



# A New Mathematical Framework for Language Representation, Association, Processing, and Understanding

**Robert Hecht-Nielsen**

**Principal Investigator**

[r@hnc.com](mailto:r@hnc.com)

University of California, San Diego  
and

HNC Software Inc.  
5935 Cornerstone Court  
San Diego, CA 92121

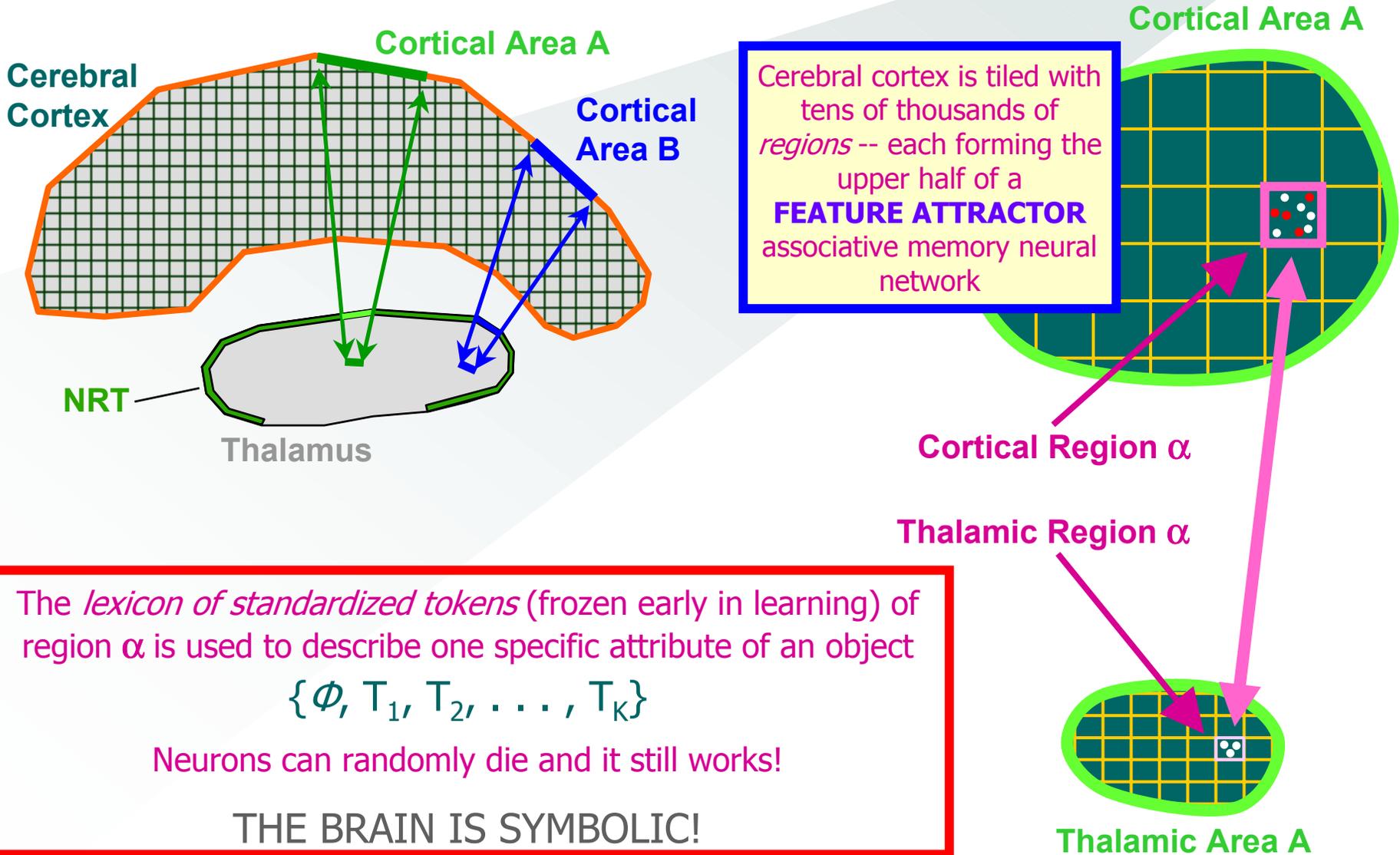
AQUAINT Phase I Mid-Year Workshop  
Monterey, California  
10 – 13 June 2002

# THE BIG QUESTIONS

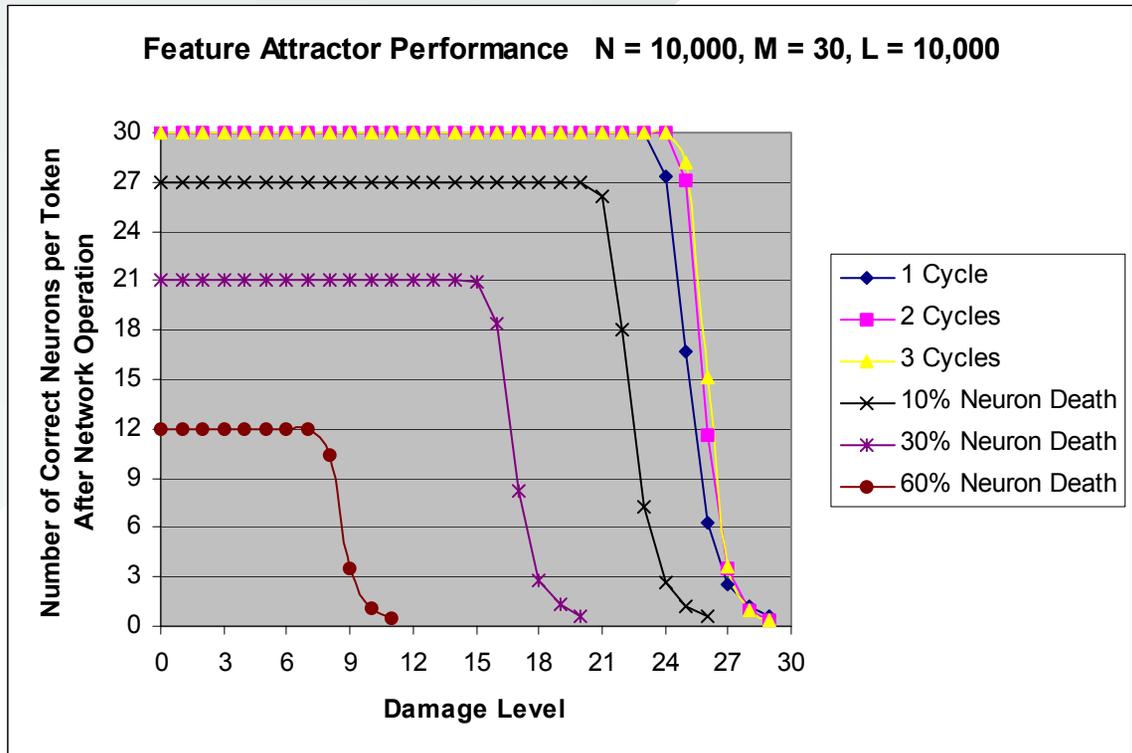
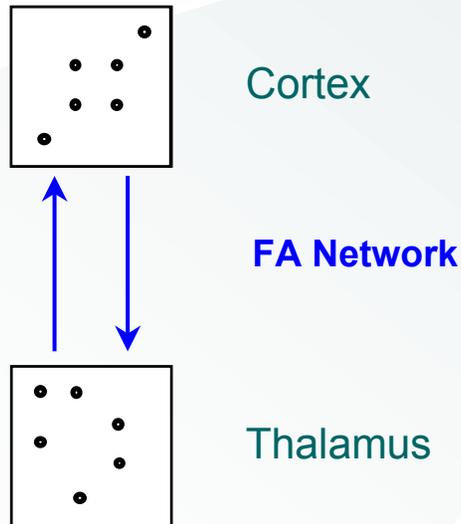
- How does thalamocortex represent objects?
- How does cerebral cortex learn and store information about objects?
- How does thalamocortex process information?
- In which brain nucleus does *Homunculina* reside?

