on AI=DN+RL vs. the real brain and behavior
executive summary
on AI=DN+RL vs. the real brain & behavior

(1) a terminological preamble

(2) Deep Networks
  ✦ brain: task trade-offs, divisions, hierarchies, and more
  ✦ behavior: beyond stimulus/response

(3) Reinforcement Learning
  ✦ brain: what lies below the cortex
  ✦ behavior: what do animals learn when they do RL?

(4) two general lessons for cognitive science and AI

(5) concluding remarks
Google’s A.I. is training itself to count calories in food photos pops.ci/PiwQcb
AI and the brain
Attempting to Code the Human Brain

Startups, Tech Giants Expand World of Artificial Intelligence; Software

By EVELYN M. RUSLI
Updated Feb. 3, 2014 7:59 p.m. ET

Somewhere, in a glass building several miles outside of San Francisco, a computer is imagining what a cow looks like.

Its software is visualizing cows of varying sizes and poses, then drawing crude digital renderings, not from a collection of photographs, but rather from the software's "imagination."
AI Masters Go, Gets Artsy, Reads into Human Behavior, and More – This Week in Artificial Intelligence 03-12-16
on AI=DN+RL vs. the real brain & behavior

(1) a terminological preamble

(2) Deep Networks
  ✦ brain: task trade-offs, divisions, hierarchies, and more
  ✦ behavior: beyond stimulus/response

(3) Reinforcement Learning
  ✦ brain: what lies below the cortex
  ✦ behavior: what do animals learn when they do RL?

(4) two general lessons for cognitive science and AI

(5) concluding remarks
Deep Networks

very high level representation:

\[
\begin{array}{ccc}
\text{MAN} & \text{SITTING} & \ldots \\
\end{array}
\]

... etc ...

slightly higher level representation

raw input vector representation:

\[\mathbf{x} = \begin{bmatrix} 23 & 19 & 20 & \cdots & 18 \end{bmatrix} \]

\[x_1, x_2, x_3, \ldots, x_n\]
Deep Networks

Convolutional Neural Network

very high level representation:

\[
\begin{array}{c}
\text{MAN} \\
\text{SITTING} \\
\text{...}
\end{array}
\]

... etc ...

slightly higher level representation

raw input vector representation:

\[ \mathbf{x} = [23, 19, 20, \ldots, 18] \]
Deep Networks

Deep Belief Network

Convolutional Neural Network

very high level representation:

MAN  SITTING  ...  

...  etc  ...

slightly higher level representation

raw input vector representation:

\[ \mathbf{v} = [23, 19, 20, \ldots, 18] \]
Deep Networks, and the “mammal brain”

“Whereas most current learning algorithms correspond to shallow architectures (1, 2 or 3 levels), the mammal brain is organized in a deep architecture [173] with a given input percept represented at multiple levels of abstraction, each level corresponding to a different area of cortex.”

Deep Networks, and the “mammal brain”

“Whereas most current learning algorithms correspond to shallow architectures (1, 2 or 3 levels), the mammal brain is organized in a deep architecture [173] with a given input percept represented at multiple levels of abstraction, each level corresponding to a different area of cortex.”

let’s stand back for a sec...

“Whereas most current learning algorithms correspond to shallow architectures (1, 2 or 3 levels), the mammal brain is organized in a deep architecture [173] with a given input percept represented at multiple levels of abstraction, each level corresponding to a different area of cortex.”


Felleman & Van Essen (2001)
let’s stand back for a sec...

"Whereas most current learning algorithms correspond to shallow architectures (1, 2 or 3 levels), the mammal brain is organized in a deep architecture with a given input percept represented at multiple levels of abstraction, each level corresponding to a different area of cortex."

Y. Bengio (2009).

Learning deep architectures for AI.

Foundations and trends in machine learning 2:1-127

Cotterill (2001)

Felleman & Van Essen (2001)
let’s stand back for a sec…

Whereas most current learning algorithms correspond to shallow architectures (1, 2 or 3 levels), the mammal brain is organized in a deep architecture with a given input percept represented at multiple levels of abstraction, each level corresponding to a different area of cortex.

