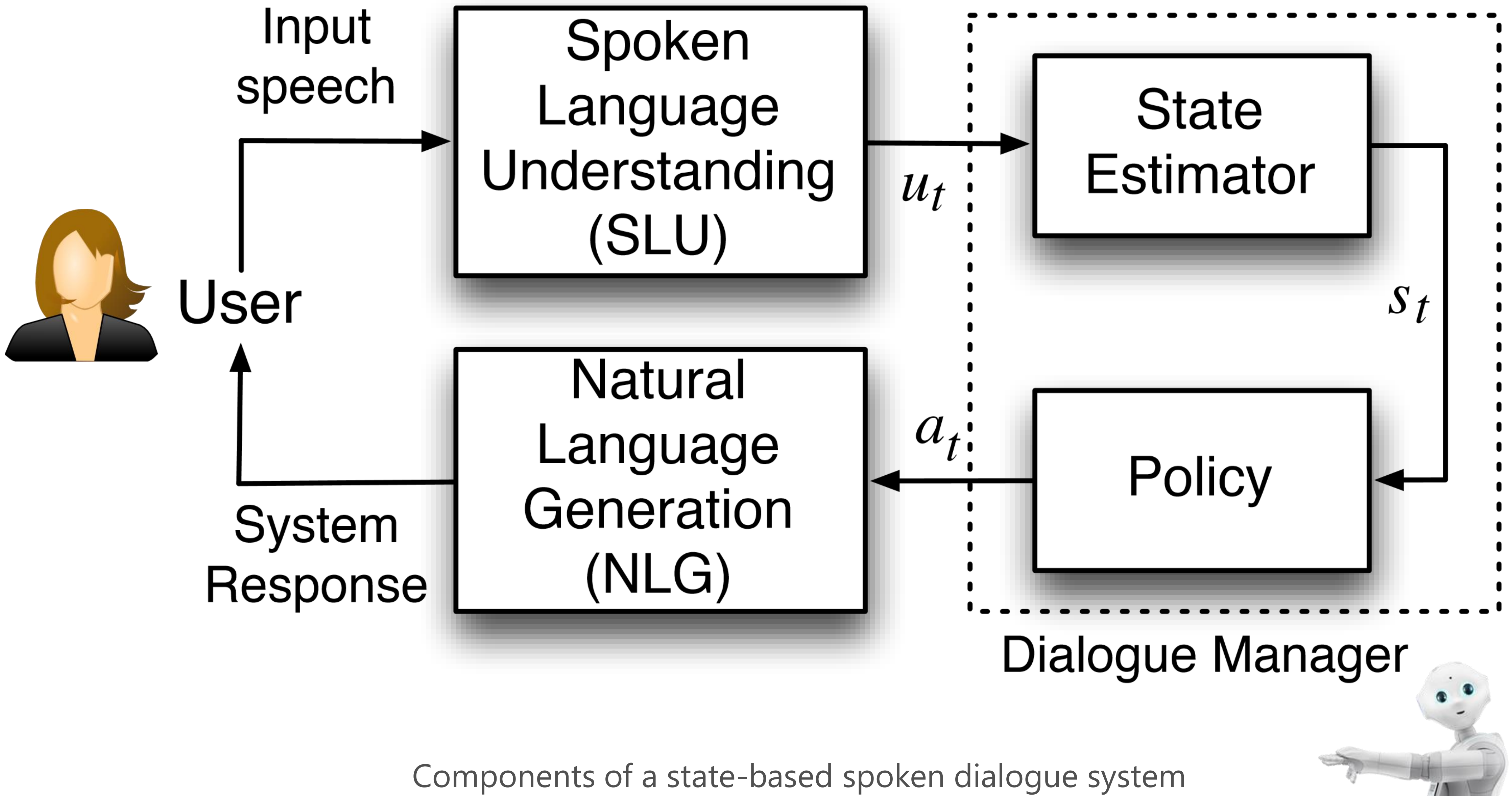
POMDP-based Statistical Spoken Dialogue Systems: a Review

Steve Young, *Fellow, IEEE,* Milica Gašić, *Member, IEEE,* Blaise Thomson, *Member, IEEE,* and Jason D Williams, *Member, IEEE*

(Invited Paper)

Abstract—**Statistical dialogue systems are motivated by the need for a data-driven framework that reduces the cost of laboriously hand-crafting complex dialogue managers and that provides robustness against the errors created by speech recognisers operating in noisy environments. By including an explicit Bayesian model of uncertainty and by optimising the policy via**

the dialogue and the attribute values (often called slots) determine the user's requirements. In conventional syst the policy is usually defined by a flow chart with n representing states and actions, and arcs representing inputs[5], [6].