And Now, The Answers!

# How Does Cerebral Cortex Represent Objects?

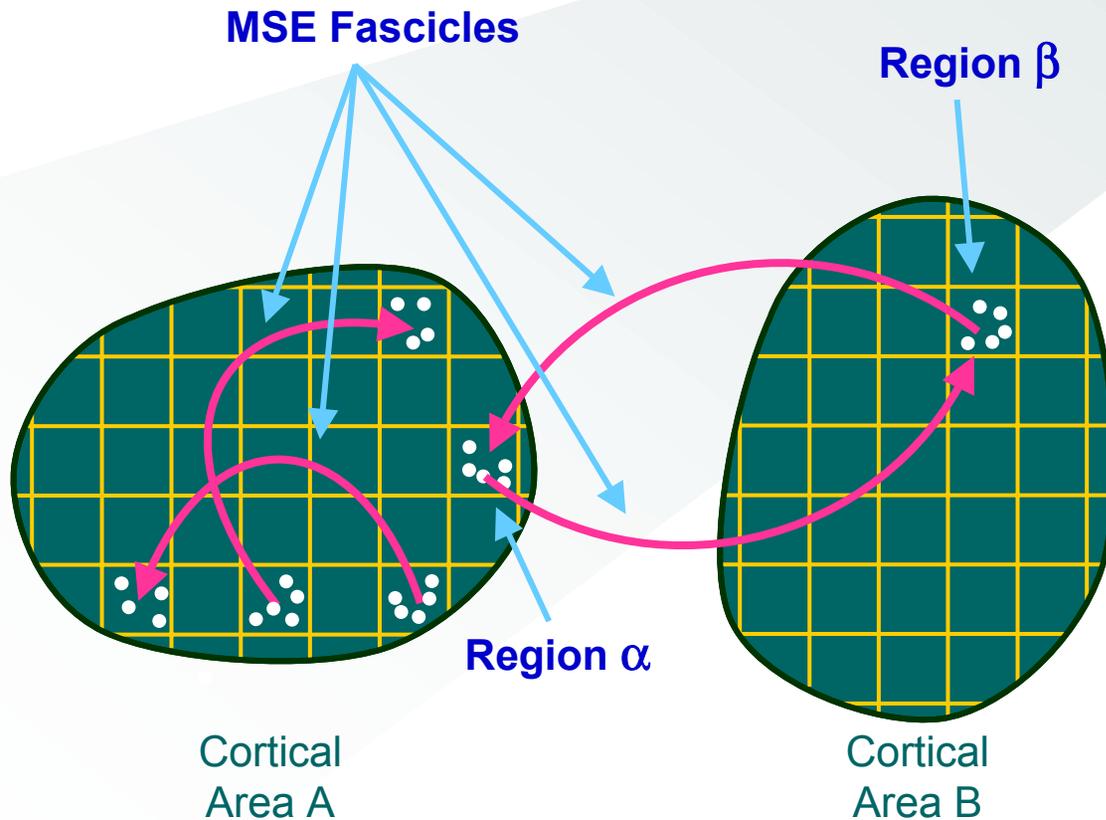


# Feature Attractor Network Performance



# How Does Cerebral Cortex Learn & Store Information?

## Mutual Significance Evaluator Associative Memory Neural Network



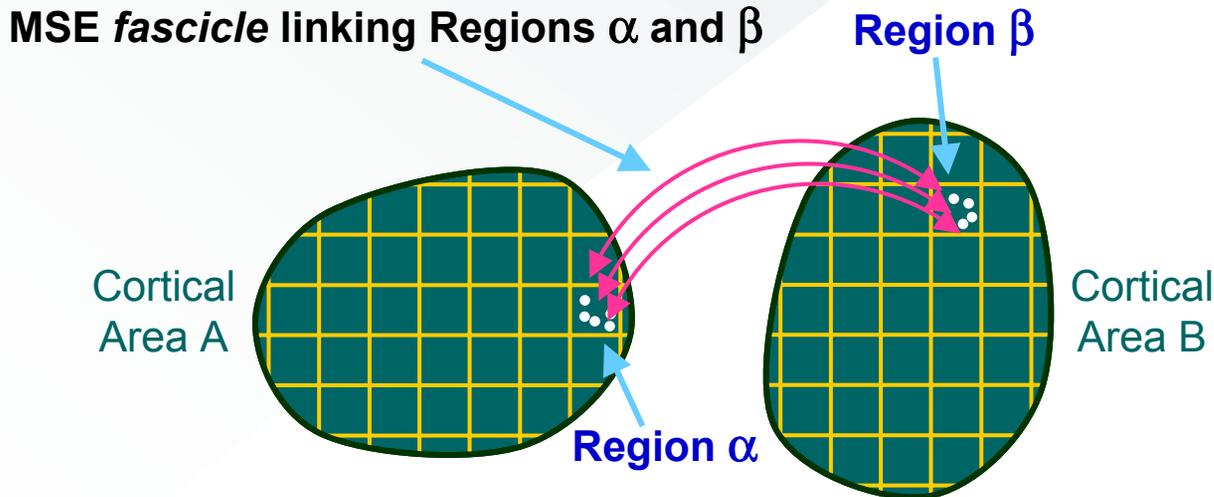
$s(i,j)$  is the *mutual significance* of tokens  $i$  and  $j$  on regions  $\alpha$  and  $\beta$ . It is the ratio of the actual probability of co-occurrence of tokens  $i$  and  $j$  to the probability that these tokens would co-occur randomly.

$$s(i, j) \cong \frac{p(i, j)}{p(i) p(j)}$$

MSE fascicles learn a monotonically increasing function of mutual significance, with low values set to zero. Thus, only a minute fraction of the token pairs of each MSE network need be connected.

# How Does Cerebral Cortex Learn & Store Information?

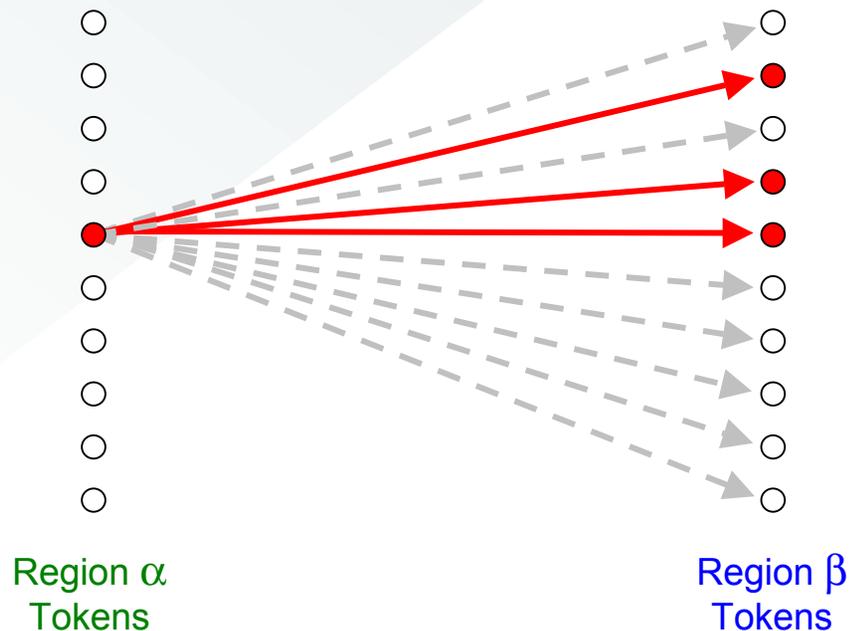
- Significance is a new information theoretic quantity. Human cortex has roughly a million mutual significance evaluator networks; each linking a pair of regions.
- Significance values are guaranteed to converge because they are a ratio of probabilities that depend only on the information environment (internal and external). If the information environment changes (or if a particular token pair is exercised repeatedly – as in memory consolidation during sleep), the significances change. Significance learning remains active throughout life.
- Token pair mutual significance is the only information learned and stored in the cerebral cortex.