Y. Bengio (2009). Learning deep architectures for AI.

Felleman & Van Essen (2001)

Cotterill (2001)
let’s stand back for a sec…

"Whereas most current learning algorithms correspond to shallow architectures (1, 2 or 3 levels), the mammal brain is organized in a deep architecture \[173\] with a given input percept represented at multiple levels of abstraction, each level corresponding to a different area of cortex."


Felleman & Van Essen (2001)

Cotterill (2001)
ok, let’s focus just on the cortex… wait, which?

anterior / frontal

posterior
posterior cortex: a tripartite sensorimotor hierarchy

Merker (2004)
hippocampus: the crossroads of sensorimotor pathways

Merker (2004)
hippocampus: the crossroads of sensorimotor pathways

Merker (2004)
tasks and trade-offs

Computational tradeoffs associated with
— posterior cortex (PC),
— hippocampus (HC), and
— frontal cortex (FC):

**Large overlapping circles in PC:**
— “where would it be good to park?”

**Small separated circles in HC:**
— “where have I parked today?”

**Isolated, self-connected representations in FC:**
— “must get to the car... must get to the car...”
— the **basal ganglia** also play a critical role in the FC system by modulating (“gating”) activations there based on learned reinforcement history.

frontal cortex: abstraction/planning hierarchy

(a) Domain-specific maintenance
   Domain-general monitoring
   Abstract plan/schema/internal monitoring

(b) Concrete features (F1 and F2)
   First order relationships F1 = F1
   Second order relationships (F1 = F1) ~ (F1 = F2)

(c) Sensory control
   Contextual control
   Episodic control
   Branching control

(d) Response conflict
   Feature conflict
   Dimension conflict
   Context conflict

Badre (2008)
frontal and posterior cortical hierarchies

Fuster (2008)
frontal and posterior cortical hierarchies

Fuster (2008)

The perception/action cycle
is it all just a [stimulus / response] mapping?

Botvinick (2008)
is it all just a [stimulus / response] mapping?

Botvinick (2008)
on AI=DN+RL vs. the real brain & behavior

(1) a terminological preamble

(2) Deep Networks
   ✦ brain: task trade-offs, divisions, hierarchies, and more
   ✦ behavior: beyond stimulus/response

(3) Reinforcement Learning
   ✦ brain: what lies below the cortex
   ✦ behavior: what do animals learn when they do RL?

(4) two general lessons for cognitive science and AI

(5) concluding remarks
the psychology of trying to understanding behavior

“I'm guided by the beauty of our weapons”

— Leonard Cohen, *First we take Manhattan*
the psychology of trying to understanding behavior

“I'm guided by the beauty of our weapons”
— Leonard Cohen, *First we take Manhattan*

“I suppose it is tempting, if the only tool you have is a hammer, to treat everything as if it were a nail.”
the psychology of trying to understanding behavior

“I'm guided by the beauty of our weapons”
— Leonard Cohen, *First we take Manhattan*

“I suppose it is tempting, if the only tool you have is a hammer, to treat everything as if it were a nail.”
the hammer: efficient methods for function approximation
with hammer in hand, look for nails

In behavior, if function approximation is the answer, what is the question?
with hammer in hand, look for nails

In behavior, if function approximation is the answer, what is the question?

“How can I best map a **stimulus** to a **response**?”
the brain as a stimulus/response, or S/R, machine: 1911

“The structural unit of the nervous system is in fact a triad, neither of whose elements has any independent existence. The sensory impression exists only for the sake of awaking the central process of reflection, and the central process of reflection exists only for the sake of calling forth the final act.”

— William James (1911)
S/R: 1949

Organization
of
Behavior
by
D. O. Hebb

Stimulus and response — and what occurs in the brain in the interval between them
T. Poggio
A Theory of How the Brain Might Work
MIT A.I. Memo No. 1253, December 1990
T. Poggio, J. Mutch, J. Leibo, L. Rosasco, and A. Tacchetti

The computational magic of the ventral stream: sketch of a theory (and why some deep architectures work)

MIT-CSAIL-TR-2012-035, December 29, 2012
“The present theory still is (unfortunately) in the ‘cortical chauvinism’ camp. Hopefully somebody will rescue it.”
natural scenes, natural language: still S/R

*Parsing natural scenes and natural language with recursive neural networks*
Proc. ICML

Figure 1: An example tree with a simple Recursive Neural Network: The same weight matrix is replicated and used to compute all non-leaf node representations. Leaf nodes are \( n \)-dimensional vector representations of words.