Components of a state-based spoken dialogue system

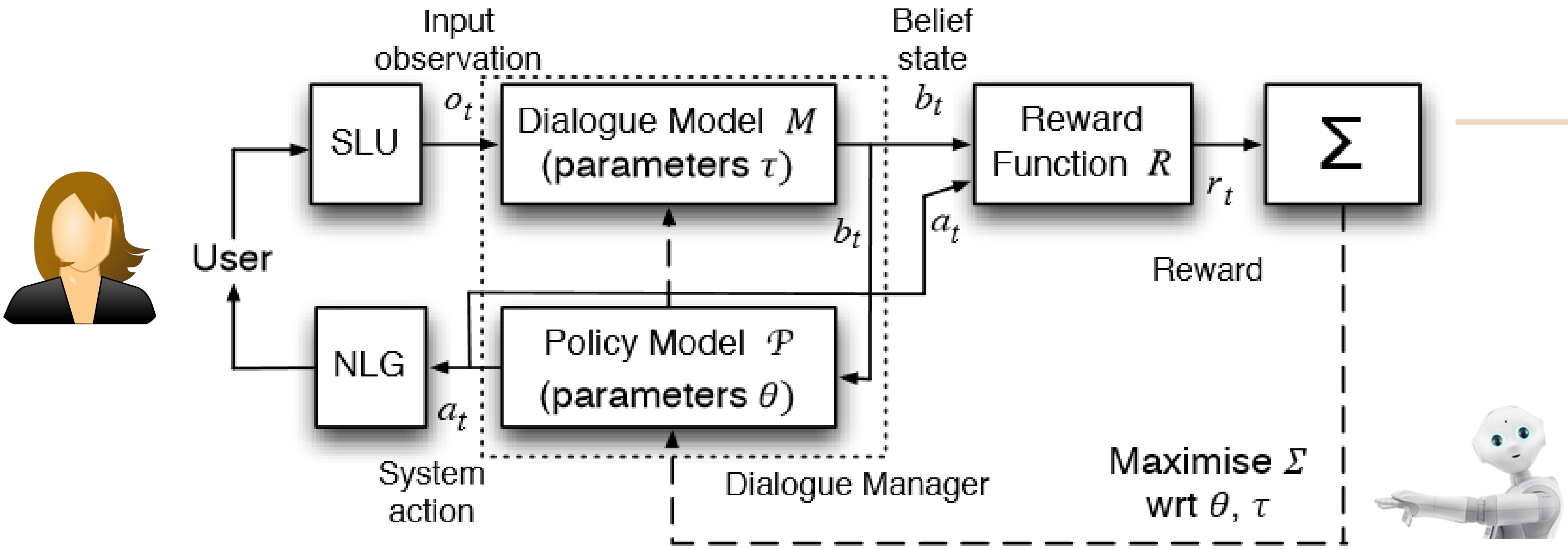


Fig. 2. Components of a POMDP-based spoken dialogue system. In contrast to Fig. 1, the decoded input speech is now regarded as a noisy observation o_t of the underlying user intent u_t. Since u_t is hidden, the system maintains a distribution b_t over all possible dialogue states and instead of trying to estimate the hidden dialogue state, the system response is determined directly from b_t. In addition, the dialogue model and policy are parameterised, and given an appropriate reward function, they can be optimised using reinforcement learning.

Generation III: Data-Driven Deep Learning

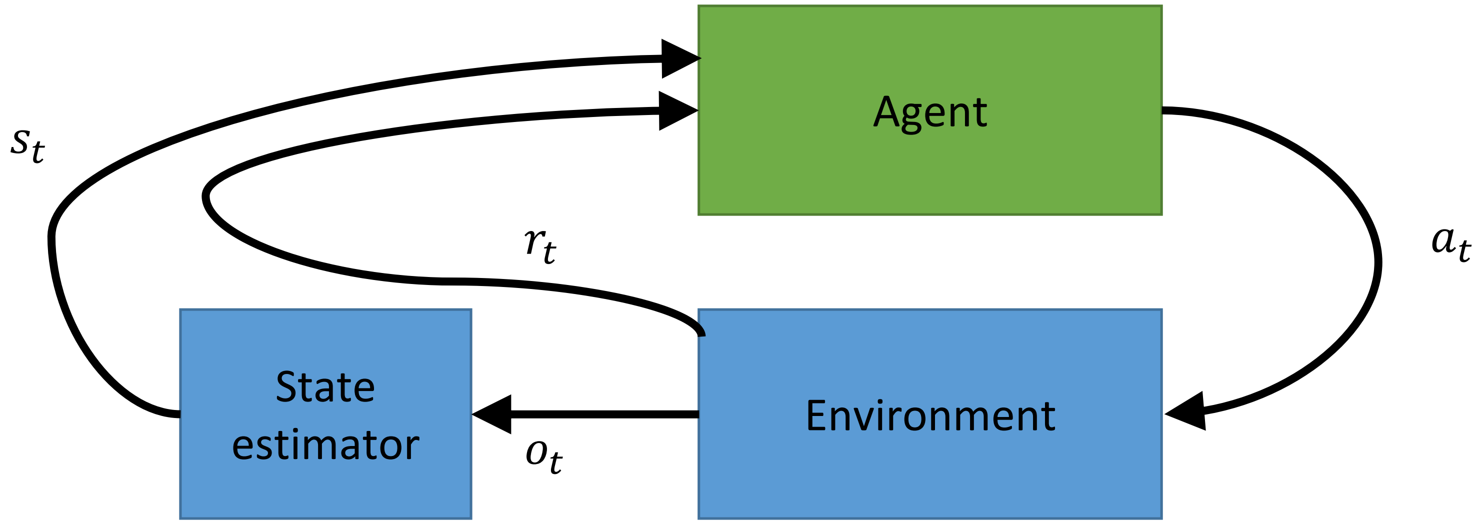
- Like Generation-II, data used to learn everything in dialogue systems
  - Reduce cost of hand-crafting complex dialogue manager
  - Robustness against speech recog errors in noisy environment & against NLU errors
  - MDP/POMDP & reinforcement learning for dialogue policy (same)
- But, neural models & representations are much more powerful
- End-to-End learning becomes feasible
- Attracted huge research efforts since 2015 (after deep learning's success in vision/speech and in deep RL shown success in Atari games)
- Limitations:
  - Still not easy to interpret, debug, and update systems
  - Lack interface btw cont. neural learning and symbolic NL structure to human users
  - Active research in scaling over domains via deep transfer learning & RL
  - No clear commercial success reported yet
- Deep RL & example research next