Connections are between neurons of one token and neurons of a paired token. However, almost all of these connections learn the **same** mutual significance value. Thus, as with the feature attractor network, the mutual significance evaluator network is inherently tolerant of extensive random neuron death.

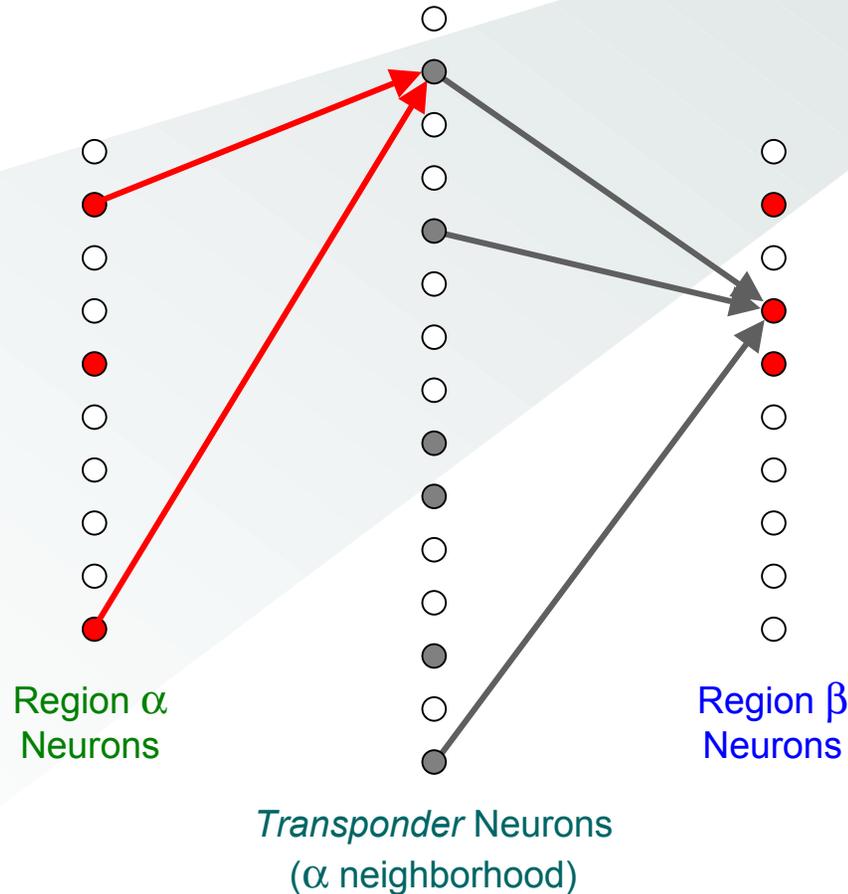
# Significance Separates the Wheat From the Chaff

- If two regions each have 10,000 tokens in their lexicons, then we might have to learn 100,000,000 pairwise token mutual significances in each direction.
- Fortunately, only a minute fraction of token pairs have mutual significances considerably above 1. These are the only connections we need implement.
- The physical implementation of these connections requires only a number of axons that grows roughly linearly with the actual number of high mutual significances (see next viewcell).
- It is the two preceding facts that allows our brain to fit within our small skulls. Otherwise, we would need a brain the size of a Lexus.



# Connections Between Cortical Region Neurons

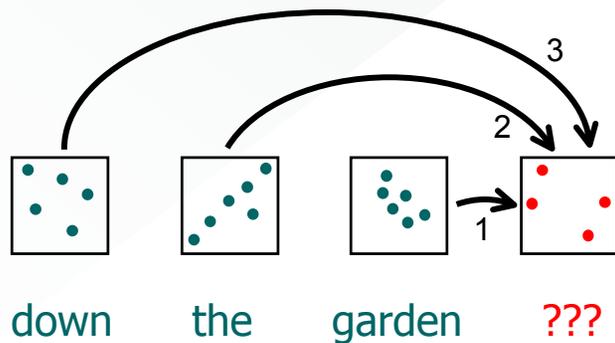
are implemented using intermediate populations of *transponder* neurons



This is like Hebb's *Cell Assemblies* or Abeles' *Synfire Chains*

# Consensus Building: The Universal Thalamocortical Information Processing Operation

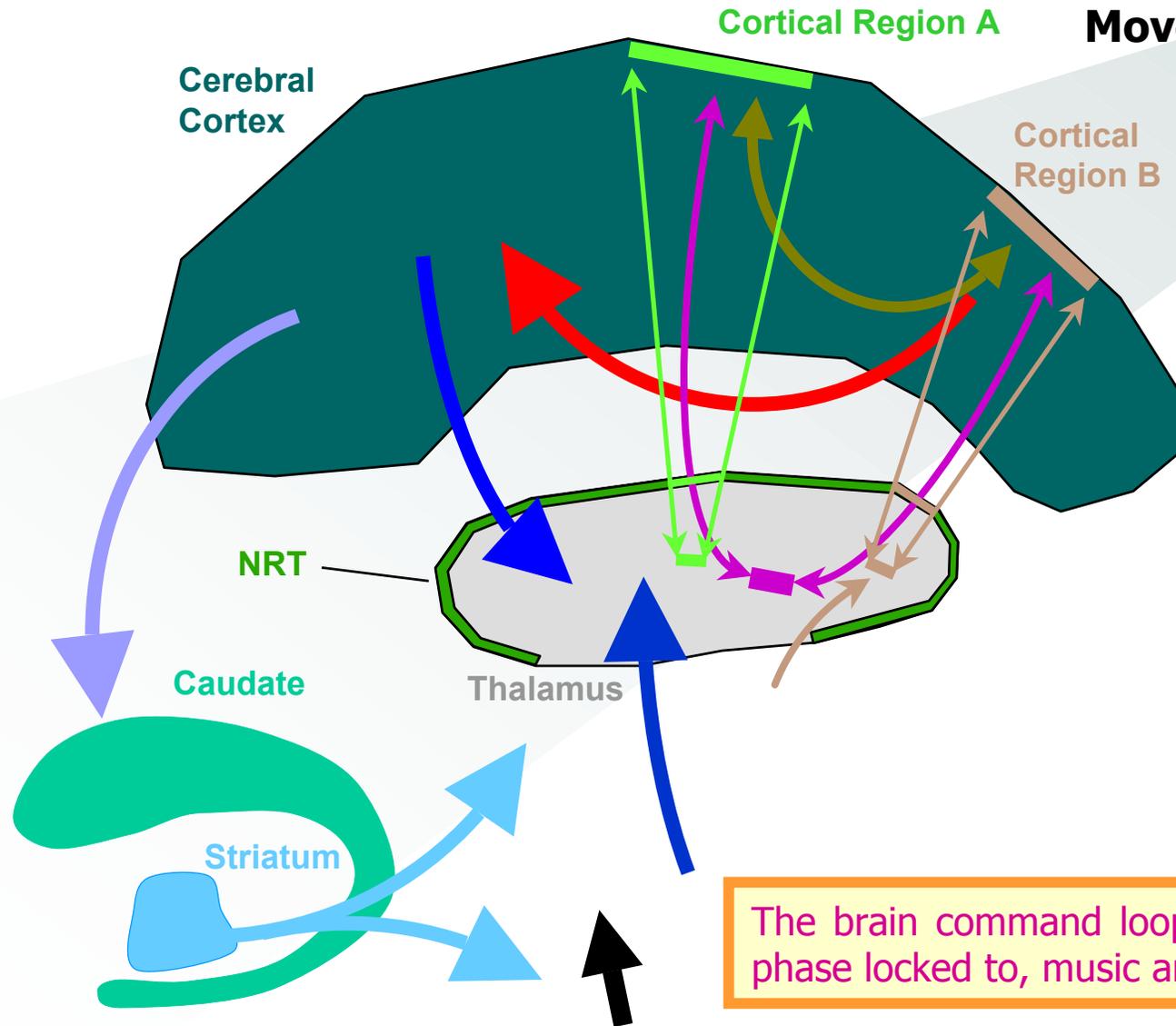
- The essence of consensus building is to find the answer token(s) that have the highest consistent level of significance with **all** of the specified known facts in terms of **all** of the specified knowledge bases (mutual significance evaluator fascicles) being employed.
- Human cognition thus always proceeds on the basis of finding specified types of answers which, when considered using a specified body of knowledge, are uniformly highly consistent (in terms of high mutual significance) with all the assumed facts. Any inconsistency causes failure of the cognitive process.
- Unanswerable questions are usually resolved by further restricting the set of considered facts and/or the body of knowledge being used.



