The Fourth Conference on Artificial General Intelligence

AGI Research Progress in Intelligence Science Lab at Chinese Academy of Sciences

Zhongzhi Shi
shizz@ics.ict.ac.cn

Institute of Computing Technology
Chinese Academy of Sciences
http://www.intsci.ac.cn

INSTITUTE OF COMPUTING TECHNOLOGY 中科院计算所

Acknowledgement

- **Key Projects of National Natural Science Foundation of China:** No.61035003, 60933004
- **National Natural Science Foundation of China:** No. 61072085, 60903141, 60970088
- **China National Basic Research Priorities Programme 973:** No. 2007CB311004

Contents Outline

INSTITUTE OF COMPUTING TECHNOLOGY

- Introduction
- Mind Model CAM
- Feature Binding
- Image Understanding
- World Wide Brain
- Perspectives

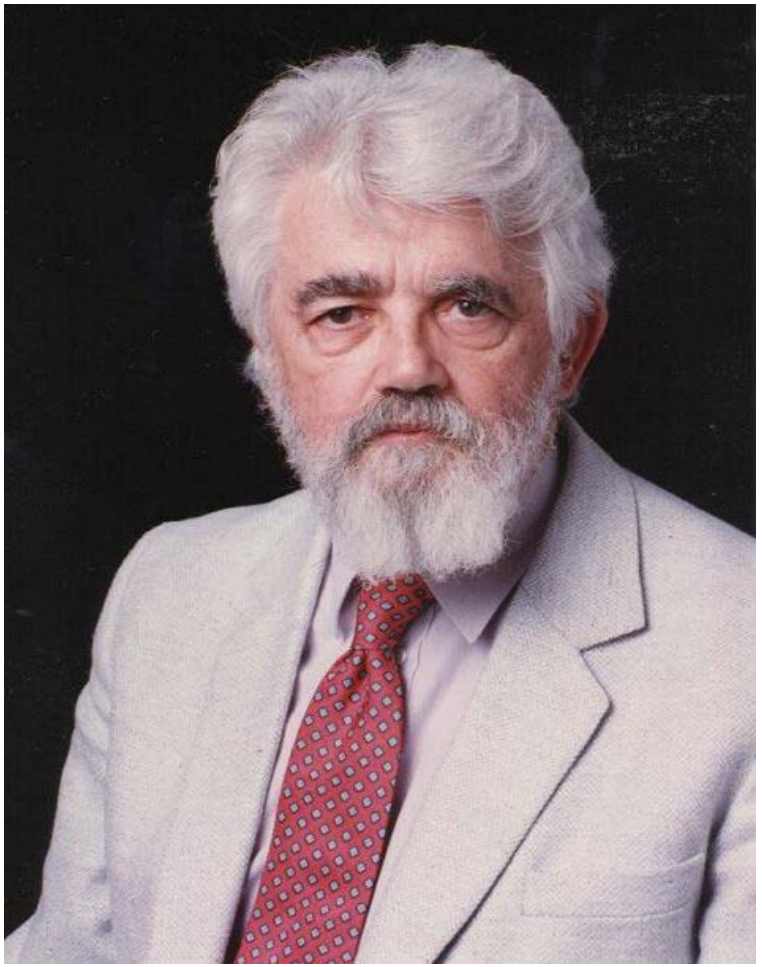
Introduction

- McCarthy declared

  the long-term goal of Artificial Intelligence is human level Artificial Intelligence

- It is not surprising that human-level AI has proved difficult and progress has been slow

John McCarthy. From here to human-level AI. Artificial Intelligence 171 (2007) 1174–1182



Artificial General Intelligence (AGI)

INSTITUTE OF COMPUTING TECHNOLOGY

"The ability to achieve complex goals in complex environments using limited computational resources"

- Autonomy
- Practical understanding of self and others
- Solving problems qualitatively different from those anticipated by the programmers