What is reinforcement learning (RL)?

- RL in Generation-II ---> not working! (with unnecessarily complex POMDP)
- RL in Generation-III ---> working! due to deep learning -- like NN vs DNN in ASR)

- RL is learning what to do so as to maximize a numerical reward signal
  - "What to do" means mapping from a situation in a given environment to an action
  - Takes inspiration from biology / psychology
- RL is a characterization of a problem class
  - Doesn't refer to a specific solution
  - There are many methods for solving RL problems
- In its most general form, RL problems:
  - <u>Have a stateful environment</u>, where actions can change the state of the environment
  - <u>Learn by trial and error</u>, *not* by being shown examples of the right action
  - <u>Have delayed rewards</u>, where an action's value may not be clear until some time after it is taken

Stateful Model for RL



$s_t = \mathrm{Summary}(o_0, a_0, o_1, a_1 \cdots, o_{t-1}, a_{t-1}, o_t)$

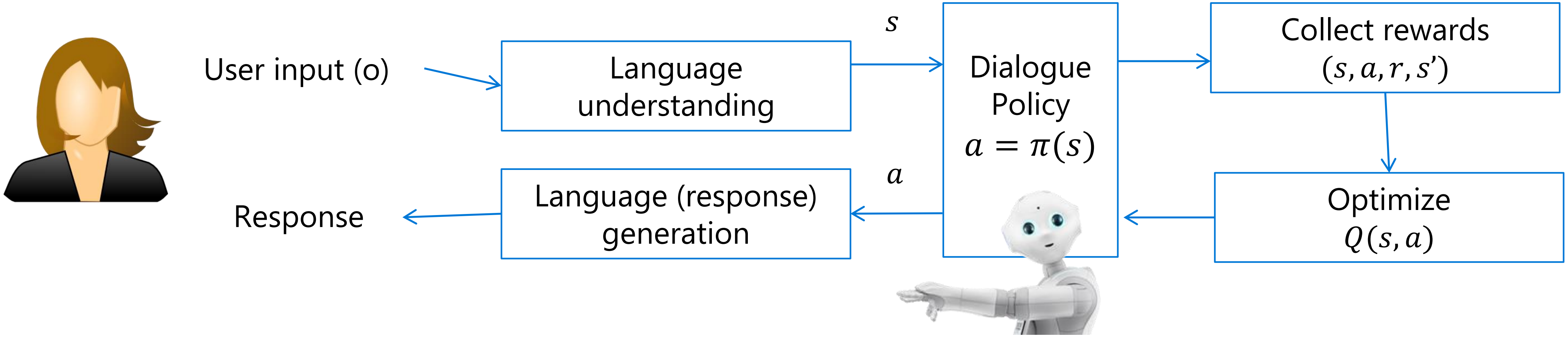
Trajectory: $a_0, r_1, s_1, a_1, r_2, s_2, a_2, \cdots$

Return: $\sum_{\tau=t+1}^{\infty} \gamma^{\tau-1} r_\tau$, $1 \geq \gamma \geq 0$

Policy: $\pi(s_t) \rightarrow a_t$

Objective: $\pi^* = \arg\max_{\pi} E[\sum_{\tau=t+1}^{\infty} \gamma^{\tau-1} r_\tau \mid \pi], \forall s_t$

Deep Reinforcement Learning to Optimize Dialogue Policy



Type pf Bots	State	Action	Reward
Social ChatBots	Chat history	System Response	# of turns maximized; Intrinsically motivated reward
InfoBots (interactive Q/A)	User current question + Context/history	Answers to current question by system	Relevance of answer; # of turns minimized
Task-Oriented Bots	User current input + Context/history	DialogAct w. SlotValue in current turn	Task success rate; # of turns minimized

