Two lessons

Lesson 1
Pogo sticks don’t fly.

Lesson 2
Wasps are strangely stupid.
Two lessons

Lesson 1
Pogo sticks don’t fly.
Two lessons

Lesson 1
Pogo sticks don’t fly.

Lesson 2
Wasps are strangely stupid.
An alternate history of flight

- It is the 19th century, and airplanes aren’t a thing yet.
- Three competitors at the first national flight competition:

Icarus Inc
An alternate history of flight

- It is the 19th century, and airplanes aren’t a thing yet.
- Three competitors at the first national flight competition:
An alternate history of flight

- It is the 19th century, and airplanes aren’t a thing yet.
- Three competitors at the first national flight competition:
  - Icarus Inc
  - Leonardo LLC
An alternate history of flight

- It is the 19th century, and airplanes aren’t a thing yet.
- Three competitors at the first national flight competition:

  - Icarus Inc
  - Leonardo LLC
An alternate history of flight

- It is the 19th century, and airplanes aren’t a thing yet.
- Three competitors at the first national flight competition:
  - Icarus Inc
  - Leonardo LLC
  - Timmy
After Timmy’s victory, **pogo sticks are all the rage**.

Better and better pogo sticks hit the market.

By 1930, the US is the world’s leading pogo stick nation.
By 1952, the US is occupied by Japan and Nazi Germany.
Its pogo sticks were no match for airplanes.
Nobody knows what happened to Timmy.
Like flight, language technology is the future.

Like pogo sticks, **hyped solutions** are not the answer:
- Deep Learning
- Big Data
- embeddings
- RNNs, LSTMs
- seq2seq

Let’s talk about them...
The current hype: Deep Learning

▶ One learning model is all over the media right now: **Deep Learning**
▶ Deep learning = very large and complex neural networks
▶ Neural networks imitate the human brain.

Standard model of the human brain

▶ connected network of neurons
▶ input activates neurons, which start “firing” (= emitting electrical current)
▶ current activates other neurons ⇒ activation patterns
▶ learning = strengthening connection between specific neurons
The current hype: Deep Learning

One learning model is all over the media right now: **Deep Learning**

- Deep learning = very large and complex neural networks
- Neural networks imitate the human brain.

Standard model of the human brain

- connected network of neurons
- input activates neurons, which start “firing” (= emitting electrical current)
- current activates other neurons ⇒ activation patterns
- learning = strengthening connection between specific neurons
The perceptron

- **input layer:** neurons that are sensitive to input
- **output layer:** neurons that represent output values
- **connections:** weighted links between input and output layer
- most activated output neuron represents decision

![Diagram of a perceptron with connections and weights](image)

- **Hi** (input) to **ham** (output): 3
- **Dear** (input) to **ham** (output): 10
- **Dear** (input) to **spam** (output): 1
- **Emily** (input) to **spam** (output): 5
- **ham** (output) to **spam** (output): 0
The perceptron

- **input layer:** neurons that are sensitive to input
- **output layer:** neurons that represent output values
- **connections:** weighted links between input and output layer
- most activated output neuron represents decision

![Perceptron Diagram](image-url)
The perceptron

- **input layer**: neurons that are sensitive to input
- **output layer**: neurons that represent output values
- **connections**: weighted links between input and output layer
- most activated output neuron represents decision
Perceptrons are linear functions (matrix multiplication).

Example (Computing *Hi Dear*)

\[
\begin{pmatrix} 1 & 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 3 & 1 \\ 0 & 5 \\ 10 & 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}
\]

Since matrix multiplication is associative, every *multi-layer perceptron can be reduced to one layer*.

Example (Adding more weight to spam score)

\[
\begin{pmatrix} 3 & 6 \end{pmatrix} \otimes \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 9 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} \times \begin{pmatrix} 9 & 1 \\ 0 & 5 \\ 30 & 0 \end{pmatrix}
\]
Modern neural networks are just the perceptron on steroids.