---Ben Goertzel

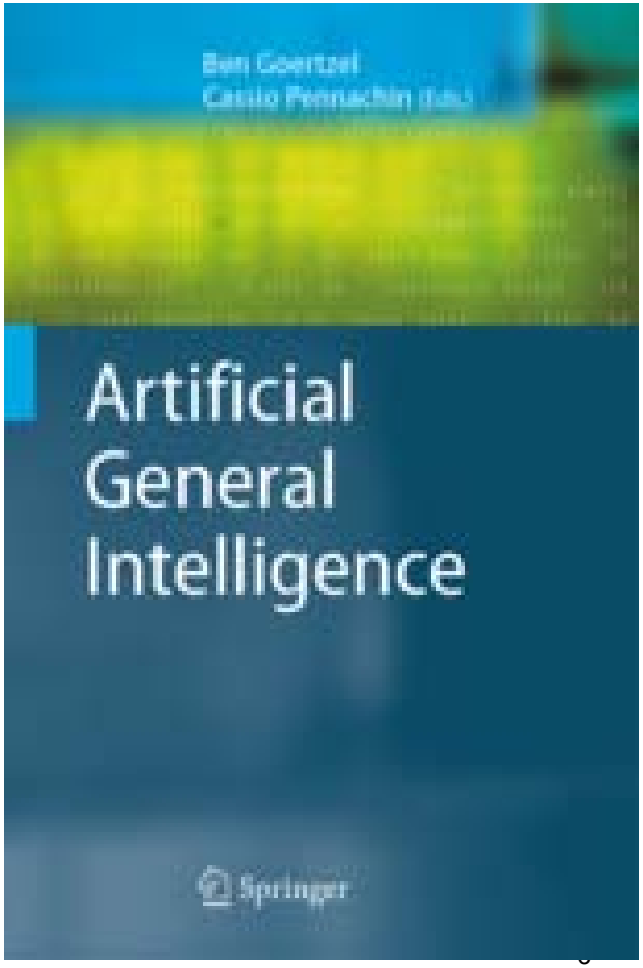
Artificial General Intelligence (AGI)

Book

Editor» Goertzel, Ben (et al.)

Artificial General Intelligence. Springer, 2007

This is the first book on current research on artificial general intelligence (AGI), work explicitly focused on engineering general intelligence – autonomous, self-reflective, self-improving, commonsensical intelligence.



Artificial General Intelligence (AGI)

INSTITUTE OF COMPUTING TECHNOLOGY

- AGI-2010: Lugano, Switzerland, March 5-8 (In Memoriam Ray Solomonoff)
- AGI-2009: Arlington, Virginia, March 6-9
- AGI-2008: University of Memphis, March 1-3

Artificial General Intelligence (AGI)

INSTITUTE OF COMPUTING TECHNOLOGY

A Roadmap Toward Human Level Artificial General Intelligence

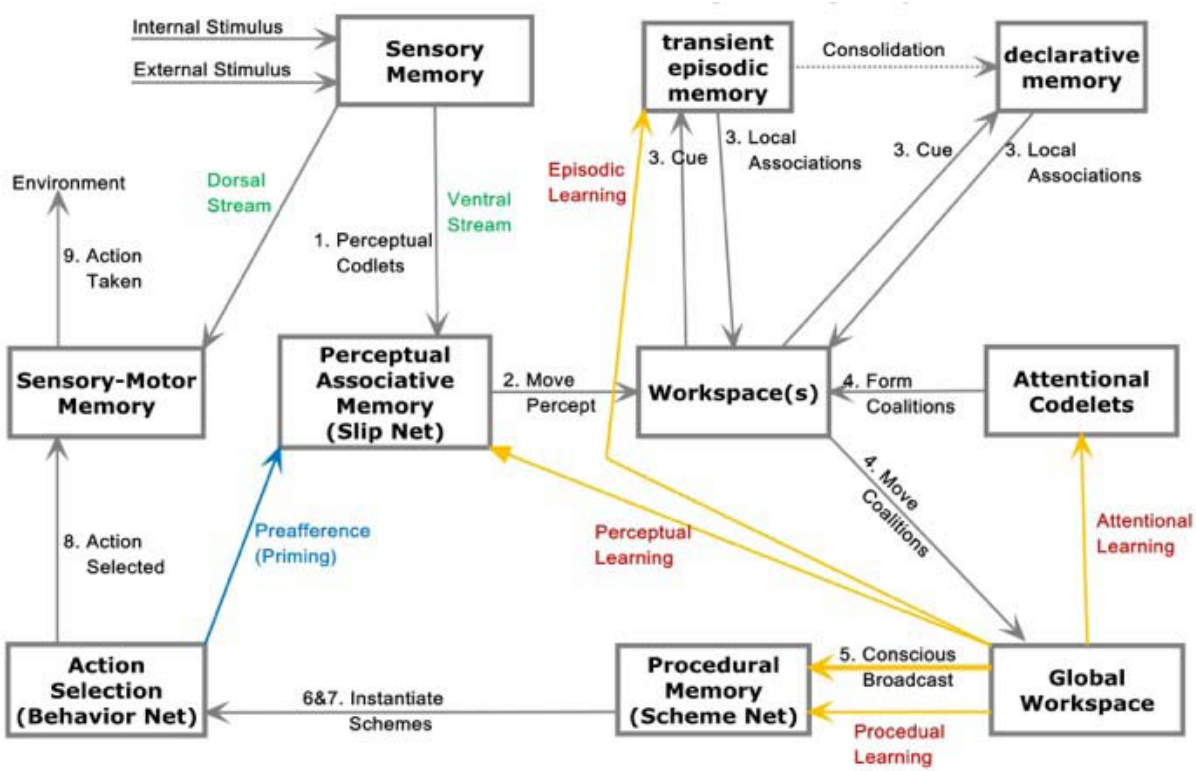
Sam Adams, Itamar Arel, Joscha Bach, Robert Coop, Rod Furlan, Ben Goertzel, J. Storrs Hall, Alexei Samsonovich, Matthias Scheutz, Matt Schlesinger, Stuart Shapiro, John Sowa