Q-Learning [Sutton & Barto 98]

- Assume $Q(s, a)$ for all (s, a) can be represented in a table

1. Initialize an array $Q(s, a)$ arbitrarily
2. Choose actions based on Q such that all actions are taken in all states (infinitely often in the limit)
3. On each time step, update one element of the array:

$$\Delta Q(s_t, a_t) = \alpha r_{t+1} + \gamma(\max_{a'} Q(s_{t+1}, a'))$$

- Model-free learning:
  - Learning long-term optimal behavior without model of the environment

Function Approximation

- In many tasks, (s, a) is too large for tabular representation
- Estimate the action-value function approximately as $Q(s,a;\theta) \approx Q^*(s,a)$
- θ: a linear function (baseline)
- θ: a DNN, aka Deep Q-Network (DQN)
  - Optimize θ using SGD w.r.t loss

$$L_i(\theta_i) = \mathbb{E}_{s,a\sim\rho(\cdot)}\left[(y_i - Q(s,a;\theta_i))^2\right]$$

$$y_i = \mathbb{E}_{s'\sim\mathcal{E}}\left[r + \gamma \max_{a'} Q(s',a';\theta_{i-1}) | s, a\right]$$

$$\nabla_{\theta_i} L_i(\theta_i) = \mathbb{E}_{s,a\sim\rho(\cdot);s'\sim\mathcal{E}}\left[\left(r + \gamma \max_{a'} Q(s',a';\theta_{i-1}) - Q(s,a;\theta_i)\right)\nabla_{\theta_i} Q(s,a;\theta_i)\right]$$

Q-Learning for DQN [DeepMind 15]

- Learning becomes unstable
  - Correlations present in the sequence of states
  - Small updates to Q leads to significant change of policy and data distribution
  - Correlations btw the to-be-learned Q and the target value $r + \max_{a'} Q(s, a')$
- Solution
  - Experience replay: randomize training samples (s, a, r, s')
  - Use a separate Q function to compute targets y

User goal – ***Two tickets for "Deadpool" tomorrow 9PM at AMC Pacific Place 11 theater, Seattle.***

RULE BASED SYSTEM

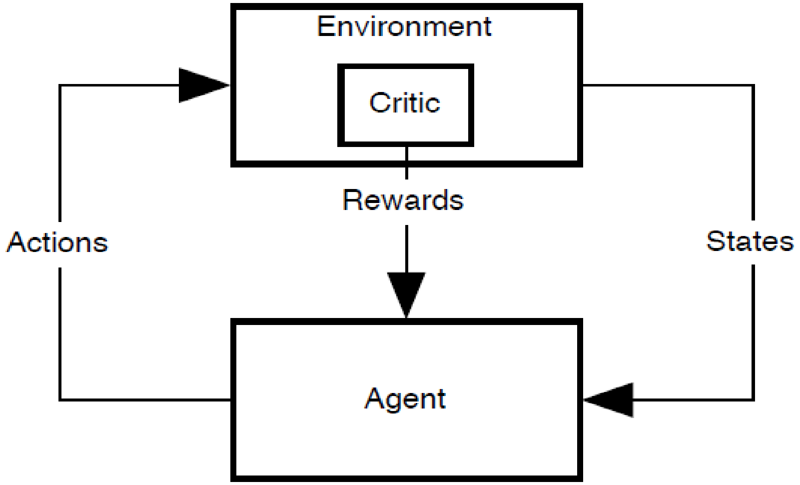
User: Do you have any ticket for movie 'Deadpool' at 9:00 pm?
Agent: Which theater do you want?
User: AMC pacific place 11 theater, please!
Agent: How many tickets do you want?
User: Two!
Agent: What date?
User: Tomorrow!
Agent: Which city do you want?
User: Seattle!
Agent: Do you have any specification for zip code?
User: No, I do not care.
Agent: Do you have any distance constraint?
User: No, I do not care.
Agent: How about video format? 2d? Imax 3d?
User: Any kind is okay.
Agent: Do you have any constraint to price?
User: No!
Agent: Okay, I will issue 2 tickets for you, tomorrow 9:00 pm at AMC pacific place 11 theater, Seattle,
movie 'Deadpool'.
User: Okay, thanks.
Agent: Thanks!

