Neurosymbolic AI

David Cox, Ph.D.
IBM Director, MIT-IBM Watson AI Lab
IBM Research
“Artificial Intelligence”
The evolution of AI

Narrow AI
Emerging

Broad AI
Disruptive and Pervasive

General AI
Revolutionary

▼ We are here

2050 and beyond

IBM Research AI © 2018 IBM Corporation
The evolution of AI

**Narrow AI**
- Single task, single domain
- Superhuman accuracy and speed for certain tasks

**Broad AI**
- Multi-task, multi-domain
- Multi-modal
- Distributed AI
- Explainable

**General AI**
- Cross-domain learning and reasoning
- Broad autonomy
The evolution of AI

**Narrow AI**
- Single task, single domain
- Superhuman accuracy and speed for certain tasks

**Broad AI**
- Multi-task, multi-domain
- Multi-modal
- Distributed AI
- Explainable

**General AI**
- Cross-domain learning and reasoning
- Broad autonomy
Elon Musk Compares Building Artificial Intelligence To “Summoning The Demon”

Posted Oct 26, 2014 by Greg Kumparak (@grg)

17.6k SHARES

Technology

Stephen Hawking warns artificial intelligence could end mankind

By Rory Cellan-Jones Technology correspondent

2 December 2014 Technology
The evolution of AI

**Narrow AI**
Single task, single domain
Superhuman accuracy and speed for certain tasks

**Broad AI**
Multi-task, multi-domain
Multi-modal
Distributed AI
Explainable

**General AI**
Cross-domain learning and reasoning
Broad autonomy
The path to a “Broad AI” toolbox

**Explainability**

Learn more from small data

Learn to transfer

**Security**

Reasoning

**Ethics**

Infrastructure

**Physics of AI**

**Platform for AI Lifecycle**

Compute  Data & Models  Applications  Workflow
The evolution of AI

- Narrow AI: Emerging
- Broad AI: Disruptive and Pervasive
- General AI: Revolutionary

▼ We are here

2050 and beyond
So what’s “narrow” about today’s AI toolbox?
Tech 2015: Deep Learning And Machine Intelligence Will Eat The World
man in black shirt is playing guitar.

construction worker in orange safety vest is working on road.

Karpathy and Li, 2015
“Teddy Bear”

Meret Oppenheim, *Le Déjeuner en fourrure*
Wang et al. 2018
man in black shirt is playing guitar.

construction worker in orange safety vest is working on road.
a man riding a motorcycle on a beach

Lake, Ullman, Tenenbaum & Gershman, 2016
What’s this?
ObjectNet
ObjectNet

- ~50K images
- ~300 object classes
- 4 different room types
Testing ImageNet-trained models on ObjectNet

![Graph showing the accuracy of ImageNet and ObjectNet models over time. The graph includes lines for ImageNet Top-5, ImageNet Top-1, Overlap Top-5, Overlap Top-1, ObjectNet Top-5, and ObjectNet Top-1. There is a notable drop in performance between ImageNet and ObjectNet models, with a 40-45% drop observed.]
Original Top-3 inferred captions:
1. A red stop sign sitting on the side of a road.
2. A stop sign on the corner of a street.
3. A red stop sign sitting on the side of a street.

Adversarial Top-3 captions:
1. A brown teddy bear laying on top of a bed.
2. A brown teddy bear sitting on top of a bed.
3. A large brown teddy bear laying on top of a bed.

Chen et al. 2018
How many blocks are on the right of the three-level tower?

Will the block tower fall if the top block is removed?
Are there more trees than animals?

What is the shape of the object closest to the large cylinder?
Neural Networks / Deep Learning

→ “Apple”
Symbolic AI

- apple
  - origin
    - apple tree
  - structure
    - body
      - shape
        - round
      - size
      - color
        - red
        - green
      - taste
  - kind
    - fruit
  - stem
    - hand
  - apple
Neural-symbolic AI
Disentangling reasoning from vision and language understanding

Question: Are there an equal number of large things and metal spheres?

Program: equal_number(count(filter_size(Scene, Large)), count(filter_material(filter_shape(Scene, Sphere), Metal)))

Answer: Yes
End-to-End Visual Reasoning

Visual Question Answering
Q: What’s the shape of the red object?