There’s a lot of jargon:
- RNNs
- LSTMs
- embeddings
- encoder/decoder
- seq2seq

Mathematically, modern neural networks intersperse linear functions (= perceptron layers) with non-linear functions.

And that’s it.
Are neural networks right for you?

- **Data hungry**
  If you don’t have tons of data, don’t even try.

- **Resource hungry**
  Large networks take forever to train.

- **Black boxes**
  Nobody knows what they do. It’s *trial and error*.

- **Do not scale**
  If your objectives change, you’re back to square 1.

- **No safeties**
  If something goes wrong, it often goes really wrong.
This is your machine learning system?

Yup! You pour the data into this big pile of linear algebra, then collect the answers on the other side.

What if the answers are wrong?

Just stir the pile until they start looking right.
A common reply

We’re not aiming for perfection, it just has to be good enough.

Every engineer ever

My reply

1 It’s still a bad choice for engineering:
   ▶ expensive (resources, time, labor)
   ▶ scales badly
   ▶ not modular

2 Your notions of “good enough” are wrong:
   ▶ precision
   ▶ recall
   ▶ F-score

They all ignore error quality.
Lesson 2

Wasps are strangely stupid.
Let’s go to wasp school

Lesson 2
Wasps are strangely stupid.

Volunteer needed!
What we learned from wasp school

- Users endow systems with human-like qualities.
- When human biases are violated, the illusion breaks down.
- Breaking the illusion is jarring.

The true task of language technology

- Trick humans into considering you human-like.
- Minimize errors that violate human biases.
Biases in human cognition: a quick experiment

- blip
- not blip
- not gnok
- not bnik
- glob
- blok

- blip
- not blip
- gnok
- not bnik
- glob
- blok

- blip
- not blip
- gnok
- glob
- blok

- blip
- not blip
- gnok
- glob
- blok
Biases in human cognition: a quick experiment

- blip
- not gnok
- not bnik
- glob
- blok
- blip
- not blip
- not gnok
- not bnik
- glob
- blok
- blip
- gnok
- bnik
- glob
- blok
- blip
- gnok
- bnik
- glob
- blok
- blip
Biases in human cognition: a quick experiment

- blip
- not gnok
- not bnik
- goblop
- blok

- not blip
- not gnok
- not bnik
- goblop
- blok

- blip
- gnok
- bnik
- not goblop
- goblop
- blok

- not blip
- gnok
- bnik
- not goblop
- goblop
- blok

- blip
- blip
- blip
<table>
<thead>
<tr>
<th></th>
<th>blip</th>
<th>not blip</th>
<th>blip</th>
<th>not blip</th>
<th>blip</th>
<th>blip</th>
<th>blip</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>blip</td>
<td>not gnok</td>
<td>not gnok</td>
<td>gnok</td>
<td>gnok</td>
<td>blip</td>
<td>blip</td>
</tr>
<tr>
<td>2</td>
<td>not bnik</td>
<td>not bnik</td>
<td>bnik</td>
<td>not bnik</td>
<td>bnik</td>
<td>bnik</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>glop</td>
<td>not glop</td>
<td>glop</td>
<td>glop</td>
<td>glop</td>
<td>glop</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td></td>
</tr>
</tbody>
</table>

**Biases in human cognition: a quick experiment**
Biases in human cognition: a quick experiment

blip  not blip  blip  not blip  blip  blip  blip
not gnok  not gnok  gnok  gnok  blip  blip  blip
not bnik  not bnik  bnik  not bnik  blip  blip  blip
glop  glop  not glop  glop  blip  blip  blip
blok  blok  blok  blok  blip  blip  blip
blip  not red
Biases in human cognition: a quick experiment

- blip
- not blip
- blip
- not blip
- blip
- not blip
- blip
- not blip
- blip
- blip
- blip
- blip

- not gnok
- not gnok
- gnok
- gnok
- gnok
- gnok

- not bnik
- not bnik
- bnik
- not bnik
- bnik
- not bnik

- glob
- glob
- not glob
- not glob
- glob
- glob

- blok
- blok
- blok
- blok
- not red
Biases in human cognition: a quick experiment