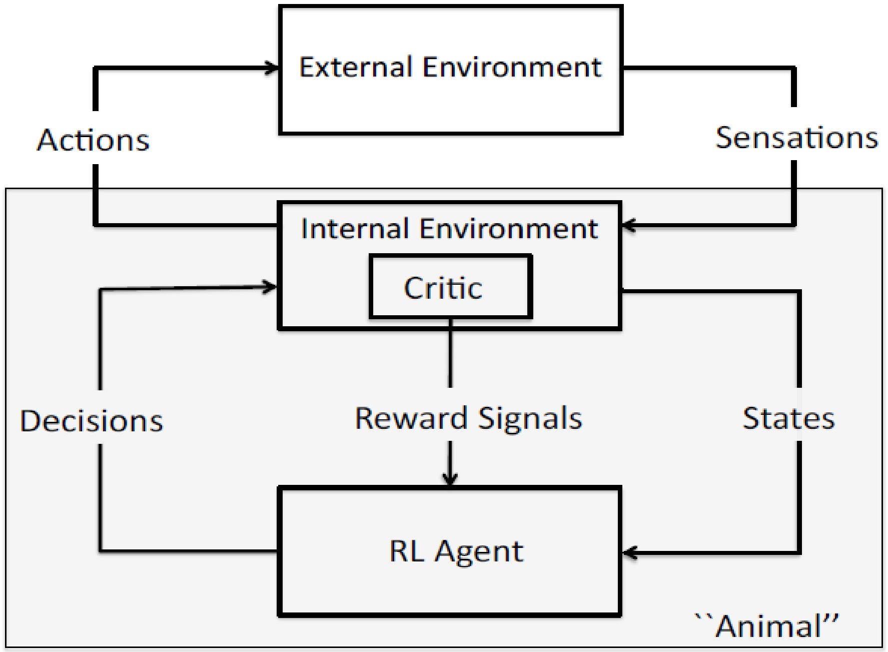
Wasteful!
(Very likely the user does not care.)

REINFORCEMENT LEARNING SYSTEM

User: I want two tickets for movie 'Deadpool'.
Agent: What is the starttime you prefer?
User: 9:00 PM
Agent: Which day?
User: Tomorrow
Agent: Which city?
User: Seattle
Agent: Which theater do you prefer?
User: AMC Pacific Place 11 theater.
Agent: Okay, I will issue 2 tickets for you, tomorrow 9:00 pm at AMC pacific place 11 theater, Seattle,
movie 'Deadpool'.
User: Okay, thanks.
Agent: Thanks!

Extrinsically / Intrinsically motivated RL





- **Extrinsic motivation** – do sth for specific reward
- RL is driven by task-specific reward
- Learning task-specific skills, don't cope flexibly with new problems

- **Intrinsic motivation** – do sth inherently enjoyable
- RL is driven by curiosity w/o explicit reward
- Developing board competence, which makes learning task-specific skill more easily

Task Completion Bots **InfoBots** **Social Bots**

References on deep-learning dialogue systems (Generation-Ⅲ technology)

- Li Deng & Yang Liu (Edited Book) 2017. **Deep Learning in Natural Language Processing, Springer**, Aug 2017 (scheduled)
- Pararth Shah, Dilek Hakkani-Tür, Larry Heck. 2017. Interactivereinforcementlearningfortask-oriented dialoguemanagement. arXiv.
- Dilek Hakkani-Tur, Gokhan Tur, Asli Celikyilmaz, YunNung Chen, Jianfeng Gao, Li Deng, and Ye-Yi Wang. 2016. Multi-domain joint semantic frame parsing using bi-directional RNN-LSTM. INTERSPEECH.
- Antoine Bordes and Jason Weston. 2016. Learning end-to-end goal-oriented dialog. arXiv.
- Mehdi Fatemi, Layla El Asri, Hannes Schulz, Jing He, Kaheer Suleman. 2016. Policy Networks with Two-Stage Training for Dialogue Systems. SIGDIAL.
- Layla El Asri, Jing He, Kaheer Suleman. 2016. A Sequence-to-Sequence Model for User Simulation in Spoken Dialogue Systems. INTERSPEECH.
- Yun-Nung Chen, Dilek Hakkani-Tur, Gokhan Tur, Jianfeng Gao, and Li Deng. 2016. End-to-end memory networks with knowledge carryover for multi-turn spoken language understanding. INTERSPEECH.
- **Bhuwan Dhingra, Lihong Li, Xiujun Li, Jianfeng Gao Yun-Nung Chen, Faisal Ahmed, Li Deng. 2016. End-to-end reinforcement learning of dialogue agents for information access. to submit to ACL.**
- Xuesong Yang, Yun-Nung Chen, Dilek Hakkani-Tur, Paul Crook, Xiujun Li, Jianfeng Gao, Li Deng. 2016. End-to-end joint learning of natural language understanding and dialogue manager. *arXiv*
- Zachary C. Lipton, Jianfeng Gao, Lihong Li, Xiujun Li, Faisal Ahmed, Li Deng. 2016. Efficient Exploration for Dialogue Policy Learning with BBQ Networks & Replay Buffer Spiking. ArXiv.
- Jason D Williams and Geoffrey Zweig. 2016. End to-end LSTM-based dialog control optimized with supervised and reinforcement learning. arXiv
- Tiancheng Zhao and Maxine Eskenazi. 2016. Towards end-to-end learning for dialog state tracking and management using deep reinforcement learning. arXiv preprint arXiv
- Pei-Hao Su, MilicaGasic, Nikola Mrksic, Lina Rojas-Barahona, Stefan Ultes, David Vandyke, Tsung-Hsien Wen and Steve Young. 2016. On-line active reward learning for policy optimisation in spoken dialogue systems. ACL.
- Tsung-Hsien Wen, Milica Gasic, Nikola Mrksic,Pei-Hao Su, David Vandyke, and Steve Young. 2015. Semantically conditioned LSTM-based natural language generation for spoken dialogue systems. EMNLP.
- Gregoire Mesnil, Yann Dauphin, Kaisheng Yao, Yoshua Bengio, Li Deng, Dilek Hakkani-Tur, Xiaodong He, Larry Heck, Gokhan Tur, Dong Yu, and Geoffrey Zweig. 2015. “Using recurrent neural networks for slot filling in spoken language understanding,” IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 23, no. 3, pp. 530–539.