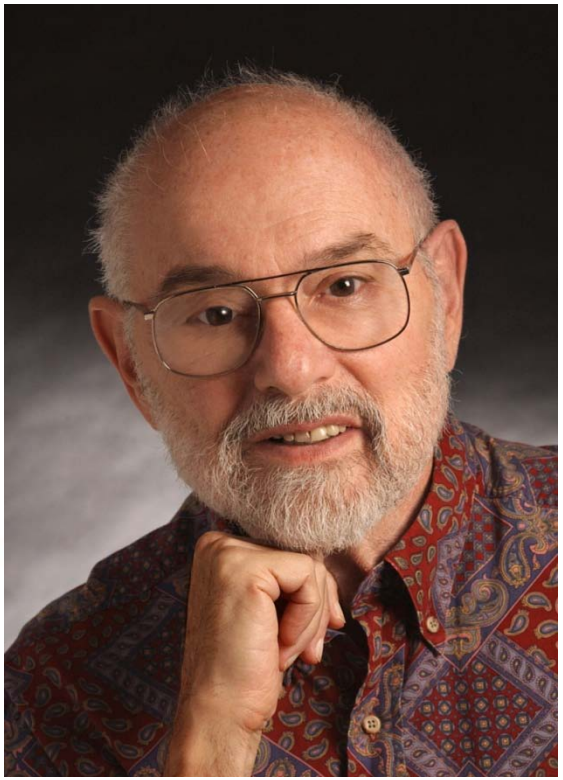
June 25, 2010

Abstract

We present the broad outlines of a roadmap toward Human Level artificial general intelligence (AGI). The roadmap is not targeted at any specific design or architecture, but is rather intended to encourage the development of multiple AGI systems along the same pathways, toward the same end goals and passing through the same set of milestones. A number of specific "scenarios" for the development of Human Level AGI are considered (including virtual-world, robotic and purely text/image-based "school" scenarios at various levels, along with others involving real-world robotics and video-game learning), along with a set of competency areas that may be explored in multiple scenarios and that together constitute a large fraction of Human Level intelligent functionality. It is proposed to develop a series of concrete tasks within each of the described scenarios, implementing the described competency areas. Such task series would then serve as a concrete, detailed roadmap to Human Level AGI for researchers accepting the given scenario as a useful one.