- **blip**
- not **blip**
- not **gnok**
- not **bnik**
- **glop**
- **blok**

- **blip**
- not **blip**
- not **gnok**
- not **bnik**
- **glop**
- **blok**

- **blip**
- **gnok**
- **bnik**
- not **bnik**
- **glop**
- **blok**

- **blip**
- **gnok**
- **gnok**
- **gnok**
- **gnok**
- **gnok**

- blip  not red
Biases in human cognition: a quick experiment

<table>
<thead>
<tr>
<th>blip</th>
<th>not blip</th>
<th>blip</th>
<th>not blip</th>
<th>blip</th>
<th>blip</th>
<th>blip</th>
</tr>
</thead>
<tbody>
<tr>
<td>not gnok</td>
<td>not gnok</td>
<td>gnok</td>
<td>gnok</td>
<td>gnok</td>
<td>gnok</td>
<td>gnok</td>
</tr>
<tr>
<td>not bnik</td>
<td>not bnik</td>
<td>bnik</td>
<td>not bnik</td>
<td>bnik</td>
<td>bnik</td>
<td>bnik</td>
</tr>
<tr>
<td>glop</td>
<td>glop</td>
<td>not glop</td>
<td>glop</td>
<td>glop</td>
<td>glop</td>
<td>glop</td>
</tr>
<tr>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
</tr>
</tbody>
</table>

blip not red
Biases in human cognition: a quick experiment

- blip
- not gnok
- not bnik
- glop
- blok

- not blip
- blip
- not gnok
- gnok
- not bnik
- glop
- blok

- blip
- not blip
- blip
- gnok
- not bnik
- not glop
- blok

- blip
- not red

- gnok
- brown or rectangular
Biases in human cognition: a quick experiment

blip  not blip  blip  not blip  blip  blip  blip
not gnok  not gnok  gnok  gnok  gnok  gnok  gnok
not bnik  not bnik  bnik  not bnik  bnik  bnik  bnik
glop  glop  not glop  glop  glop  glop  glop
blok  blok  blok  blok  blok  blok  blok

blip  not red

gnok  brown or rectangular
Biases in human cognition: a quick experiment

<table>
<thead>
<tr>
<th>blip</th>
<th>not blip</th>
<th>blip</th>
<th>not blip</th>
<th>blip</th>
<th>blip</th>
<th>blip</th>
</tr>
</thead>
<tbody>
<tr>
<td>not gnok</td>
<td>not gnok</td>
<td>gnok</td>
<td>gnok</td>
<td>gnok</td>
<td>gnok</td>
<td>gnok</td>
</tr>
<tr>
<td>not bnik</td>
<td>not bnik</td>
<td>bnik</td>
<td>not bnik</td>
<td>bnik</td>
<td>bnik</td>
<td>bnik</td>
</tr>
<tr>
<td>glob</td>
<td>glob</td>
<td>not glob</td>
<td>glob</td>
<td>glob</td>
<td>glob</td>
<td>glob</td>
</tr>
<tr>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
</tr>
</tbody>
</table>

blip not red

gnok brown or rectangular
### Biases in human cognition: a quick experiment

<table>
<thead>
<tr>
<th>blip</th>
<th>not blip</th>
<th>blip</th>
<th>not blip</th>
<th>blip</th>
<th>blip</th>
<th>blip</th>
</tr>
</thead>
<tbody>
<tr>
<td>not gnok</td>
<td>not gnok</td>
<td>gnok</td>
<td>gnok</td>
<td>gnok</td>
<td>gnok</td>
<td>gnok</td>
</tr>
<tr>
<td>not bnik</td>
<td>not bnik</td>
<td>bnik</td>
<td>not bnik</td>
<td>bnik</td>
<td>bnik</td>
<td>bnik</td>
</tr>
<tr>
<td>glop</td>
<td>glop</td>
<td>not glop</td>
<td>glop</td>
<td>glop</td>
<td>glop</td>
<td></td>
</tr>
<tr>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
</tr>
</tbody>
</table>