Figure 1: An example tree with a simple Recursive Neural Network: The same weight matrix is replicated and used to compute all non-leaf node representations. Leaf nodes are $n$-dimensional vector representations of words.
AI “beats humans on an IQ test”? 

[analogy I] 
Isotherm is to temperature as isobar is to? 
(i) atmosphere, (ii) wind, (iii) pressure, (iv) latitude, (v) current.

[analogy II] 
Identify the connection (analogy) in the following pairs: 
(analysis) book, verse, read, ACT (stage, audience, play).

[classification] 
Which is the odd one out? (i) calm, (ii) quiet, (iii) relaxed, (iv) serene, (v) unrue.

[synonym] 
Which word is closest to IRRATIONAL? 
(i) intransigent, (ii) irredeemable, (iii) unsafe, (iv) lost, (v) nonsensical.

[antonym] 
Which word is most opposite to MUSICAL? 
(i) discordant, (ii) loud, (iii) lyrical, (iv) verbal, (v) euphonious.

Solving Verbal Comprehension Questions in IQ Test by Knowledge-Powered Word Embedding 
H. Wang et al., ArXiV (2015)
AI “beats humans on an IQ test”?

[analogy I]
Isotherm is to temperature as isobar is to?
(i) atmosphere, (ii) wind, (iii) pressure, (iv) latitude, (v) current.

[analogy II]
Identify two words (one from each set of brackets) that form a connection (analogy) when paired with the words in capitals: CHAPTER (book, verse, read), ACT (stage, audience, play).

[classification]
Which is the odd one out? (i) calm, (ii) quiet, (iii) relaxed, (iv) serene, (v) unrueed.

[synonym]
Which word is closest to IRRATIONAL?
(i) intransigent, (ii) irredeemable, (iii) unsafe, (iv) lost, (v) nonsensical.

[antonym]
Which word is most opposite to MUSICAL?
(i) discordant, (ii) loud, (iii) lyrical, (iv) verbal, (v) euphonious.

Solving Verbal Comprehension Questions in IQ Test by Knowledge-Powered Word Embedding
H. Wang et al., ArXiV (2015)
AI “beats humans on an IQ test”?

[analogy I]
Isotherm is to temperature as isobar is to?
(i) atmosphere, (ii) wind, (iii) pressure, (iv) latitude, (v) current.

[analogy II]
Identify two words (one from each set of brackets) that form a connection (analogy) when paired with the words in capitals: CHAPTER (book, verse, read), ACT (stage, audience, play).

[classification]
Which is the odd one out? (i) calm, (ii) quiet, (iii) relaxed, (iv) serene, (v) unraveled.

[synonym]
Which word is closest to IRRATIONAL?
(i) intransigent, (ii) irredeemable, (iii) unsafe, (iv) lost, (v) nonsensical.

[antonym]
Which word is most opposite to MUSICAL?
(i) discordant, (ii) loud, (iii) lyrical, (iv) verbal, (v) euphonious.
natural language: treating translation as a S/R task

<table>
<thead>
<tr>
<th>ねこ</th>
<th>Cat</th>
</tr>
</thead>
<tbody>
<tr>
<td>ねこねこ</td>
<td>cat</td>
</tr>
<tr>
<td>ねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
<tr>
<td>ねこねこねこねこねこねこねこねこねこねこねこねこねこねこ</td>
<td>Cat</td>
</tr>
</tbody>
</table>

| Nekonebom cat  | Astonishingly funny |
| Right-handed Personnel | Astonishingly funny |
| Right now it's kind of you | Astonishingly funny |
| I'm afraid it's kind |astonishingly funny |
| It is a kind of catch | Astonishingly funny |
| I'm afraid it's kind | Astonishingly funny |
| I'm afraid it's kind | Astonishingly funny |
| It is a funny thing | Astonishingly funny |
| It is a human nature | Astonishingly funny |
| It is a kind of weakness | Astonishingly funny |
| It is neko-naiko nputko nputko | Astonishingly funny |
| It is neko-naiko nputko nputko | Astonishingly funny |
| It is impossible for you to do something | Astonishingly funny |
| I'm aptly aware that this is not the case | Astonishingly funny |
| I'm aptly aware that this is not the case | Astonishingly funny |
| I'm aptly aware that this is not the case | Astonishingly funny |
| It is weak | Astonishingly funny |
| It is weak | Astonishingly funny |
| It is weak | Astonishingly funny |
natural language: treating translation as a S/R task
natural language: treating discourse as an S/R task
natural language: treating discourse as an S/R task