End-to-End Neural Network

A: Sphere.

NMN [Andreas et al., 2016]
IEP [Johnson et al., 2017]
FiLM [Perez et al., 2018],
MAC [Hudson & Manning, 2018]
Stack-NMN [Hu et al., 2018]
TbD [Mascharka et al. 2018]
End-to-End Visual Reasoning

Visual Question Answering
Q: What’s the shape of the red object?

Concept (e.g., colors, shapes)
Reasoning (e.g., count)

NMN [Andreas et al., 2016]
IEP [Johnson et al., 2017]
FiLM [Perez et al., 2018],
MAC [Hudson & Manning, 2018]
Stack-NMN [Hu et al., 2018]
TbD [Mascharka et al. 2018]
End-to-End Visual Reasoning

Visual Question Answering
Q: What's the shape of the red object?

Concept (e.g., colors, shapes)  Entangled  Reasoning (e.g., count)

NMN [Andreas et al., 2016]
IEP [Johnson et al., 2017]
FiLM [Perez et al., 2018],
MAC [Hudson & Manning, 2018]
Stack-NMN [Hu et al., 2018]
TbD [Mascharaka et al. 2018]
End-to-End Visual Reasoning

Visual Question Answering
Q: What’s the shape of the red object?

Concept (e.g., colors, shapes)

Reasoning (e.g., count)

Hard to transfer

Image Captioning

Instance Retrieval

NMN [Andreas et al., 2016]
IEP [Johnson et al., 2017]
FiLM [Perez et al., 2018],
MAC [Hudson & Manning, 2018]
Stack-NMN [Hu et al., 2018]
TbD [Maschanka et al. 2018]
Task: **Visual Reasoning**

**Question:** Are there an equal number of large things and metal spheres?
Task: Visual Reasoning

Question: Are there an equal number of large things and metal spheres?
Task: Visual Reasoning

Question: Are there an equal number of large things and metal spheres?

3 large things!
Task: Visual Reasoning

Question: Are there an equal number of large things and metal spheres?

3 large things!

3 metal spheres!
Task: **Visual Reasoning**

**Question:** Are there an equal number of large things and metal spheres?

Equal? Yes!

- 3 large things!
- 3 metal spheres!
Task: Visual Reasoning

Question Understanding

Question: Are there an equal number of large things and metal spheres?

Visual Perception

Logic Reasoning

Equal? Yes!

3 large things!

3 metal spheres!
Vision (CNN) → CNN → Structured Representation

I. Scene Parsing (de-rendering)

Language (RNN)

II. Question Parsing (Program Generation)

(d) Question

How many cubes that are behind the cylinder are large?

Encoder

LSTM

1. filter_shape(scene, cylinder)

LSTM

3. filter_shape(scene, cube)

LSTM

4. filter_size(scene, large)

LSTM

5. count(scene)

(e) Program

III. Program Execution

1. filter_shape
2. relate
3. filter_shape
4. filter_size
5. count

Answer: 3

Structured Program
Incorporate Concepts in Visual Reasoning

NS-VQA [Yi et al. 2018]

Vision

Scene Parsing

Language

Q: What’s the shape of the red object?
Incorporate Concepts in Visual Reasoning

Vision

Scene Parsing

<table>
<thead>
<tr>
<th>ID</th>
<th>Color</th>
<th>Shape</th>
<th>Material</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Green</td>
<td>Cube</td>
<td>Metal</td>
</tr>
</tbody>
</table>

Language

Q: What’s the **shape** of the **red** object?
Incorporate Concepts in Visual Reasoning

Vision

<table>
<thead>
<tr>
<th>ID</th>
<th>Color</th>
<th>Shape</th>
<th>Material</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Green</td>
<td>Cube</td>
<td>Metal</td>
</tr>
<tr>
<td>2</td>
<td>Red</td>
<td>Sphere</td>
<td>Rubber</td>
</tr>
</tbody>
</table>

Language

Q: What’s the shape of the red object?