- **blip** not red
- **gnok** brown or rectangular
Biases in human cognition: a quick experiment

- **blip**
- not blip
- not gnok
- not bnik
- glop
- blok

- **gnok**
- brown or rectangular

- **bnik**
- blip and gnok

- **blip** not red

- **gnok** brown or rectangular

- **bnik** blip and gnok
Biases in human cognition: a quick experiment

- **blip**
- not **blip**
- **blip**
- not **blip**
- **blip**
- not **blip**
- **blip**
- not **blip**

- not **gnok**
- not **gnok**
- **gnok**
- **gnok**
- **gnok**
- **gnok**
- **gnok**

- not **bnik**
- not **bnik**
- **bnik**
- **bnik**
- **bnik**

- **glop**
- **glop**
- not **glop**
- **glop**
- **glop**

- **blok**
- **blok**
- **blok**
- **blok**

- **blip** not red
- **gnok** brown or rectangular
- **bnik** blip and gnok
Biases in human cognition: a quick experiment

- **blip**
- **not blip**
- **blip**
- **not blip**
- **blip**
- **blip**
- **blip**

- **not gnok**
- **not gnok**
- **gnok**
- **gnok**
- **gnok**
- **gnok**
- **gnok**

- **not bnik**
- **not bnik**
- **bnik**
- **bnik**
- **bnik**
- **bnik**
- **bnik**

- **glop**
- **glop**
- **not glop**
- **glop**
- **glop**
- **glop**
- **glop**

- **blok**
- **blok**
- **blok**
- **blok**
- **blok**
- **blok**
- **blok**

**blip** not red

**gnok** brown or rectangular

**bnik** blip and gnok
Biases in human cognition: a quick experiment

- **blip**
  - not **blip**
  - not **gnok**
  - not **bnik**
  - **glop**
  - **blik**

- **not blip**
  - **not gnok**
  - **not bnik**
  - **glop**
  - **blik**

- **blip**
  - **gnok**
  - **bnik**
  - **glop**
  - **blik**

- **not blip**
  - **gnok**
  - **not bnik**
  - **not gloop**
  - **blik**

- **blip**
  - **gnok**
  - **bnik**
  - **glop**
  - **blik**

- **blip**
  - **gnok**
  - **bnik**
  - **glop**
  - **glop**

**blip** not red

**gnok** brown or rectangular

**bnik** blip and gnok
Biases in human cognition: a quick experiment

- **blip**: not red
- **gnok**: brown or rectangular
- **bnik**: blip and gnok
- **glop**: if bnik and not brown, then not rectangular
### Biases in human cognition: a quick experiment

<table>
<thead>
<tr>
<th>blip</th>
<th>not blip</th>
<th>blip</th>
<th>not blip</th>
<th>blip</th>
<th>blip</th>
<th>blip</th>
</tr>
</thead>
<tbody>
<tr>
<td>not gnok</td>
<td>not gnok</td>
<td>gnok</td>
<td>gnok</td>
<td>gnok</td>
<td>gnok</td>
<td>gnok</td>
</tr>
<tr>
<td>not bnik</td>
<td>not bnik</td>
<td>bnik</td>
<td>not bnik</td>
<td>bnik</td>
<td>bnik</td>
<td>bnik</td>
</tr>
<tr>
<td>glob</td>
<td>glob</td>
<td>not glob</td>
<td>glob</td>
<td>glob</td>
<td>glob</td>
<td>glob</td>
</tr>
<tr>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>not blok</td>
<td>not blok</td>
<td>not blok</td>
</tr>
</tbody>
</table>

- **blip**: not red
- **gnok**: brown or rectangular
- **bnik**: blip and gnok
- **glop**: if bnik and not brown, then not rectangular
Biases in human cognition: a quick experiment

- **blip**: not red
- **gnok**: brown or rectangular
- **bnik**: blip and gnok
- **glop**: if bnik and not brown, then not rectangular
**Biases in human cognition: a quick experiment**