Human: what is your job?
natural language: treating discourse as an S/R task

**Human:** what is your job?

**Machine:** i’m a lawyer
natural language: treating discourse as an S/R task

**Human:** what is your job?
**Machine:** i’m a lawyer
**Human:** what do you do?
natural language: treating discourse as an S/R task

Human: what is your job?
Machine: i’m a lawyer
Human: what do you do?
Machine: i’m a doctor
Context: Both its sun-speckled shade and the cool grass beneath were a welcome respite after the stifling kitchen, and I was glad to relax against the tree’s rough, brittle bark and begin my breakfast of buttery, toasted bread and fresh fruit. Even the water was tasty, it was so clean and cold.

Target sentence: It almost made up for the lack of ____.

Target word: coffee

Context: “Yes, I thought I was going to lose the baby.” “I was scared too,” he stated, sincerity flooding his eyes. “You were?” “Yes, of course. Why do you even ask?” “This baby wasn’t exactly planned for.”

Target sentence: “Do you honestly think that I would want you to have a ____?”

Target word: miscarriage

Context: The battery on Logan’s radio must have been on the way out. So he told himself. There was no other explanation beyond Cygan and the staff at the White House having been overrun. Lizzie opened her eyes with a flutter. They had been on the icy road for an hour without incident.

Target sentence: Jack was happy to do all of the ____.

Target word: driving
S/R: one task to rule them all
Montezuma’s Revenge

Mnih et al. (2015).
Human-level control through deep reinforcement learning
Nature 518:529-533
Montezuma’s Revenge

Mnih et al. (2015).
Human-level control through deep reinforcement learning
Nature 518:529-533
Montezuma’s Revenge

Mnih et al. (2015).
*Human-level control through deep reinforcement learning*
Nature 518:529-533
the world is neither a board game, nor a set of board games

R. Munroe (xkcd)
structured problems (e.g., free analogy) do not reduce to S/R

Bongard’s problem #75 (1970)
“What we have is a circuit, not an arc or broken segment of a circle. [...] The motor response determines the stimulus, just as truly as sensory stimulus determines movement. Indeed, the movement is only for the sake of determining the stimulus, of fixing what kind of a stimulus it is, of interpreting it. [...] There is simply a continuously ordered sequence of acts, all adapted in themselves and in the order of their sequence, to reach a certain objective end, the reproduction of the species, the preservation of life, locomotion to a certain place. The end has got thoroughly organized into the means.”

John Dewey (1896)
The Reflex Arc Concept in Psychology
Psychological Review, 3, 357-370
Thurstone knew it: the “S/R fallacy” (1923)

L. L. Thurstone (1923).
The stimulus-response fallacy in psychology
Psychological Review 30:354-369
Thurstone knew it: the “S/R fallacy” (1923)

L. L. Thurstone (1923).
The stimulus-response fallacy in psychology
Psychological Review 30:354-369
Thurstone knew it: the “S/R fallacy” (1923)

THE STIMULUS-RESPONSE FORMULA

The stimulus ——— Human wants ——— The response

THE SELF-EXPRESSION FORMULA

Human wants ——— The stimulus ——— Conduct

L. L. Thurstone (1923).  
The stimulus-response fallacy in psychology  
Psychological Review 30:354-369
the S/R doctrine: “a red herring” (2014)
the S/R doctrine: “a red herring” (2014)

“Drosophila got introduced to many branches of neuroscience and has made fascinating contributions to these fields. Yet, has any of that spectacular progress revealed how brains work? Rather not. Why? What is the problem with brain research? The problem is the input-output doctrine. It is the wrong dogma, the red herring. The hard-wired reflex, the innate response to a stimulus, is not the essence of brain function, not the basic building block of behavior. Behavior is problem solving. It is part of the evolutionary process. It is fundamentally active.”