Framework of LIDA

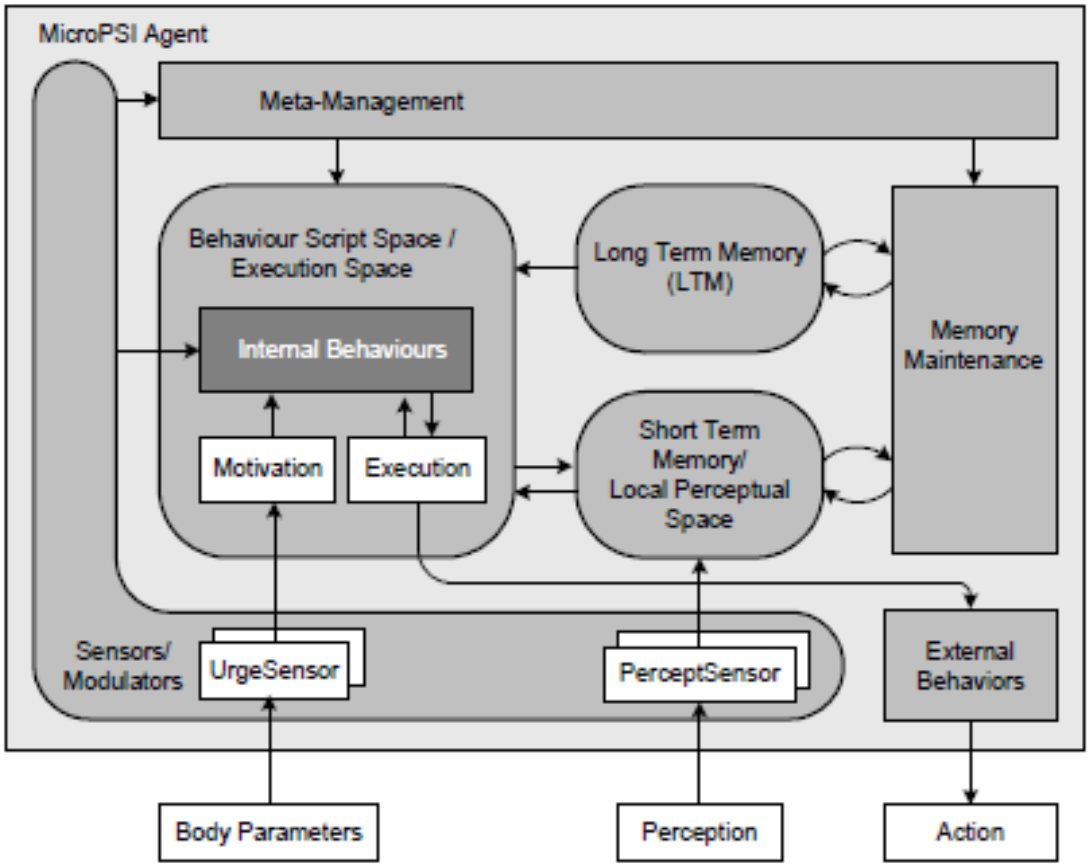




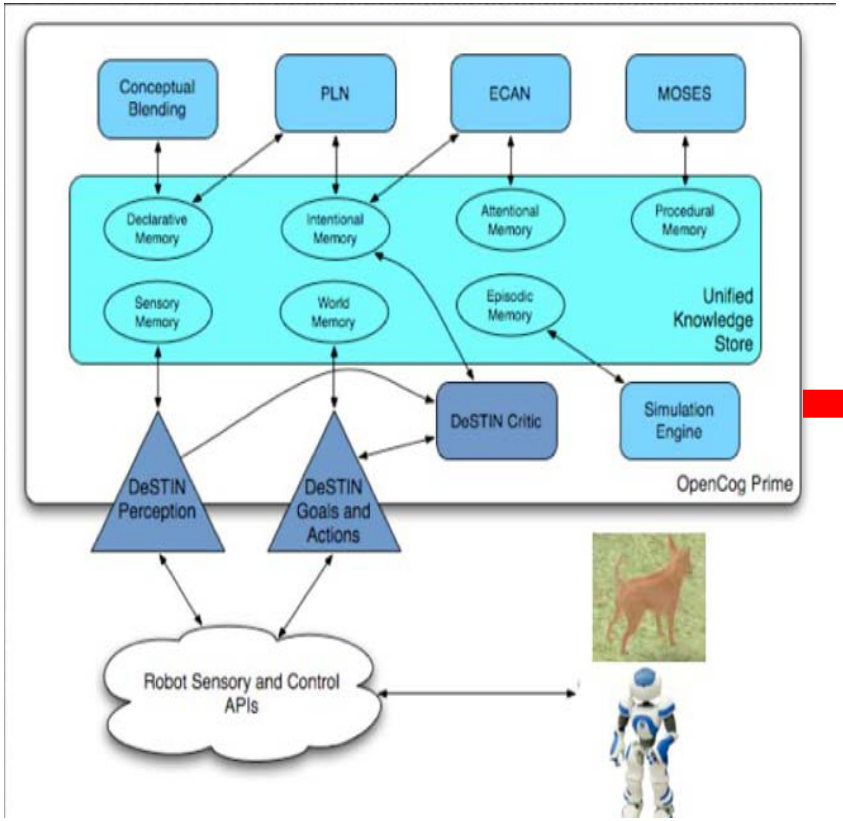
Javier Snaider, Ryan Mccall and Stan Franklin. The LIDA Framework as a General Tool for AGI. AGI2011