$$s(i, j) \cong \frac{p(i, j)}{p(i) p(j)}$$

# Exorcising Homunculina

## Movements and Thoughts are Siblings

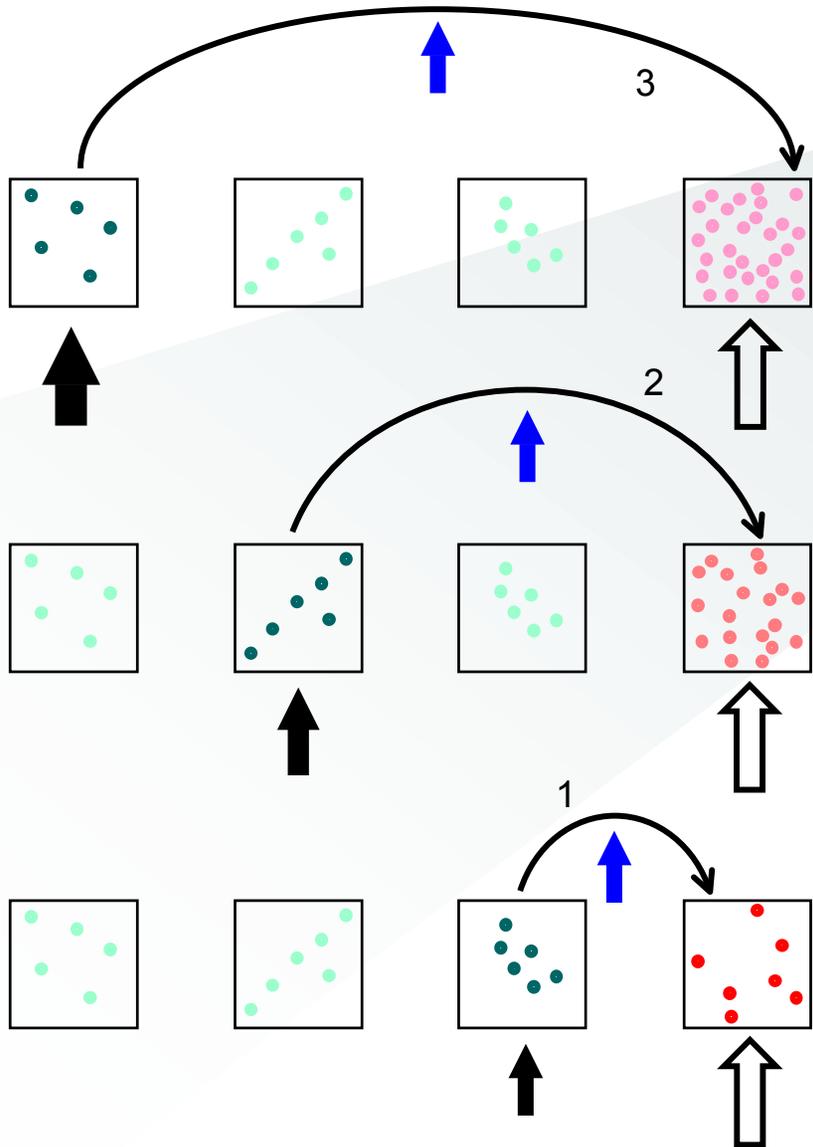


The world state (internal and external) is constantly being expressed across cortex. Layer 5 projections from primary motor cortex go to spinal motor nuclei. Layer 5 projections from the rest of cortex go to subcortical nuclei which control cortical feature attractor and mutual significance evaluator network activations (just as if they were muscles).

The brain command loop can be entrained by, and phase locked to, music and dance.

The *brain command loop* runs continuously during wakefulness

# Implementation of Consensus Building: *Honing*

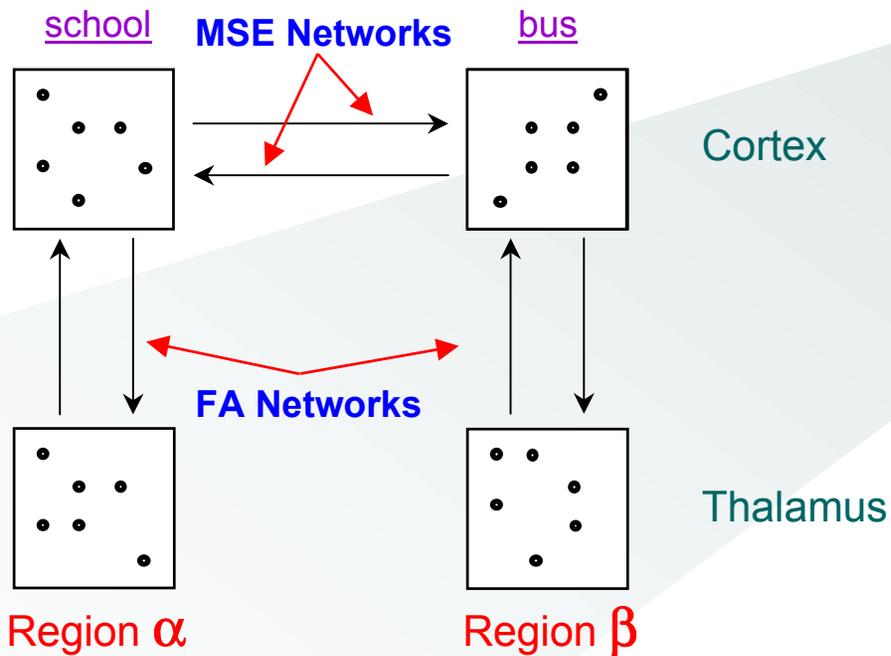


Consensus building involves *sequential activation* of groups of regions in accordance with a specified *precedence ordering*. At each stage of this sequence, each answer region is *honed*. Honing involves turning off (for the remainder of the CB operation) those neurons which do not receive significance input above a fixed minimum. Neurons which remain on have their activation level set to their new significance input value; except that this level can never increase; only decrease, during the CB process. Additional answers can be discerned by turning off answer neurons and repeating the CB process. The entire CB process typically requires only a few hundred milliseconds.