M. Heisenberg (2014).
*The Beauty of the Network in the Brain and the Origin of the Mind in the Control of Behavior*
how can we make sense of behavior?
an example of real-life behavior: foraging

Figure 3.9. Animals are faced with a choice of plants at a feeding station which offer different potential instantaneous intake rates, nutrient density, and secondary compounds.
an example of real-life behavior: foraging

Figure 3.9. Animals are shown foraging for food at different potential internal considerations.

Figure 3.4. Hierarchy of large grazers’ physiological and behavioral needs which affect patterns of landscape use. These threshold levels trigger initiation and velocity of movement and frequency of encounter of locals within a landscape.
an example of real-life behavior: foraging

internal considerations

external constraints
an example of real-life behavior: hunting

“Salticids are also distinctive for development of behavioral flexibility, including conditional predatory strategies, the use of trial-and-error to solve predatory problems, and the undertaking of detours to reach prey.” — Jackson & Pollard, Ann. Rev. Entomol. (1996)
an example of real-life behavior: hunting

“Salticids are also distinctive for development of behavioral flexibility, including conditional predatory strategies, the use of trial-and-error to solve predatory problems, and the undertaking of detours to reach prey.” — Jackson & Pollard, Ann. Rev. Entomol. (1996)
an example of real-life behavior: hunting

“Salticids are also distinctive for development of behavioral flexibility, including conditional predatory strategies, the use of trial-and-error to solve predatory problems, and the undertaking of detours to reach prey.” — Jackson & Pollard, Ann. Rev. Entomol. (1996)
an example of real-life behavior: hunting

“Salticids are also distinctive for development of behavioral flexibility, including conditional predatory strategies, the use of trial-and-error to solve predatory problems, and the undertaking of detours to reach prey.” — Jackson & Pollard, Ann. Rev. Entomol. (1996)
on AI=DN+RL vs. the real brain & behavior

(1) a terminological preamble

(2) Deep Networks
   ✦ brain: task trade-offs, divisions, hierarchies, and more
   ✦ behavior: beyond stimulus/response

(3) Reinforcement Learning
   ✦ brain: what lies below the cortex
   ✦ behavior: what do animals learn when they do RL?

(4) two general lessons for cognitive science and AI

(5) concluding remarks
reinforcement learning (RL) in sequential behavior

![Diagram of TD-learning](image)

Scholarpedia (2008)
need to address: serial order, hierarchy, action selection

Scholarpedia (2008)

Hull (1934)
need to address: serial order, hierarchy, action selection

TD-learning

Scholarpedia (2008)

Hull (1934)

Botvinick, Niv, & Barto (2009)
on AI=DN+RL vs. the real brain & behavior

(1) a terminological preamble

(2) Deep Networks
   ✦ brain: task trade-offs, divisions, hierarchies, and more
   ✦ behavior: beyond stimulus/response

(3) Reinforcement Learning
   ✦ brain: what lies below the cortex
   ✦ behavior: what do animals learn when they do RL?

(4) two general lessons for cognitive science and AI

(5) concluding remarks
RL and the basal ganglia circuits

Botvinick, Niv, & Barto (2009)
RL and the basal ganglia circuits

Botvinick, Niv, & Barto (2009)
RL and the basal ganglia circuits

Botvinick, Niv, & Barto (2009)
basal ganglia: the Go/NoGo model

By enhancing synaptic plasticity, dopamine release during unexpected rewards can drive the animal to learn to perform the action that led to the reward.

Atallah, Frank & O’Reilly (2004):

The cortico–striato–thalamo–cortical loops include the direct and indirect pathways of the basal ganglia. In the striatum, the “Go” cells facilitate the execution of an action represented in cortex. The “NoGo” cells have an opposing effect, suppressing actions from getting executed.
Recent anatomical investigations have revealed a rather more complex organization in which the transformations that are applied to the inputs to generate outputs are less easy to predict.”