Research Frontiers

- Speech-based vs text-based
  - Errors in speech recog NOT just as "noise" in text as input to text-based bots
  - Solving robustness: huge opportunity for integrated end-to-end design

- Deep reinforcement learning for dialogues
  - How to scale up for large data
  - How to automatically acquire reward signals via "self play" (like AlphaGo)
  - High-quality user simulator for *exploration* phase of RL & system evaluation
- Symbolic-Neural Integration
  - Adding interpretability and reasoning to conversational interface

Neural Information Processing Systems Foundation

NIPS : Conferences : 2009 : Program

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 been dominated by a "shallow" architecture --- hidden Markov models (HMMs). Significant technological success has been achieved using complex and carefully engineered variants of HMMs. The next generation of the technology requires solutions to remaining technical challenges under diversified deployment environments. These challenges, not adequately addressed in the past, arise from the many types of variability present in the speech generation process. Overcoming these challenges is likely to require "deep" architectures with efficient learning algorithms. For speech recognition and related sequential pattern recognition applications, some attempts have been made in the past to develop computational architectures that are "deeper" than conventional HMMs, such as

nature
International weekly journal of science

Deep learning

Yann LeCun, Yoshua Bengio & Geoffrey Hinton

Affiliations | Corresponding author

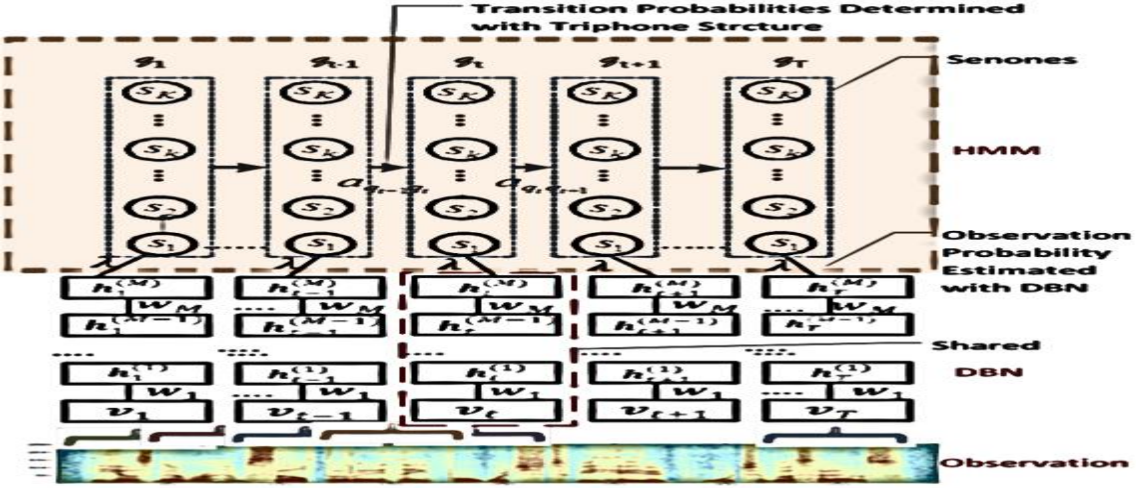
Nature **521**, 436–444 (28 May 2015) | doi:10.1038/nature14539
Received 25 February 2015 | Accepted 01 May 2015 | Published online 27