Program

Filter(Red) → Query(Shape)
Incorporate Concepts in Visual Reasoning

**Vision**

<table>
<thead>
<tr>
<th>ID</th>
<th>Color</th>
<th>Shape</th>
<th>Material</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Green</td>
<td>Cube</td>
<td>Metal</td>
</tr>
<tr>
<td>2</td>
<td>Red</td>
<td>Sphere</td>
<td>Rubber</td>
</tr>
</tbody>
</table>

**Language**

Q: What’s the shape of the red object?

**Semantic Parsing**

Filter(\textcolor{red}{Red}) \quad \Rightarrow \quad \text{Query(Shape)}

**Program**

NS-VQA [Yi et al. 2018]
Incorporate Concepts in Visual Reasoning

Vision

Scene Parsing

Language
Q: What’s the shape of the red object?

Semantic Parsing

ID | Color | Shape | Material
---|-------|-------|----------
1  | Green | Cube  | Metal    
2  | Red   | Sphere| Rubber   

Filter(\text{Red})

Query(\text{Shape})

Program

NS-VQA [Yi et al. 2018]
Incorporate Concepts in Visual Reasoning

Visual Reasoning

<table>
<thead>
<tr>
<th>ID</th>
<th>Color</th>
<th>Shape</th>
<th>Material</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Green</td>
<td>Cube</td>
<td>Metal</td>
</tr>
<tr>
<td>2</td>
<td>Red</td>
<td>Sphere</td>
<td>Rubber</td>
</tr>
</tbody>
</table>

Language

Q: What’s the shape of the red object?

Program

Filter(Red)

Query(Shape)

Symbolic Reasoning

NS-VQA [Yi et al. 2018]
Incorporate Concepts in Visual Reasoning

**Vision**

<table>
<thead>
<tr>
<th>ID</th>
<th>Color</th>
<th>Shape</th>
<th>Material</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Green</td>
<td>Cube</td>
<td>Metal</td>
</tr>
<tr>
<td>2</td>
<td>Red</td>
<td>Sphere</td>
<td>Rubber</td>
</tr>
</tbody>
</table>

**Language**

Q: What’s the **shape** of the **red** object?

**Semantic Parsing**

Filter(\textcolor{red}{Red})

Query(Shape)

Symbolic Reasoning

**Program**

\textcolor{red}{Sphere}

NS-VQA [Yi et al. 2018]
Advantage 1: **High Accuracy**

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>92.6</td>
</tr>
<tr>
<td>RN</td>
<td>95.5</td>
</tr>
<tr>
<td>IEP</td>
<td>96.9</td>
</tr>
<tr>
<td>FiLM</td>
<td>97.6</td>
</tr>
<tr>
<td>MAC</td>
<td>98.9</td>
</tr>
<tr>
<td>TbD</td>
<td>99.1</td>
</tr>
<tr>
<td>NS-VQA (Ours)</td>
<td><strong>99.8</strong></td>
</tr>
</tbody>
</table>

Advantage 2: **Data Efficiency**

High accuracy when trained with just 1% the of the data that other methods require

![Graph showing data efficiency comparison](image_url)

[Yi et al. NeurIPS 2018]
Advantage 3: Transparency and Interpretability

**Question:** Are there more yellow matte things that are right of the gray ball than cyan metallic objects?

**Answer:** no

[Yi et al. NeurIPS 2018, Johnson et al. ICCV 2017]
NeurIPS 2018: Neurosymbolic VQA: Properties (e.g. “color”) and values (“red”) predefined

ICLR 2019: Neurosymbolic Concept Learner: Properties predefined, can learn new values autonomously

NeurIPS 2019: Neurosymbolic Metaconcept Learner: Autonomously learns new concepts

ICML 2020 (target submission): Real world images

less predefined, more autonomous →
Q: What's the **shape** of the **red** object?
Q: What's the shape of the red object?
Q: What’s the **shape** of the **red** object?
Q: What's the **shape** of the **red** object?
Symbolic AI

Reproduced from Minsky, 1991
### Meta-concept Learning

**Visual reasoning questions**

<table>
<thead>
<tr>
<th><strong>color:</strong> red</th>
<th>Q: Is there any <strong>red cube</strong>?</th>
<th>A: Yes.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>color:</strong> green</td>
<td>Q: Is there any <strong>green block</strong>?</td>
<td>A: Yes</td>
</tr>
</tbody>
</table>