Redgrave et al. (2010)
basal ganglia: not one loop, but many
Functional territories represented at the level of cerebral cortex are maintained throughout the basal ganglia nuclei and thalamic relays. However, for each loop, the stages in the cortex, basal ganglia and thalamus offer opportunities for activity inside the loop to be modified or modulated by signals from outside the loop.

Redgrave et al. (2011)
basal ganglia: not one loop, but many

“Functional territories represented at the level of cerebral cortex are maintained throughout the basal ganglia nuclei and thalamic relays. However, for each loop, the stages in the cortex, basal ganglia and thalamus offer opportunities for activity inside the loop to be modified or modulated by signals from outside the loop.”

Redgrave et al. (2011)
We suggest that the ventral striatum, at the convergence of hippocampal formation and prefrontal cortex circuits for replay of memory and dopamine signals for learning, is necessary for controlling the behavioural expression of episodic memory replay and, via the dual shell pathway mechanism above, eliciting phasic dopamine bursts to reinforce each stage of behaviour as a part of an action sequence.

Humphries, M. D., & T. J. Prescott (2010). The ventral basal ganglia, a selection mechanism at the crossroads of space, strategy, and reward. Progress in Neurobiology 90:385-417
basal ganglia: even more complicated

Humphries & Prescott (2010)
basal ganglia: further complications

Humphries & Prescott (2010)
basal ganglia: insanely complicated

Humphries & Prescott (2010)
on AI=DN+RL vs. the real brain & behavior

(1) a terminological preamble

(2) Deep Networks
   - brain: task trade-offs, divisions, hierarchies, and more
   - behavior: beyond stimulus/response

(3) Reinforcement Learning
   - brain: what lies below the cortex
   - behavior: what do animals learn when they do RL?

(4) two general lessons for cognitive science and AI

(5) concluding remarks
what do animals actually learn when they do RL?

- **How to balance habits and goals**
  [because merely following the reinforcement path leads to addiction]

- **How the world works**
  [because utility, and therefore reward, is context- and experience-dependent]
“Instrumental conditioning is mediated in the brain by two separate computational processes, each with independent sub-goals:

(i) a mechanism that determines whether an unpredicted sensory event is caused by the agent (agency), and if so, the agent is motivated through trial-and-error to discover the causal components of its behaviour; and

(ii) a mechanism whereby reward modulates the processing in structures that provide afferent signals to the basal ganglia, in particular, the striatum.”

Redgrave et al. (2011)
goal-directed vs. habitual control and causality

“Early analyses of instrumental behaviour stressed the importance of reinforced associations between stimuli and responses. However, during the past 20 years it has become clear that animals can also encode causal relationships between actions and outcomes.”

Redgrave et al. (2010)
goal-directed vs. habitual control and causality

“Early analyses of instrumental behaviour stressed the importance of reinforced associations between stimuli and responses. However, during the past 20 years it has become clear that animals can also encode causal relationships between actions and outcomes.”

Redgrave et al. (2010)
goal-directed vs. habitual control and causality

“Early analyses of instrumental behaviour stressed the importance of reinforced associations between stimuli and responses. However, during the past 20 years it has become clear that animals can also encode causal relationships between actions and outcomes.”

Redgrave et al. (2010)
goal-directed vs. habitual control and causality

“Early analyses of instrumental behaviour stressed the importance of reinforced associations between stimuli and responses. However, during the past 20 years it has become clear that animals can also encode causal relationships between actions and outcomes.”

Redgrave et al. (2010)
Most of the work has been and is being done within the S/R paradigm. In contrast, realistic behavior is 

*agentic,* 

*sequentially and hierarchically structured,* 

*situated,* and 

*dynamic.*

What problem(s) does it give rise to?
the overarching problem in managing real-life behavior

Most of the work has been and is being done within the S/R paradigm. In contrast, realistic behavior is

*agentic,*

*sequentially and hierarchically structured,*

*situated,* and

*dynamic.*

What problem(s) does it give rise to?

The overarching problem in managing behavior is:

*mapping a stimulus to a response*

*deciding what to do next.*
the overarching problem in managing real-life behavior

Most of the work has been and is being done within the S/R paradigm. In contrast, realistic behavior is

agentic,
sequentially and hierarchically structured,
situated, and
dynamic.