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

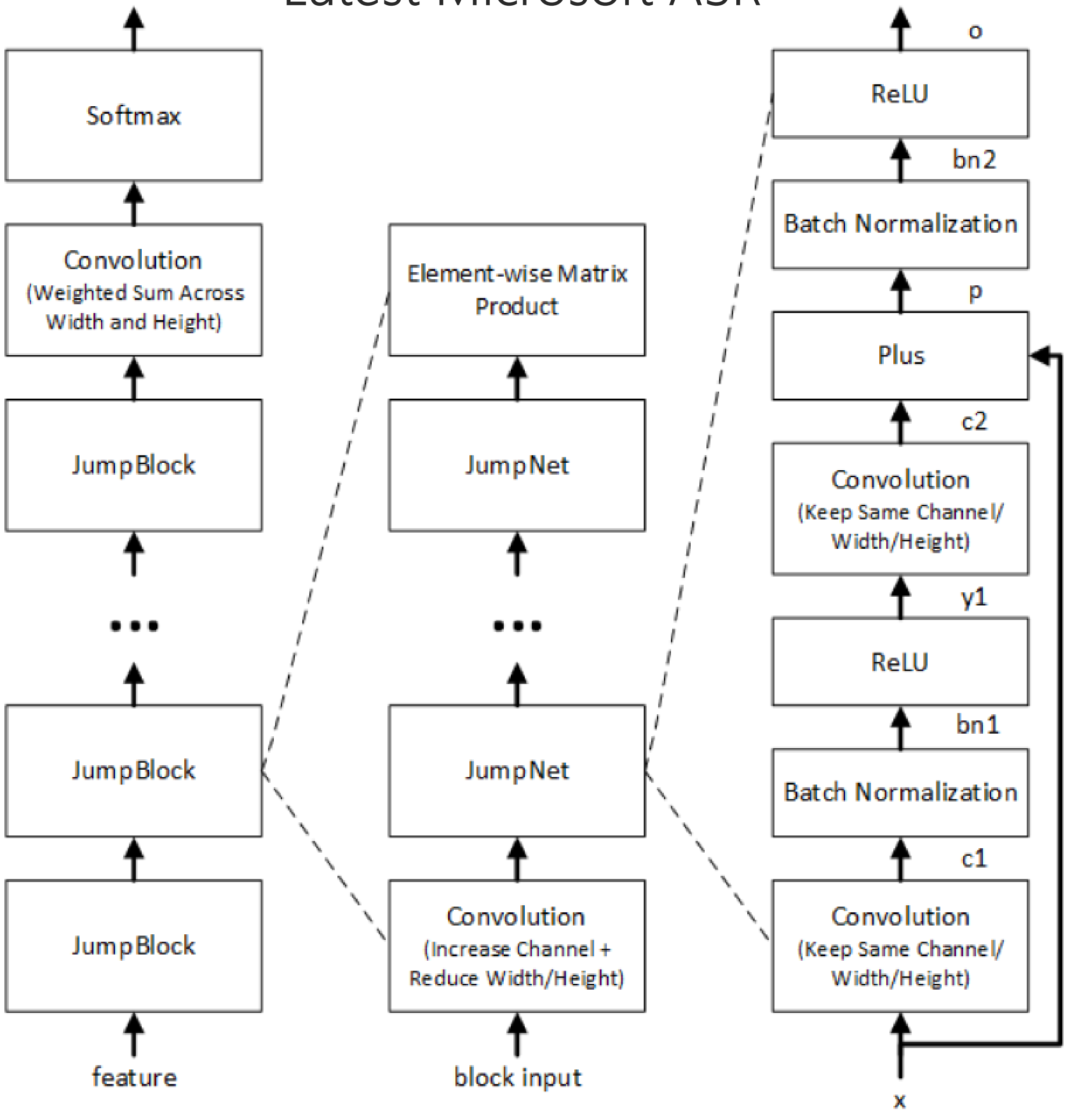
[Geoffrey Hinton, Li Deng, Dong Yu, George E. Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N. Sainath, and Brian Kingsbury]

Deep Neural Networks for Acoustic Modeling in Speech Recognition

[The shared views of four research groups]



Latest Microsoft ASR



(a) (b) (c)

Achieving Human Parity in Conversational Speech Recognition Xiong, et al., MSR-TR-2016-71, 2016

- From DNN□HMM hybrid to (CNN + LSTM)□HMM hybrid
- CNN with attentional layer-wise context expansion (LACE)
- Resnet-like jump links and VGG version of CNN
- Lattice-free MMI training
- Use i-vectors for speaker adaptation
- Bi-directional RNN LM rescoring
- A spatial smoothing method motivated by neuroscience
- Use of letter trigrams (DSSM coding scheme) as LM inputs
- Use of CNTK (Microsoft Cognitive Toolkit)
- Ensemble learning

- Lowest ASR error rate on SWBD: 5.9%
- On par with human SR error rate on the same: 5.9%

- But, learning is based on minimizing word error rate, not dialogue' systems' user satisfaction

5 areas of potential ASR breakthrough

1. better modeling for end-to-end and other specialized architectures capable of disentangling mixed acoustic variability factors (e.g. sequential GAN)
2. better integrated signal processing and neural learning to combat difficult far-field acoustic environments especially with mixed speakers

3. **use of neural language understanding to model long-span dependency for semantic and syntactic consistency in speech recognition outputs, use of semantic understanding in spoken dialogue systems to provide feedbacks to make acoustic speech recognition easier**
4. use of naturally available **multimodal** "labels" such as images, printed text, and handwriting to supplement the current way of providing text labels to synchronize with the corresponding acoustic utterances (NIPS Multimodality Workshop)
5. development of ground-breaking deep **unsupervised learning** methods for exploitation of potentially unlimited amounts of naturally found acoustic data of speech without the otherwise prohibitively high cost of labeling based on the current deep supervised learning paradigm

INVITED PAPER

Speech-Centric Information Processing: An Optimization-Oriented Approach

The authors present a statistical framework for the end-to-end system design where the interactions between automatic speech recognition and downstream text-based processing tasks are fully incorporated and design consistency established.