CLEVR (Johnson et al. 2017)

**Metaconcept questions**

<table>
<thead>
<tr>
<th>Q: Is red a <strong>same kind</strong> of concept as green?</th>
<th>A: Yes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q: Is cube a <strong>synonym</strong> of block?</td>
<td>A: Yes.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q: Is there any <strong>Ivory Gull</strong>?</th>
<th>A: Yes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q: Is there any <strong>Laridae</strong>?</td>
<td>A: Yes.</td>
</tr>
<tr>
<td>Q: Is there any <strong>Black Tern</strong>?</td>
<td>A: Yes.</td>
</tr>
<tr>
<td>Q: Is there any <strong>Laridae</strong>?</td>
<td>A: Yes.</td>
</tr>
<tr>
<td>Q: Is Laridae a <strong>hypernym</strong> of Ivory gull?</td>
<td>A: Yes.</td>
</tr>
</tbody>
</table>

Han et al. NeurIPS 2019
Augmenting VQA with Metaconcepts

**Visual reasoning questions**

- **color:** red
  - Q: Is there any red cube?  
    A: Yes.

- **color:** green
  - Q: Is there any green block?  
    A: Yes.

**CLEVR**  
*Johnson et al. 2017*

**+ Metaconcept questions**

- Q: Is red a same kind of concept as green?  
  A: Yes.

- Q: Is cube a synonym of block?  
  A: Yes.

**Laridae**  
*Wah et al. 2011*

- Q: Is there any Ivory Gull?  
  A: Yes.
- Q: Is there any Laridae?  
  A: Yes.
- Q: Is there any Black Tern?  
  A: Yes.
- Q: Is there any Laridae?  
  A: Yes.

- Q: Is Laridae a hypernym of Ivory gull?  
  A: Yes.
Visual reasoning questions

Q: Is there any red object?  

P: \text{Exist( Filter(\text{red}) )}

Metaconcept questions

Q: Is red a same kind of concept as yellow?  

P: \text{MetaVerify( red, yellow, same-kind )}

Answer: Yes
Generalization

Metaconcept Generalization

Q: Is there any *airplane*?  
A: Yes

Q: Is there any *plane*?  
A: Yes

Q: Is there any *kid*?  
A: Yes

Q: Is there any *child*?  
A: Yes

Q: Is airplane a synonym of plane?  
A: Yes

Q: Is kid a synonym of child?  
A: Yes

Training

Testing: metaconcepts on unseen pairs of concepts
# Generalization

## Metaconcept Generalization: Results

<table>
<thead>
<tr>
<th>Q: Is there any <em>airplane</em>?</th>
<th>A: Yes</th>
<th>Q: Is there any <em>kid</em>?</th>
<th>A: Yes</th>
<th>Q: Is airplane a <em>synonym</em> of <em>plane</em>?</th>
<th>A: Yes</th>
<th>Q: Is kid a <em>synonym</em> of <em>child</em>?</th>
<th>A: Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q: Is there any <em>plane</em>?</td>
<td>A: Yes</td>
<td>Q: Is there any <em>child</em>?</td>
<td>A: Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Training

<table>
<thead>
<tr>
<th>CLEVR</th>
<th>GQA</th>
<th>CUB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synonym</td>
<td>Synonym</td>
<td>Hypernym</td>
</tr>
<tr>
<td>Same-kind</td>
<td>Same-kind</td>
<td>Meronym</td>
</tr>
<tr>
<td>GRU (Lang. Only) [Cho et al., 2014]</td>
<td>66.3±1.4</td>
<td>74.3±5.2</td>
</tr>
<tr>
<td></td>
<td>64.7±5.1</td>
<td>80.1±5.9</td>
</tr>
<tr>
<td>GRU-CNN [Zhou et al., 2015]</td>
<td>60.9±10.6</td>
<td>76.7±8.8</td>
</tr>
<tr>
<td></td>
<td>61.5±6.6</td>
<td>78.1±4.8</td>
</tr>
<tr>
<td>BERT (question ; concept) [Jacob Devlin, 2018]</td>
<td>76.2±10.2 ; 80.2±16.1</td>
<td>75.6±1.2 ; 61.7±10.3</td>
</tr>
<tr>
<td></td>
<td>75.4±5.4 ; 80.1±10.0</td>
<td>63.1±3.2 ; 72.9±9.9</td>
</tr>
<tr>
<td>Training Testing</td>
<td>100.0±0.0</td>
<td>100.0±0.0</td>
</tr>
<tr>
<td></td>
<td>92.3±4.9</td>
<td>99.3±1.0</td>
</tr>
<tr>
<td>NS-CL [Mao et al. 2019]</td>
<td>100.0±0.0</td>
<td>91.1±1.7</td>
</tr>
<tr>
<td></td>
<td>81.2±2.8</td>
<td>66.8±4.1</td>
</tr>
<tr>
<td>VCML</td>
<td>91.1±1.7</td>
<td>94.8±1.3</td>
</tr>
<tr>
<td></td>
<td>69.1±1.7</td>
<td>92.5±1.0</td>
</tr>
</tbody>
</table>
CLEVERER: CoLlision Events for Video REpresentation and Reasoning