MicroPsi Architecture

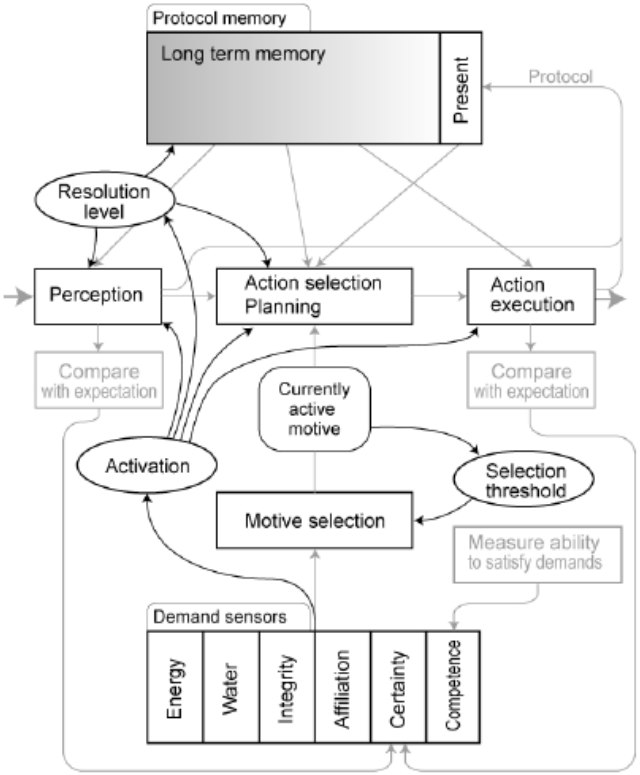
Dr. Joscha Bach



OpenPsi



