Computing Beyond Turing

How neocortical theory can lead to machine intelligence

Stanford Computer Forum
April 15, 2009
Jeff Hawkins
Numenta
What Does Cortex Do?

- People
- Cars
- Buildings
- Words
- Songs
- Ideas

World → Senses → Cortex

- light
- sound
- touch
- patterns
What Does Cortex Do?
- It builds a model of the world -

People
- Cars
- Buildings
- Words
- Songs
- Ideas

Senses
- light
- sound
- touch

Cortex
(cause1 0.22
cause2 0.07
cause3 0.00
cause4 0.63
cause5 0.08)

World
(cause)

Senses
Patterns
1 Discovers causes in the world
2 Infers causes of novel input
3 Predicts future
4 Creates motor behavior
- Knowledge is distributed hierarchically
- Self training through changing sensory patterns
- Each region is similar
- Knowledge is distributed hierarchically
- Self training through changing sensory patterns
- Each region is similar

Sensory data (skin)

Sensory data (retina)
Sensory data
Hierarchical Temporal Memory

Each node
- Learns common spatial patterns
- Learns common sequences

Sequence names passed up
Predicted spatial patterns passed down

Creates hierarchical model of causes
Bayesian methods resolve ambiguity
Belief Propagation In an HTM Node
## Belief Propagation Equations In an HTM Node

1) Calculate likelihood over coincidence-patterns.

\[
y_i(t) = P(-e_i|c_i(t)) \propto \prod_{j=1}^{M} \lambda^{r_{ij}^{(M)}}(r_{ij}^{(M)})
\]  

where coincidence-pattern \( c_i \) is the co-occurrence of \( r_{ij}^{(M)} \)'th Markov chain from child 1, \( r_{ij}^{(M)} \)'th Markov chain from child 2, \ldots, and \( r_{ij}^{(M)} \)'th Markov chain from child \( M \).

2) Feed-forward probability over Markov chains (temporal groups) using dynamic programming

\[
\lambda_t^{(g_r)}(g_r(t)) = P(-e_i|g_r(t)) \propto \sum_{c_{ij(t)} \in G^k} \alpha_t(c_{ij}, g_r)
\]

\[
\alpha_t(c_{ij}, g_r) = P(-e_i|c_{ij}(t)) \sum_{c_{ij(t-1)} \in G^k} P(c_{ij}(t)|c_{ij(t-1)}, g_r)\beta_{t-1}(c_{ij}, g_r)
\]

\[
\alpha_0(c_{ij}, g_r) = P(-e_i|c_{ij}(t = 0))P(c_{ij}(t = 0)|g_r)
\]

3) Calculate the belief distribution over coincidence patterns

\[
\text{Bel}_t(c_i) \propto \sum_{g_r \in G^k} P(g_r, g_0)\beta_t(c_i, g_r)
\]

\[
\beta_t(c_i, g_r) = P(-e_i|c_i(t)) \sum_{c_{ij(t-1)} \in G^k} P(c_{ij}(t)|c_{ij(t-1)}, g_r)\beta_{t-1}(c_{ij}, g_r)
\]

\[
\beta_0(c_i, g_r) = P(-e_i|c_i(t = 0))P(c_i(t = 0)|g_r, g_0)
\]

4) Calculate the messages to be sent to child nodes.

\[
\pi^{child} (g_m) \propto \sum_{c_i} I(c_i) \text{Bel}_t(c_i)
\]

where

\[
I(c_i) = \begin{cases} 
1, & \text{if } g_m^{child} \text{ is a component of } c_i \\
0, & \text{otherwise}
\end{cases}
\]
NuPIC, Numenta Platform for Intelligent Computing

2) Dev Tools
   - Configurator
   - Trainer
   - Debugger

1) Run time environment
   - Supervisor
   - Node Processor 1
   - Node Processor 2
   - Node Processor N
   - Gigabit switch
   - Fileserver

3) Learning Algorithms
   - Node learning algorithms

Additional information:
- Node learning algorithms
- NuPIC, Numenta Platform for Intelligent Computing
Pictures: Simple Vision System (32 x 32 Pixels)
Time Based Inference
Time Based Inference

Static inference (with noise)  19%
Time Based Inference

Static inference (with noise)

19%
Time Based Inference

Static inference (with noise) 19%
Time Based Inference

Static inference (with noise) 19%

Time based inference (with noise) 52%
Time Based Inference

Static inference (with noise) 19%

Time based inference (with noise) 52%
Time Based Inference

Static inference (static noise) 19%

Time based inference (with noise) 52%
Time Based Inference

Static inference (with noise)  19%

Time based inference (with noise)  52%

Time based inference (with dynamic noise)  40%
HTM Vision In Digital Pathology
- discriminate glands from other structures -

Glands

Not glands
Promising Early Results

• We trained a network to recognize glands
  – Training set: 195 images of glands and non-glands
  – Test set: 80 novel images
• Result: test set accuracy was 95%
• The four errors were “reasonable”
Illusory Contour
Illusory Contour
Illusory Contour

Perceived line with no bottom up support
HTM Application Areas

- Medical: Digital pathology
- Voice: Speaker/gender id
- Security: Video behavior recognition
- Auto: Lane change prediction
- Banking: Fraud detection
- Web: Analytics
- Pharma: Drug discovery
- Networks: Attacks and Failures
- Gaming: Motion capture inference, Visual object editor
- Semantic analysis of text
Technical Challenges

- Algorithms
  - Sequence Memory algorithms
  - Attention mechanisms
  - Feedback
  - Sensory pre-processing

- Performance
- Silicon vs. software
To Learn More

- Read “On Intelligence”
- Sign up for Numenta newsletter
- Download demo applications
- Download NuPIC
- Attend HTM workshop June 25, San Jose, CA
- Internships or employment at Numenta

- jhawkins@numenta.com