CB is a stored *thought process* (region and fascicle activation sequence) which is data independent. Such thought processes are learned, stored, and issued by cortex exactly like movements.

# CORTRONICS: A New Foundation for AI

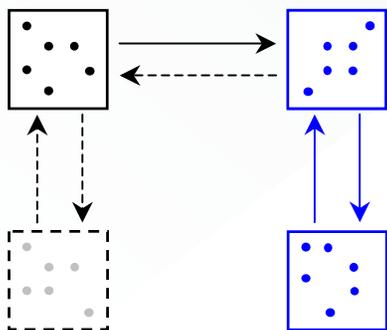
## Computer Implementations of Thalamocortical Neural Networks



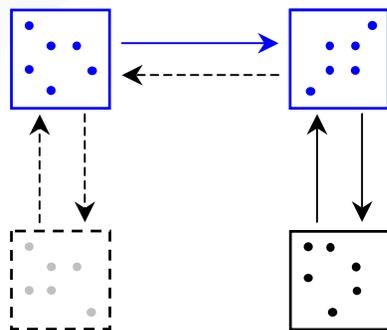
We have applied Cortronics to sound, vision, and English text. The examples presented here are for text. Experiments often use a diverse billion word training corpus equivalent to thousands of books (novels, encyclopedias, news, etc.). Many MSEs are trained in parallel during each training stage.

### Train Mutual

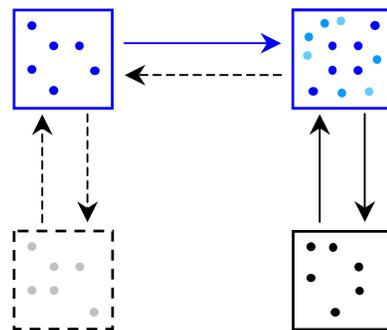
#### Build Feature Attractor



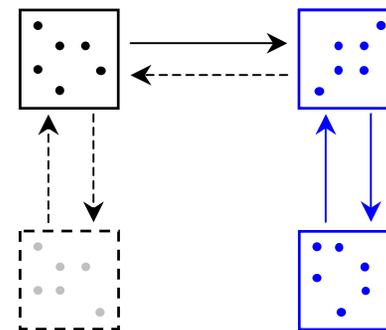
#### Significance Evaluator



#### Insert Word & Apply MSE



#### Apply Feature Attractor



# Consensus Building: Raw Knowledge Examples

## driving at top ?

driving: influence, expired, suspended  
at: university, time A.M., P.M.  
top: priority, aides, aid, executives

## man on the ?

man: woman, murdering, moon, himself  
on: #XXX (our universal numerical quantity)  
the: united, first, same, company

## all but a ?

all: sudden, bear, are, abroad  
but: said, don't, is, not  
a: lot, few, share, new

## the future looks ?

the: the, of, a, in  
future: Jerusalem, uncertain, the, of  
looks: like, forward, pretty, at

## Unites States of ?

United: Canada, Britain, Japan, India  
States: Canada, Britain, Japan, Europe  
of: the, a, its, his

## traveling around the ?

traveling: mph, speed, directions, unavailable  
around: world, country, clock, globe  
the: united, first, same, company

## all of a ?

all: sudden, bear, are, abroad  
of: and, million, #XXX, world's  
a: lot, few, share, new

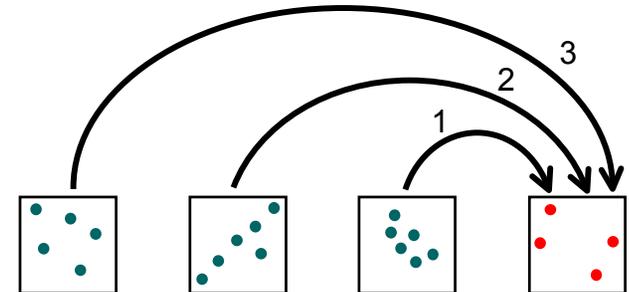
## deep blue ? (three-word example)

deep: cuts, in, the, divisions  
blue: chips, jeans, chip, shield

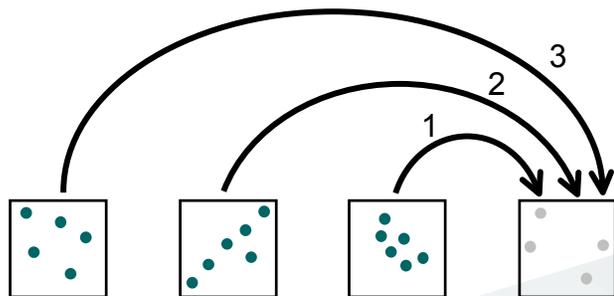
## down the garden ?

down: unchanged, #XXX, percent, from  
the: of, states, is, and  
garden: grove, city, centers, ceremony

This is THE classic AI problem: how to combine disparate fragments of 'knowledge' to yield a 'logical' answer.



# Consensus Building: Combining MSE Knowledge



MSE Fascicle 3: **driving**: influence, expired, suspended

MSE Fascicle 2: **at**: university, time, A.M., P.M.

MSE Fascicle 1: **top**: priority, aides, aid, executives

## Assumed Facts

## Consensus Answers

(from successive CB operations)

driving	at	top	speed
traveling	around	the	country, village, airport, town
man	on	the	moon, bench, street
all	of	a	sudden, common, handful
all	but	a	handful, sudden, remarkable
all	over	the	map, country, globe
	deep	blue	sky, eyes, paint, waters, flowers
the	future	looks	promising, bright
down	the	garden	path, lane

instructive mistake (a three-way tie):