- Descriptive

Q: What is the material of the last object to collide with the cyan cylinder?

A: Metal

Chuang Gan w/ Kevin Xi, Yunzhu Li, Pushmeet Kohli, Jiajun Wu, Antonio Torralba & Josh Tenenbaum
• Explanatory

Q: What is responsible for the collision between the rubber and metal cylinder?

A. The presence of the yellow sphere
B. The collision between the rubber cylinder and the red rubber sphere
• Counterfactual

Q: What will happen without the cyan cylinder?

A. The red rubber sphere and the metal sphere collide
B. The red rubber sphere and the gray object collide
Looking Ahead

How many employees have over 10 years experience but have moved location in the last year?

What factors might contribute to better output from Factory A vs. Factory B?

Why is our database down?
Neurosymbolic Planning

(a) Target (b) β-TCVAE (c) DS-VAE with perturbed Z (d) DS-VAE

Srivastava et al. 2020 (submitted)

Neurosymbolic Generative Models

Neurosymbolic Safe ML/RL

Fulton et al. AAAI 2018

Neurosymbolic NLU

Wilcox et al. NAACL 2019

Neurosymbolic Code Optimization

Shi et al. ICLR 2019

Neurosymbolic Machine Common Sense

Smith et al. NeurIPS 2019
Inducing Behavioral Insight

Inferring flexible behavioral plans/policies from temporal observation data
(:action pickup
  :parameters (?b1 ?b2 - block)
  :precondition (and (on ?b1 ?b2) (hand-clear))
  :effect (and (not (hand-clear))
             (not (on ?b1 ?b2))
             (holding ?b1))
)

Task: Induce the action theory of an environment through observations
LatPlan
Mixing symbolic planning with neural networks
LatPlan
Mixing symbolic planning with neural networks

(Example with N=25)

The latent layer converges to the categorical distrib.  ↓

The output converges to the input

;; Translates to a PDDL model below:
(:action slide-up ...
 :precondition
  (and (empty ?x ?y-old)
        (at ?x ?y-new ?p) ...)
 :effects
  (and (not (empty ?x ?y-old))
       (empty ?x ?y-new)
       (not (at ?x ?y-new ?p))
       (at ?x ?y-old ?p)))
Verifiably Safe Reinforcement Learning

\[ \phi \]

\{accel, brake, turn\}

Observe Reward

Use a theorem prover to prove:

\[ \text{init} \rightarrow [\{{\text{accel}, brake}\}; \text{ODEs}]^* \text{safe} \]

is correctly monitored by \( \phi \).

Nathan Fulton
IBM
NEURAL NETWORKS  +  SYMBOLIC AI
Causal Inference
Beyond Correlation—inferring and testing for causal relationships in complex systems

http://tylervigen.com/spurious-correlations
**Feature** ▶ Christmas 2016: Food for Thought

**Is caviar a risk factor for being a millionaire?**

*BMJ* 2016; 355 doi: [https://doi.org/10.1136/bmj.i6536](https://doi.org/10.1136/bmj.i6536) (Published 09 December 2016)
Cite this as: *BMJ* 2016;355:i6536

Anders Hultfeldt, postdoctoral scholar

**Author affiliations**

Correspondence to: A Hultfeldt ahultfel@stanford.edu