What problem(s) does it give rise to?

The overarching problem in managing behavior is:

mapping a stimulus to a response
deciding what to do next.

Computationally, this amounts to

navigation/search
in a problem space,
which is open-ended and needs to be learned and updated.
on AI=DN+RL vs. the real brain & behavior

(1) a terminological preamble

(2) Deep Networks
   - brain: task trade-offs, divisions, hierarchies, and more
   - behavior: beyond stimulus/response

(3) Reinforcement Learning
   - brain: what lies below the cortex
   - behavior: what do animals learn when they do RL?

(4) two general lessons for cognitive science and AI

(5) concluding remarks
two general lessons for cognitive science and AI

(I) Flexibility pays:
to enable complex sequential behavior, learn representations that support open-ended “reasoning.”

(II) Circuitry matters:
hard choices in localist, precisely wired representations make for easier credit assignment, decision making, and action control.
(I) why flexibility is a must: self-taught AI vs. real life

Why Self-Taught Artificial Intelligence Has Trouble With the Real World

The latest artificial intelligence systems start from zero knowledge of a game and grow to world-beating in a matter of hours. But researchers are struggling to apply these systems beyond the arcade.
why flexibility is a must: limitations of Deep RL

Deep Reinforcement Learning Doesn't Work Yet

Feb 14, 2018

This mostly cites papers from Berkeley, Google Brain, DeepMind, and OpenAI from the past few years, because that work is most visible to me. I’m almost certainly missing stuff from older literature and other institutions, and for that I apologize - I’m just one guy, after all.

Introduction

Once, on Facebook, I made the following claim.

Whenever someone asks me if reinforcement learning can solve their problem, I tell them it can't. I think this is right at least 70% of the time.

WHENEVER SOMEONE ASKS ME IF RL WORKS, I TELL THEM IT DOESN'T
(I) why flexibility is a must: RL and moving goalposts
(I) why flexibility is a must: RL and moving goalposts
(I) why flexibility is a must: RL and moving goalposts

Does it even make sense to try to learn the joint probability $P(\text{situations, utility, actions})$? [note: $P$ is timeless; life is lived in time]
(II) why circuitry matters: the need for “switching control”

Sustained sequential behavior (migration, foraging, courtship rituals, glade skiing, birdsong, language) amounts to navigating through a shifting labyrinth of discrete options.

Here, control outcome must be serially “local”: every discrete step is a causal nexus, at which one and only one choice must be made (a superposition of alternatives will not do).
In learning, resorting to localist representations may keep down the complexity of **credit assignment**.
on $\text{AI=DN+RL vs. the real brain & behavior}$

1. a terminological preamble
2. Deep Networks
   - brain: task trade-offs, divisions, hierarchies, and more
   - behavior: beyond stimulus/response
3. Reinforcement Learning
   - brain: what lies below the cortex
   - behavior: what do animals learn when they do RL?
4. two general lessons for cognitive science and AI
5. concluding remarks
don’t panic

AI ≠ DN + RL.

Now what?

• Why?
  - [problem level]
    Because it pretends that behavior reduces to stimulus/response (S/R) mapping.
  - [algorithmic level]
    Because its algorithmic methods (DL; RL) are inadequate.
  - [implementation level]
    Because its architecture is ill-suited to the tasks at hand (unlike that of the brain).

• What to do?
  - [problem level]
    Understand behavior by consulting ethology, evolution.
  - [algorithmic level]
    [...]
  - [implementation level]
    [...]

Members of the Hixon Symposium (1951)

Left to right: (seated) Halstead, Lashley, Klüver, Köhler, Lorente de Nó; (standing) Bosin, Jeffress, Weiss, Lindsley, von Neumann, Nielsen, Gerard, Liddell. Dr. McCulloch was unable to be present when this picture was taken.
Members of the Hixon Symposium (1951)
Members of the Hixon Symposium (1951)
Members of the Hixon Symposium (1951)

- Engineer
- Psychologist
- Neurologist
Members of the Hixon Symposium (1951)
Members of the Hixon Symposium (1951)
the last word

Members of the Hixon Symposium (1951)
Members of the Hixon Symposium (1951)
Members of the Hixon Symposium (1951)