United States of Canada, Japan, America

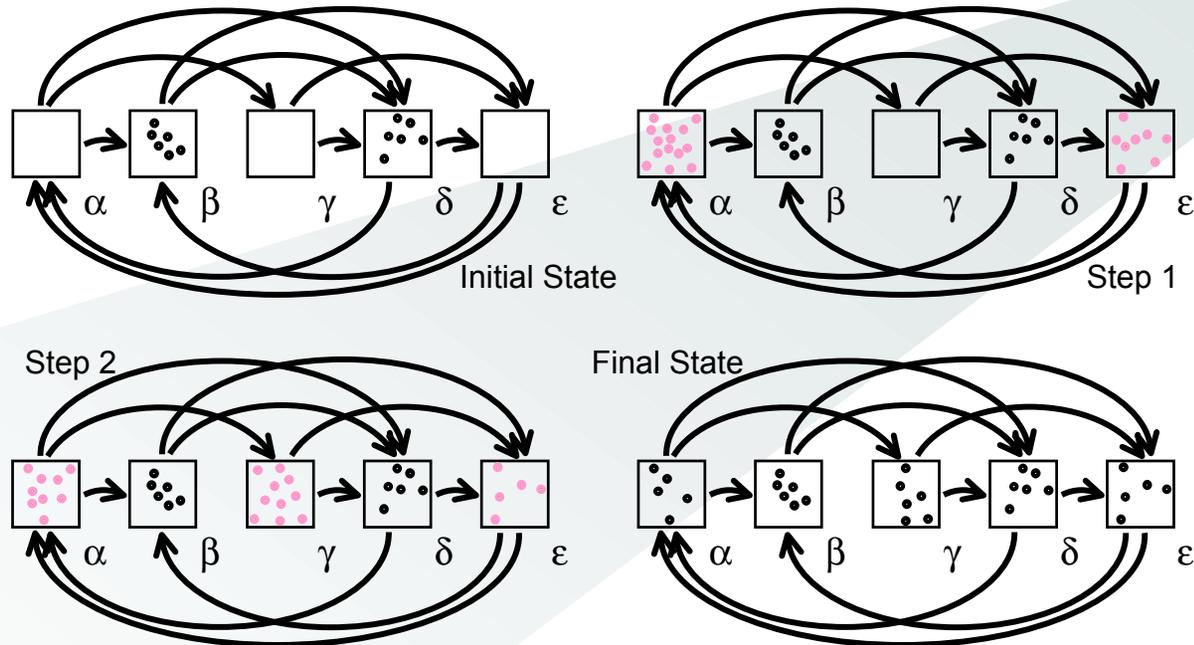
### legend

red =  
phrase  
not in  
training  
database

blue =  
phrase in  
training  
database

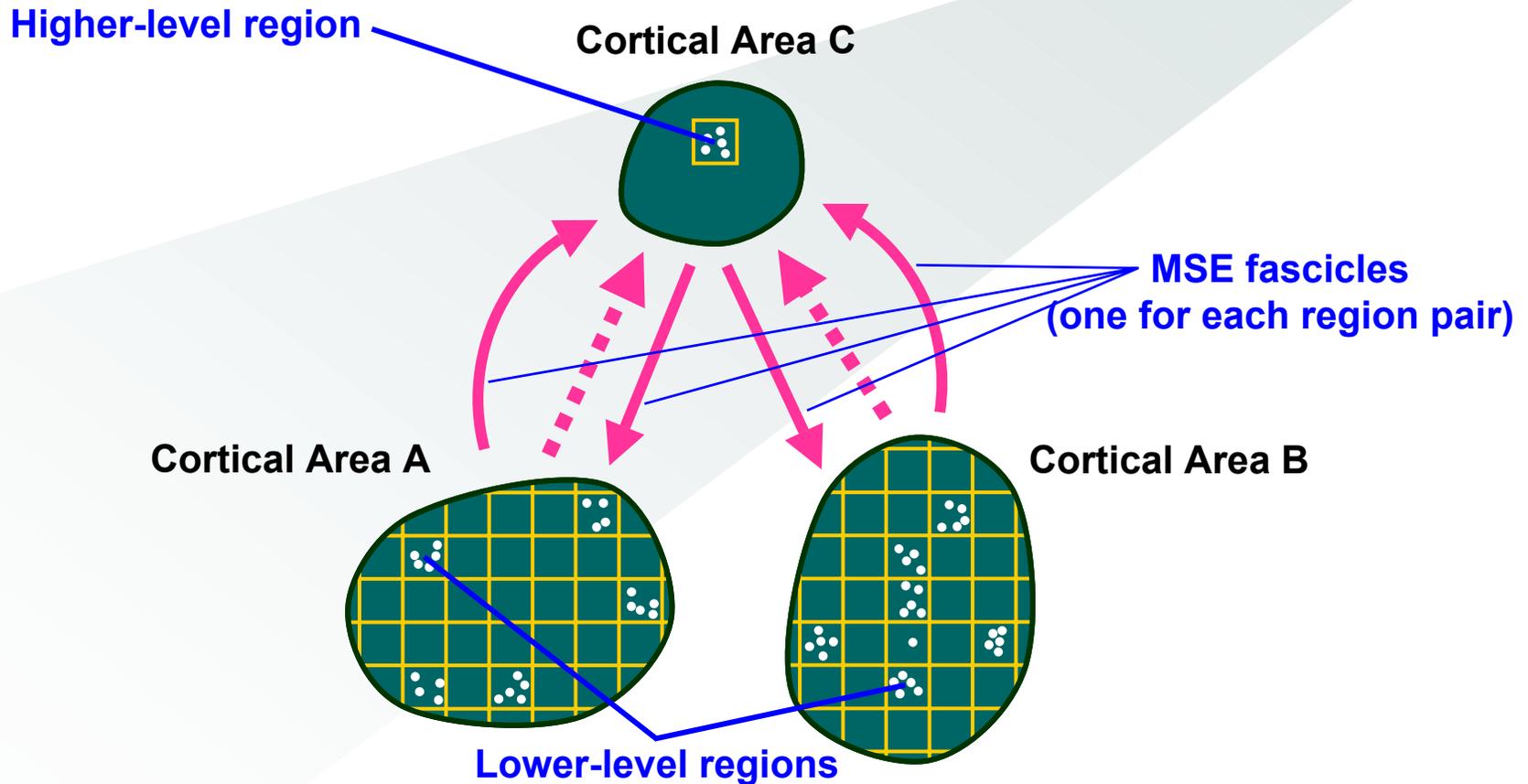
CB is the Holy Grail of AI. Savor your first drink from it.

# How Does Cerebral Cortex Process Information?



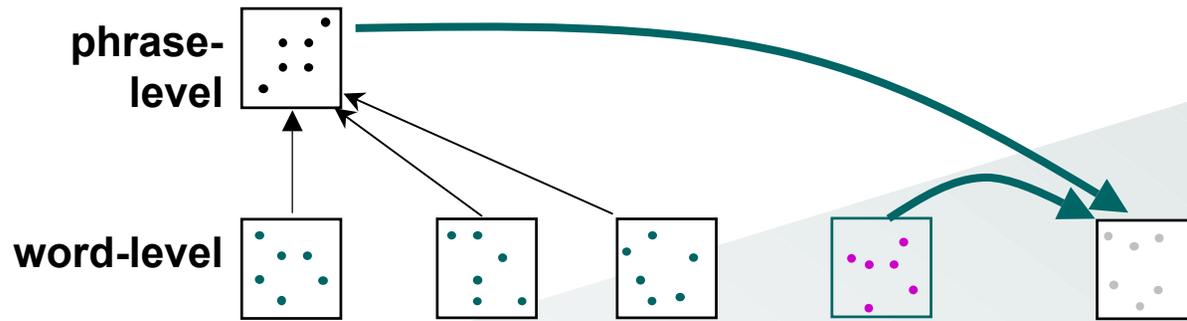
*Consensus building* involves five elements: a selected set of *constraint regions* (each expressing the token of an *assumed fact*), a selected set of *answer regions* (these are blank – this is where the answers – if any there be -- will appear), and a selected set of mutual significance evaluator networks (each considered an individual *knowledge base*), and an *operation precedence ordering*. Answers obtained by consensus building are logically consistent with ALL the assumed facts in the context of ALL the applied knowledge. Answer blanking and repetition of CB may yield additional valid answers. Consensus building is all there is.

# Hierarchical Abstractors are a Type of Derived Network



*Hierarchical abstractor* networks are built up from converging MSE networks and their FAs. Random upward connections are used when new tokens need to be formed.