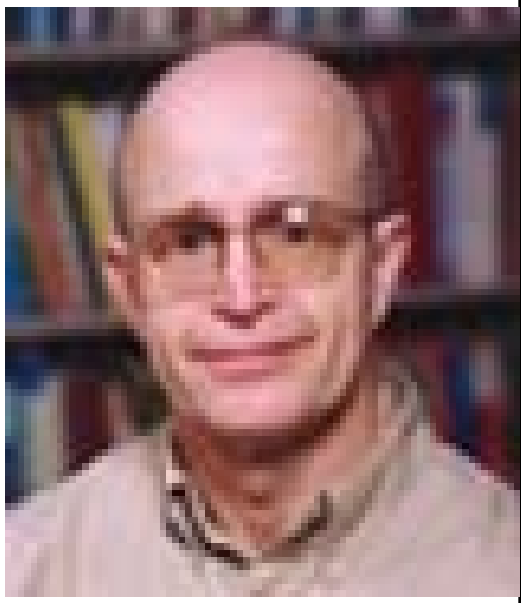
OpenCog



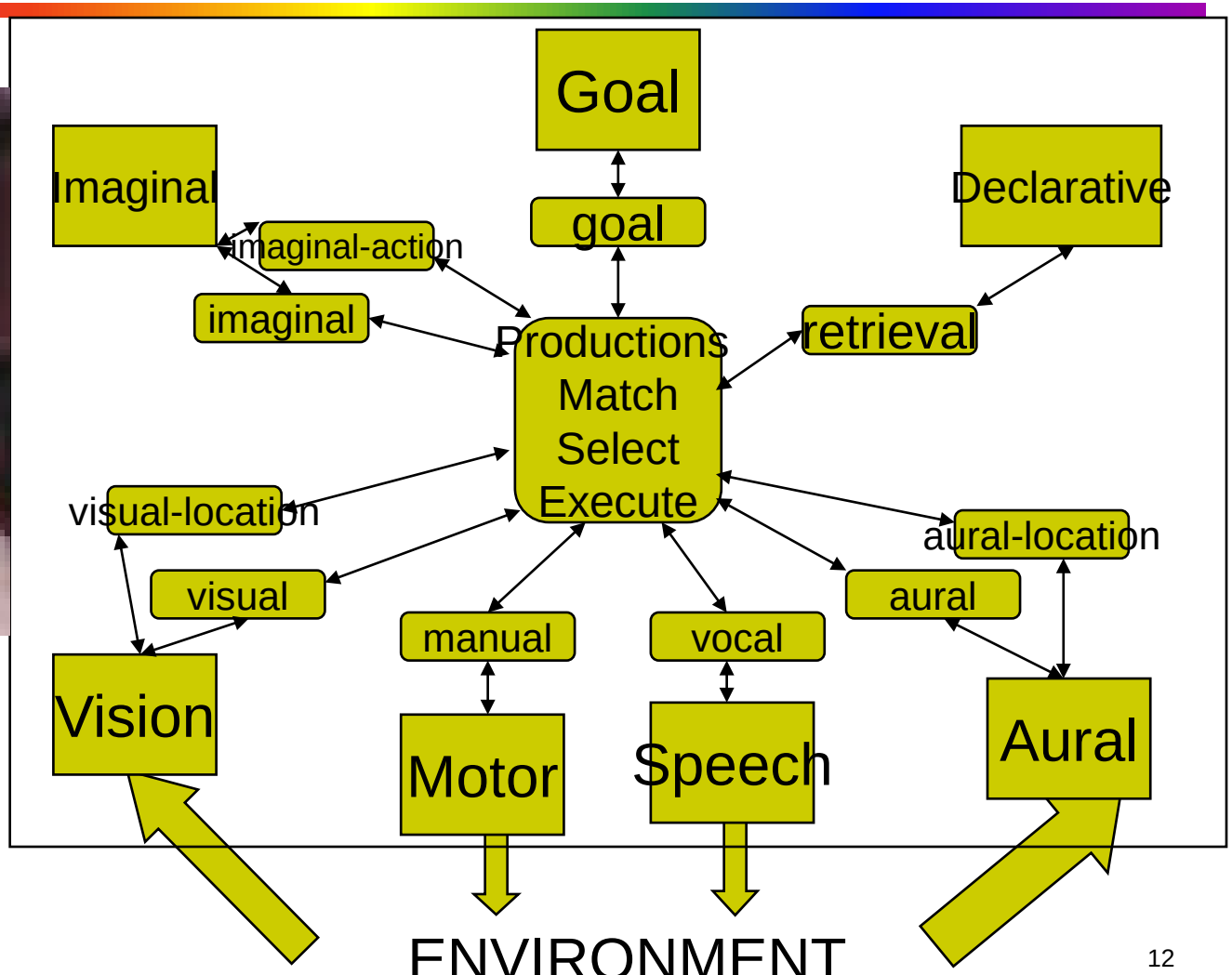
OpenPsi

Dr. Ben Goertzel et al.