<table>
<thead>
<tr>
<th>blip</th>
<th>not blip</th>
<th>blip</th>
<th>not blip</th>
<th>blip</th>
<th>blip</th>
<th>blip</th>
<th>blip</th>
</tr>
</thead>
<tbody>
<tr>
<td>not gnok</td>
<td>not gnok</td>
<td>gnok</td>
<td>gnok</td>
<td>gnok</td>
<td>gnok</td>
<td>gnok</td>
<td>gnok</td>
</tr>
<tr>
<td>not bnik</td>
<td>not bnik</td>
<td>bnik</td>
<td>not bnik</td>
<td>bnik</td>
<td>bnik</td>
<td>bnik</td>
<td>bnik</td>
</tr>
<tr>
<td>glob</td>
<td>glob</td>
<td>not glob</td>
<td>glob</td>
<td>glob</td>
<td>glob</td>
<td>glob</td>
<td>glob</td>
</tr>
<tr>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>blok</td>
<td>not blok</td>
<td>not blok</td>
<td>not blok</td>
<td>not blok</td>
</tr>
</tbody>
</table>

- **blip** not red
- **gnok** brown or rectangular
- **bnik** blip and gnok
- **glop** if bnik and not brown, then not rectangular
Biases in human cognition: a quick experiment

**blip**  not red

**gnok**  brown or rectangular

**bnik**  blip and gnok

**glop**  if bnik and not brown, then not rectangular

**blok**  bnik or glob, but not both
Human language bias and unreasonable expectations

Human bias is a much bigger issue for language than for, say, cars.
Robots’ Narrow Range of Language Competence

completely human
perfect but weird voice

Arnold
completely human
Robots’ Narrow Range of Language Competence

completely human

perfect but weird voice
Robots’ Narrow Range of Language Competence

- completely human
- perfect but weird voice
- Arnold
Fake it till you make it!

- Language technology is still largely smokes and mirrors.
- User-facing software has to fool the user.
- Neural networks can’t do that, they will trip up in weird ways.
- A hand-designed model is a better conman.

Be wrong, not weird

- 90% performance can be better than 99%.
- It depends on how unnatural the errors are.
So what’s the alternative?

▶ “Traditional methods” that build on formal language theory are still very useful.
▶ Many long-known results have been forgotten, new ones have gone unnoticed.

**Automata**

- Start state
- Transition states:
  - D: old
  - the: John, Mary
  - woman, hobo
- Final state: S

**CFGs**

- S → NP VP
- NP → Det (AP) N | PN
- AP → A (AP)
- VP → V | Vsub S
- Det → the
- A → old
- N → man | woman
- V → died
- Vsub → thinks
Using automata

- **Intersection parsing of CFGs**
  - parsing = generation with CFG that recognizes regular language containing input
  - done via Bar-Hillel construction
  - combine regex pattern matching with structural description

- **Automata approximation of CFGs**
  - convert CFG into almost equivalent automaton
  - linear-time parsing

- **Discourse parsing**
  - parse an entire text rather than individual sentences
  - simplifies meaning extraction

- **Semiring parsing**
  - modularize parsing algorithm for multiple tasks
  - recognition, structure, best $n$-structures, ...
Tree transducers

- Tree transducer = rewriting mechanism for trees
- syntax-directed translation (cf. compiling)
- transfer parses into meanings
  (e.g. Abstract Meaning Representations)
- seq2seq = neural network counterpart for string transducers
- encoding trees for neural networks is really hard
Wrapping up

- Don’t get me wrong, Deep Learning can be useful.
- But **NNs are not a magic fix**.
- The symbolic methods are still good, and it’s where we’ll see true progress.
Don’t get me wrong, Deep Learning can be useful.

But **NNs are not a magic fix**.

The symbolic methods are still good, and it’s where we’ll see true progress.

Lesson 1

Pogo sticks don’t fly.
Don’t get me wrong, Deep Learning can be useful.
But **NNs are not a magic fix**.
The symbolic methods are still good, and it’s where we’ll see true progress.

Lesson 1
Pogo sticks don’t fly.

Lesson 2
Wasps are strangely stupid.