# Consensus Building: Combining Disparate Knowledge

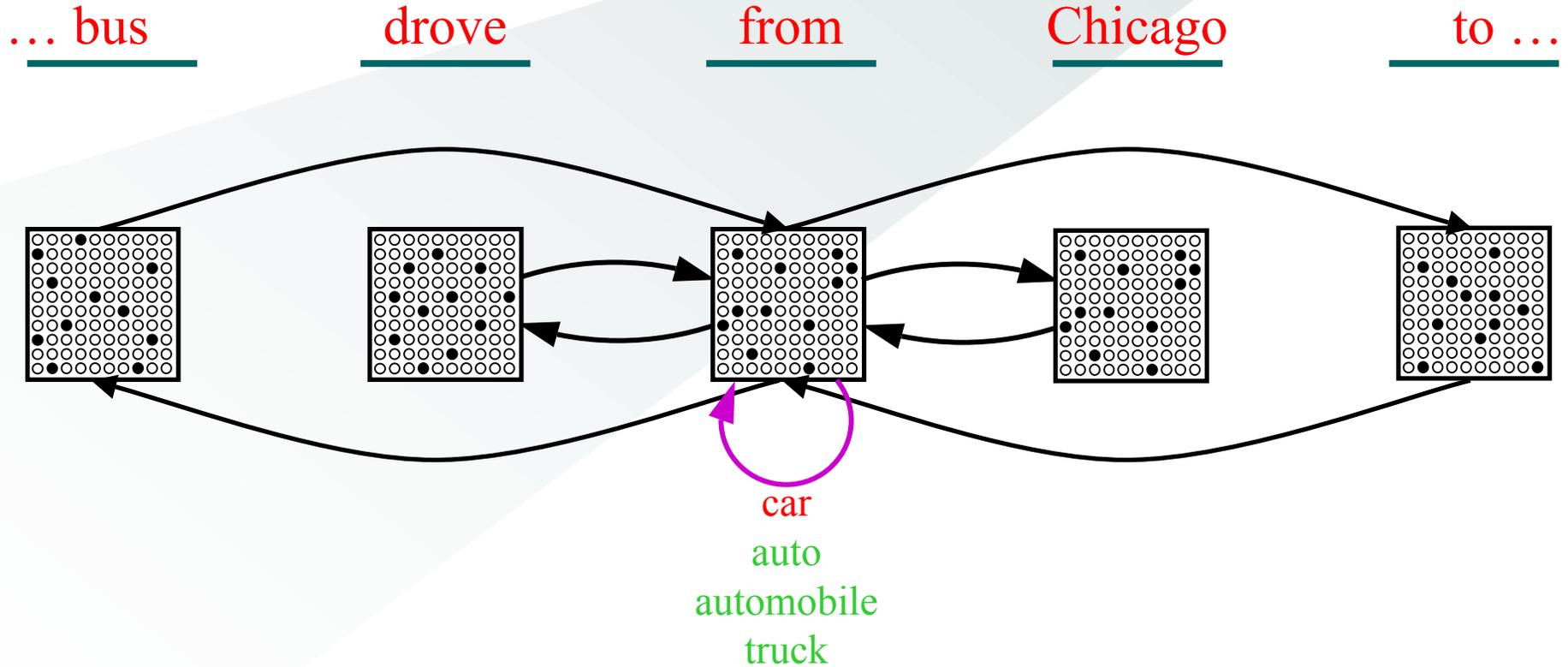


Consensus building liberates us from wimpy Markov models, Bayesian networks, N-grams, LSI, context vectors, etc. Now context of arbitrary type and quantity can be brought to bear on segmentations, classifications, and predictions.

the	wishes	of	the	school	board
spoke	on	condition		of	anonymity
##XXX	degrees			below	zero
former	Prime	Minister		Margaret	Thatcher
former	President			Ronald	Reagan
according	to	the		school	administrators, records, record
the	Dow	Jones		Industrial	Average
the	Dow	Jones		Average	rose, fell, climbed
the	destruction	of		wildlife	habitat
##XXX	on	the		Richter	scale
living	in			New	York, Jersey, Delhi, Orleans
comes	in			early	afternoon, trading, September, Spring
based	on			new	evidence
according	to			new	laws, rules
earlier	in			the	day, week, evening, session
New	York			Stock	Exchange, market, index

Yes, this is really it.  
CB is the mechanism  
of all human thought.