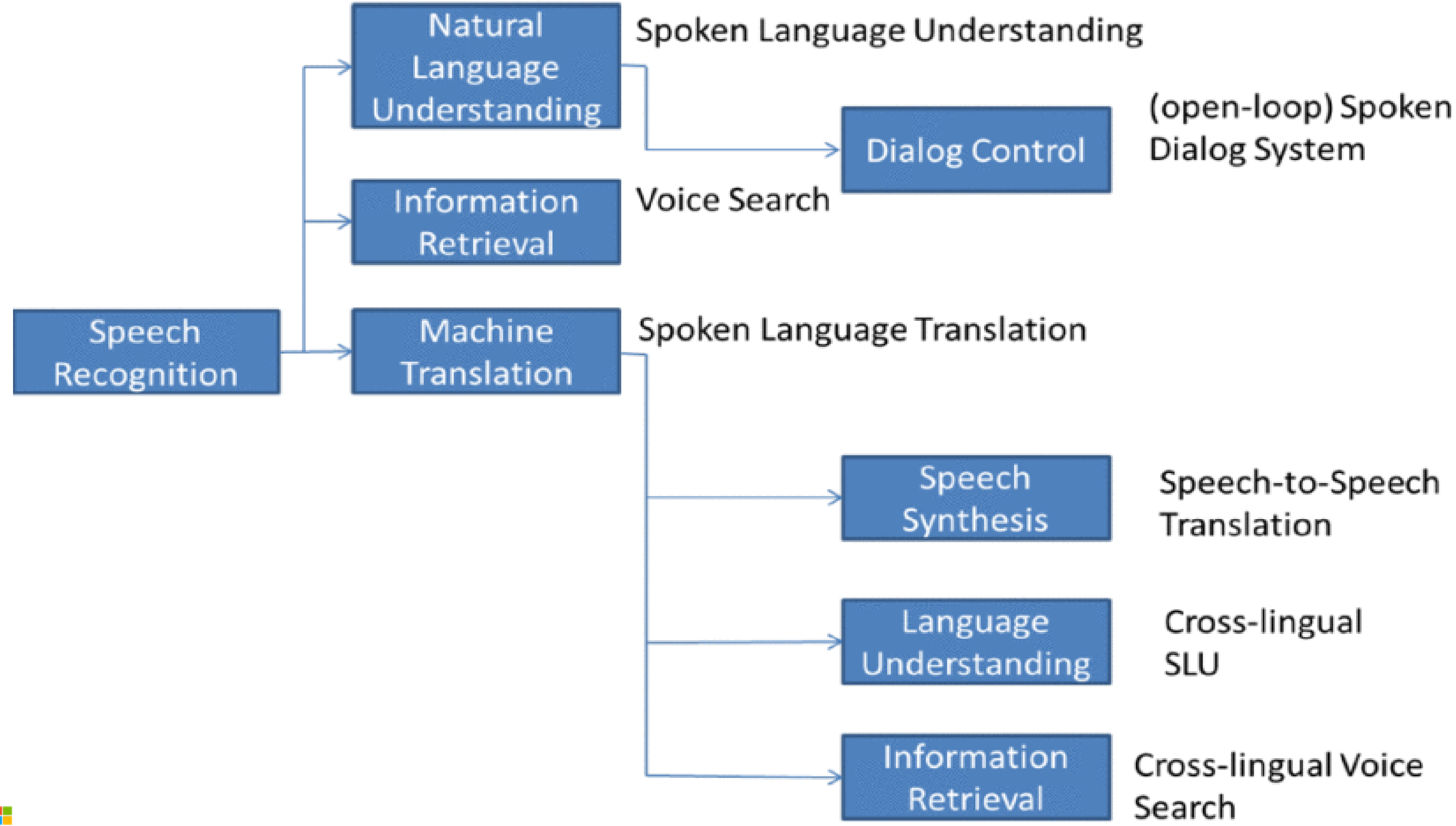
By Xiaodong He, *Senior Member IEEE*, and Li Deng, *Fellow IEEE*

ABSTRACT | Automatic speech recognition (ASR) is a central and common component of voice-driven information processing systems in human language technology, including spoken language translation (SLT), spoken language understanding (SLU), voice search, spoken document retrieval, and so on. Interfacing ASR with its downstream text-based processing tasks of translation, understanding, and information retrieval

KEYWORDS | Joint optimization; speech recognition; speech-centric information processing (SCIP); spoken language translation (SLT); spoken language understanding (SLU); voice search

I. INTRODUCTION

Microsoft


Speech Recognition
Natural Language Understanding
Spoken Language Understanding
Dialog Control
(open-loop) Spoken Dialog System
Information Retrieval
Voice Search
Machine Translation
Spoken Language Translation
Speech Synthesis
Speech-to-Speech Translation
Language Understanding
Cross-lingual SLU
Information Retrieval
Cross-lingual Voice Search

Microsoft