ACT-R 6.0

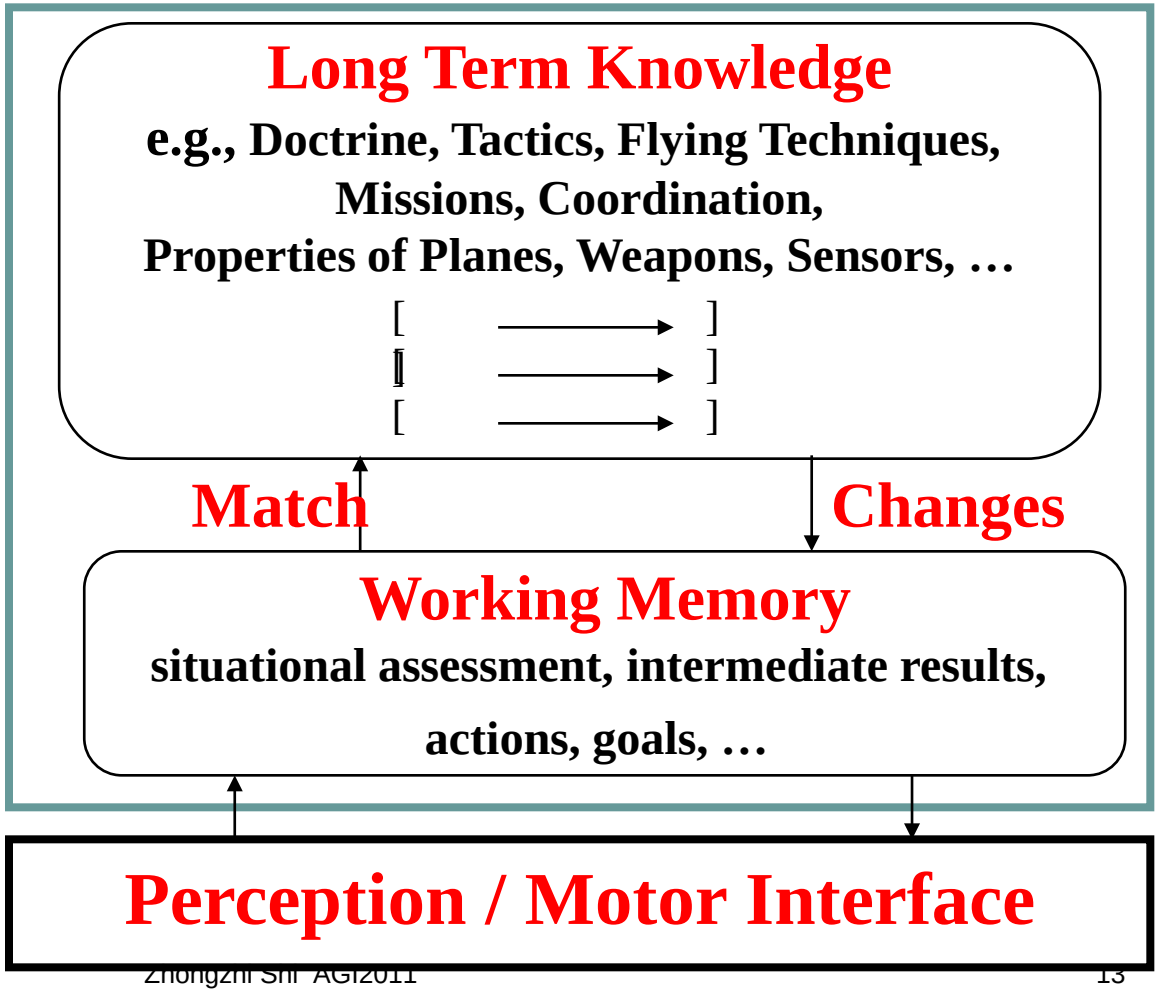


J.A. Anderson



Soar Architecture

Allen Newell

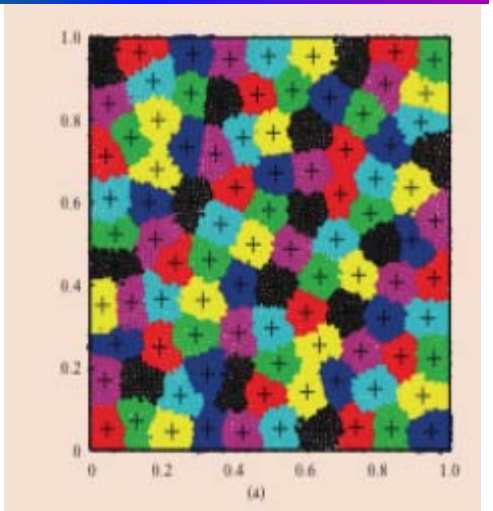


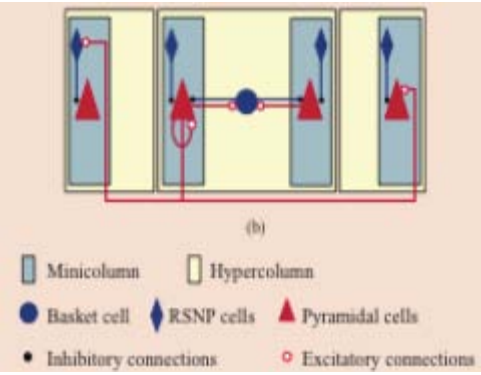
IBM Blue Brain Project

IBM Blue Brain Project(launched in 2005)

- IBM “cat brain scale” spiking neural net simulation
- 147,000 processors
- 144 terabytes of memory