# Learning Word Level Synonymy



# Word Synonymy Lists

John

Jack

Ivan

Joseph

George

Paul

Peter

Daniel

Edward

Mary

Betty

Elisabeth

Sue

Sally

Karen

Jackie

Margaret

Lisa

Detroit

Chicago

Sydney

Seattle

Phoenix

Manila

Cincinnati

Minneapolis

Vancouver

France

Portugal

China

Singapore

Haiti

Cuba

Spain

Chile

Japan

phone

telephone

telecommunications

communications

ninety

twenty

thirty

eighty

minutes

years

days

pounds

TV

television

radio

video

television

TV

radio

news

see

hear

be

get

car

vehicle

truck

jeep

college

school

university

graduate

automobile

automaker

automotive

auto

seen

got

done

been

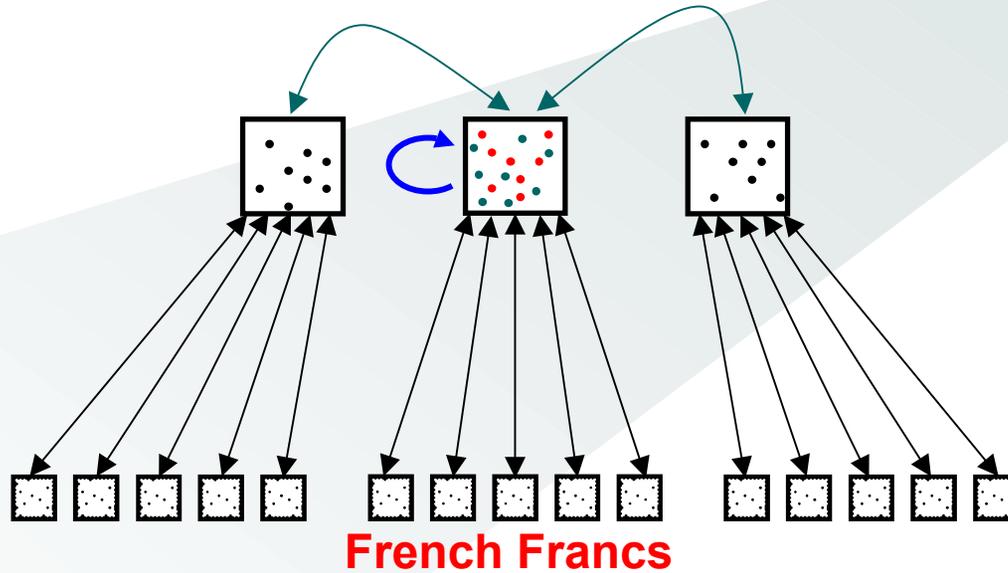
president

presidents

leader

chairman

# Phrase Level Synonymy Lists



**French Francs**  
German Marks  
Italian Lira  
Swiss Francs  
Canadian Dollars

**will be**  
should be  
have been  
are  
can be

**New York**  
Los Angeles  
Seattle  
New Orleans  
Chicago

# Welcome to Our Laboratory

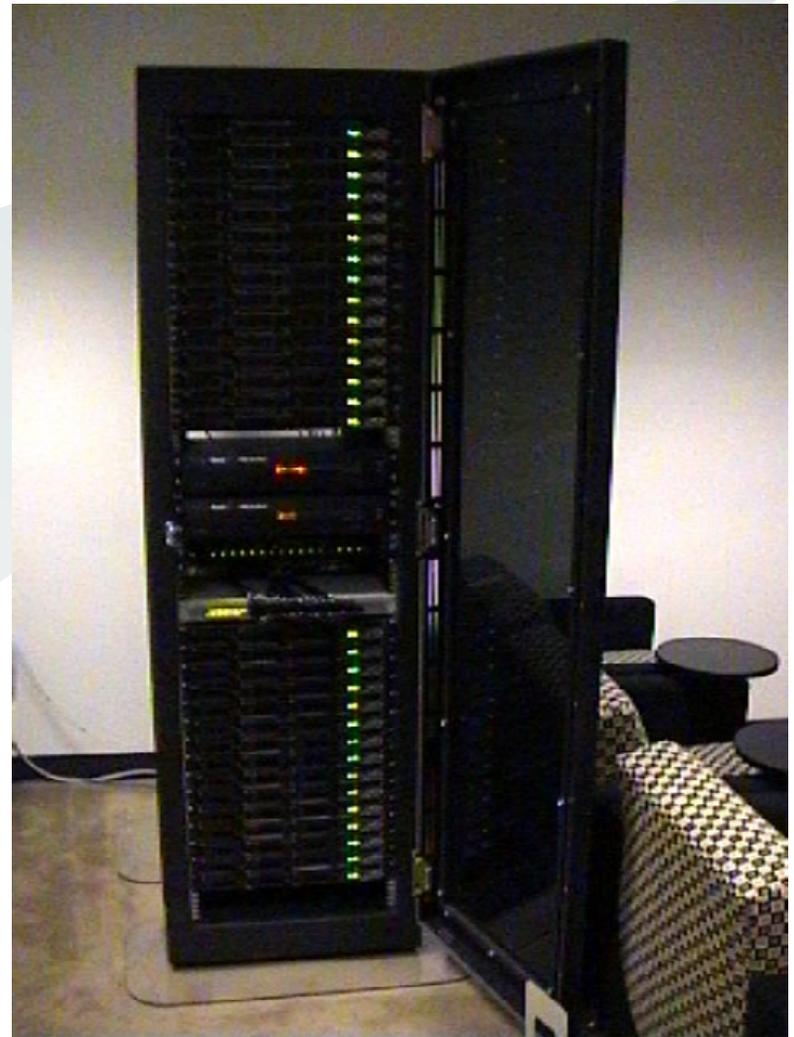


## BRAIN 1

24 Pentium Processors  
48 GBytes RAM  
1 Gbit/s Optical Ethernet

## HUMAN THALAMOCORTEX

25 Billion 1kHz Neurons  
500 GBytes Information Storage



## BRAIN 2

30 1GHz Pentium Processors  
60 GBytes RAM  
1 Gbit/s Optical Ethernet

# An Abundance of Immediate Applications

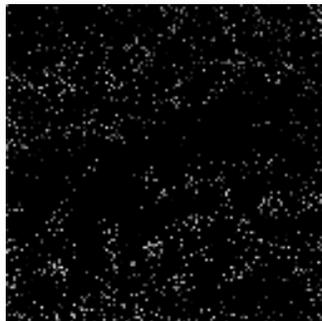
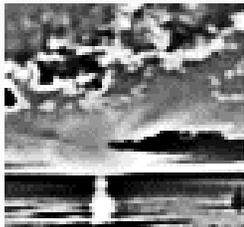
## Language Understanding

Modern Bligh set adrift Tuesday off Roosevelt Roads.

## Speech Recognition



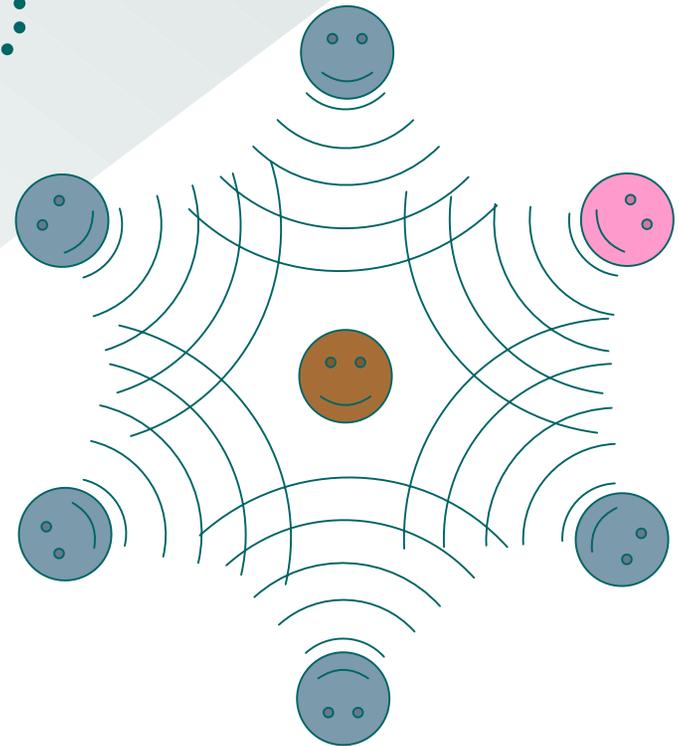
## Machine Vision



## Control

Sparse image coding  
example courtesy  
Dr. Joseph Sirosh of HNC

## Cocktail Party Problem

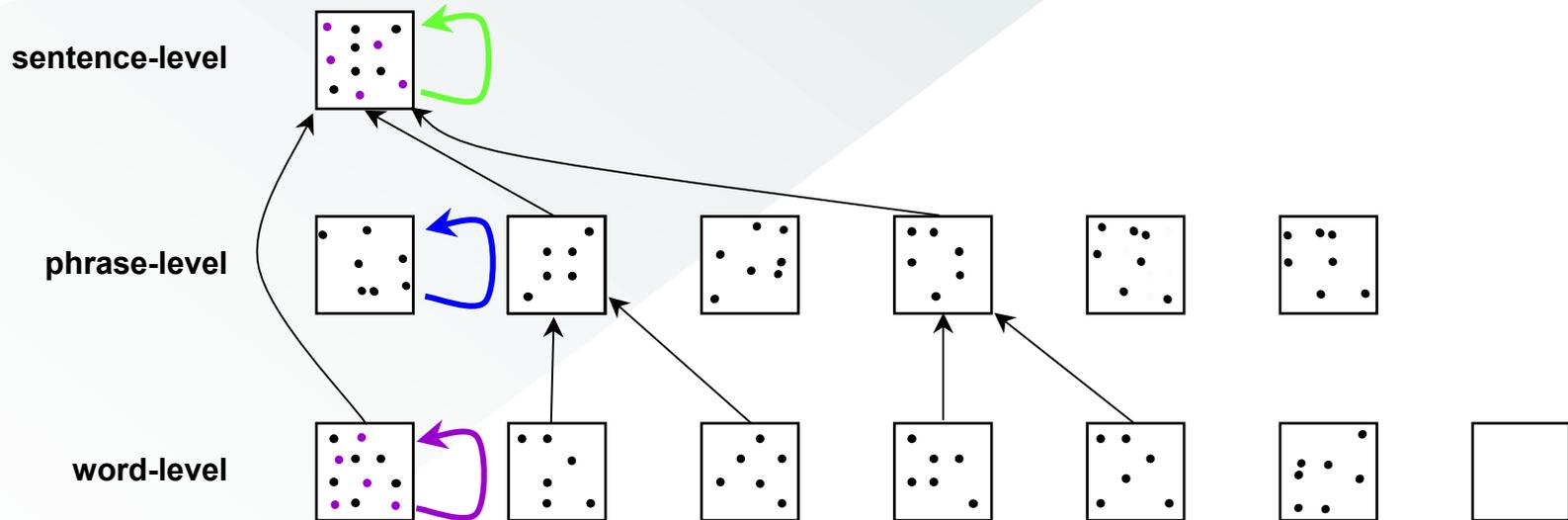


Sagi, B., et al (2001) A biologically motivated solution to the Cocktail Party Problem  
*Neural Computation*, **13**, 1575-1602.

# Question Answering Approach Sketch

Jane works for ??? (A *question answering* sentence)

Motorola employs Jane. (A *database* sentence)



Consensus building operations, learned from examples, can be used with a sentence representation structure like this to logically compare the meaning of one sentence with another (even if one has missing elements).

# The HNC AQUAINT Project Team

Dr. Robert Means -- Chief Technologist

Kate Mark - Project Coordinator

David Busby - Chief Brain Software Architect

Rion Snow - Researcher

Dr. Syrus Nemat-Nasser, Consultant

Christopher Downing, Consultant