Dharmendra Modha



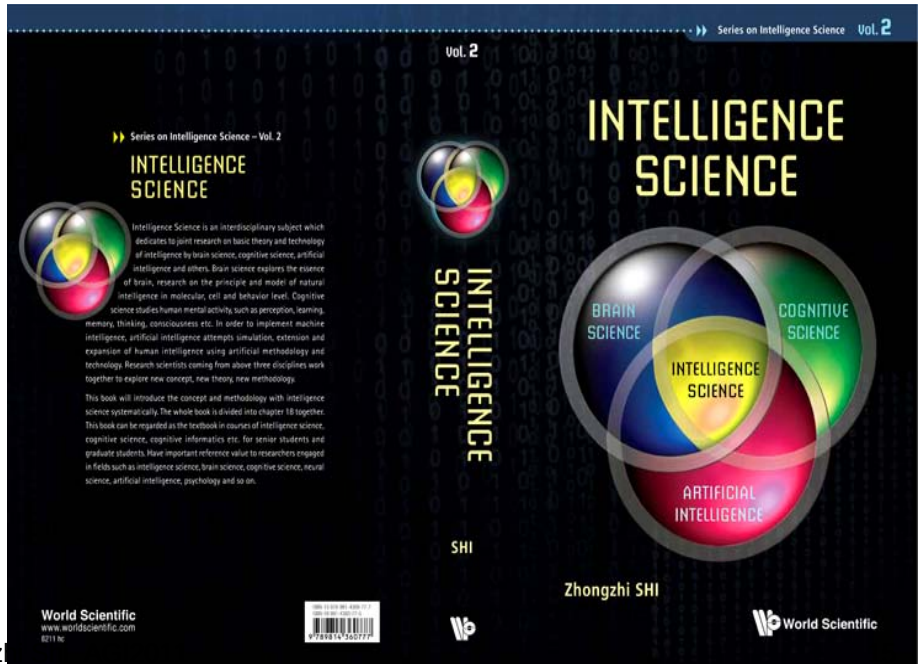


IBM Simulation

Intelligence Science

INSTITUTE OF COMPUTING TECHNOLOGY

- An interdisciplinary subject that dedicates to joint research on basic theory and technology of intelligence by brain science, cognitive science, artificial intelligence and others.



Series on Intelligence Science

INSTITUTE OF COMPUTING TECHNOLOGY

World Scientific Publishing will publish Series on Intelligence Science. Prof. Zhongzhi Shi is the Editor-in-Chief.

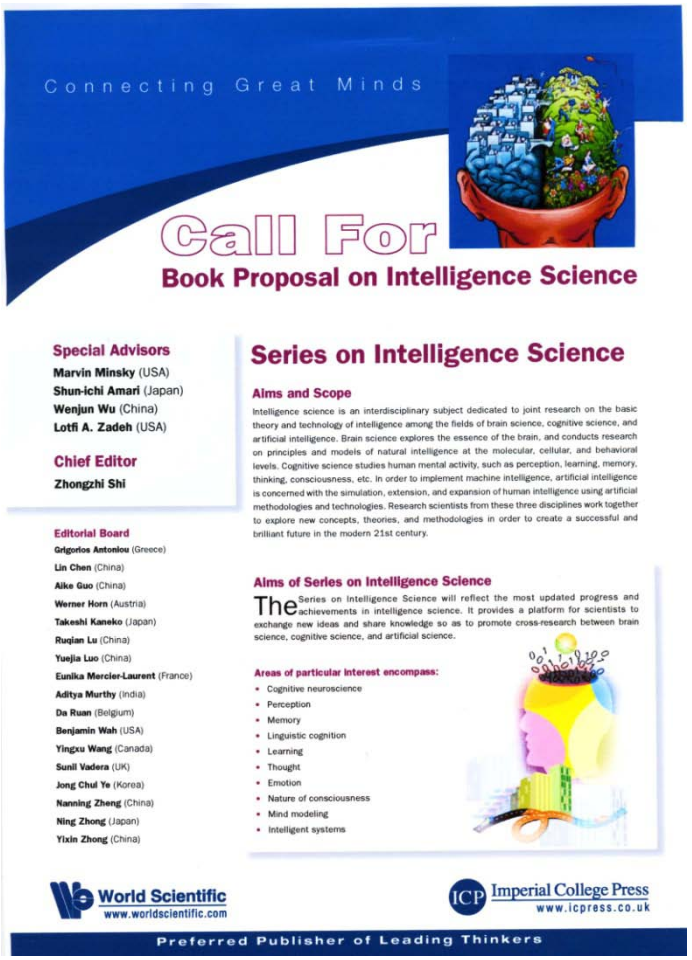
Special Advisors：

Marvin Minsky (USA)

Shun-ichi Amari (Japan)

Wenjun Wu (China)

Lotfi A. Zadeh (USA)



Basic Issues of Intelligence Science

INSTITUTE OF COMPUTING TECHNOLOGY

- How Do Brain Neural Circuits Work?
- What is Perceptual Representation and Theory of Perception?
- How Are Memories Stored and Retrieved?
- What is the Neural Basis of Language?
- How Does the Brain Learn?
- How to Think in Human Brain?
- What is the Procedure of Intelligence Development?
- What is the Nature of Emotion?
- What is the Nature of Consciousness?
- How to Build Mind Model?

Zhongzhi Shi. ***Foundations of Intelligence Science. International Journal of Intelligence Science,*** *PP.8-16, 2011*

Contents Outline

INSTITUTE OF COMPUTING TECHNOLOGY

- Introduction
- Mind Model CAM
- Feature Binding
- Understanding
- World Wide BrainImage
- Perspectives

Mind Model CAM

INSTITUTE OF COMPUTING TECHNOLOGY

CAM-**C**onsciousness **a**nd **M**emory Model is

A new mind modeling for human cognitive activities

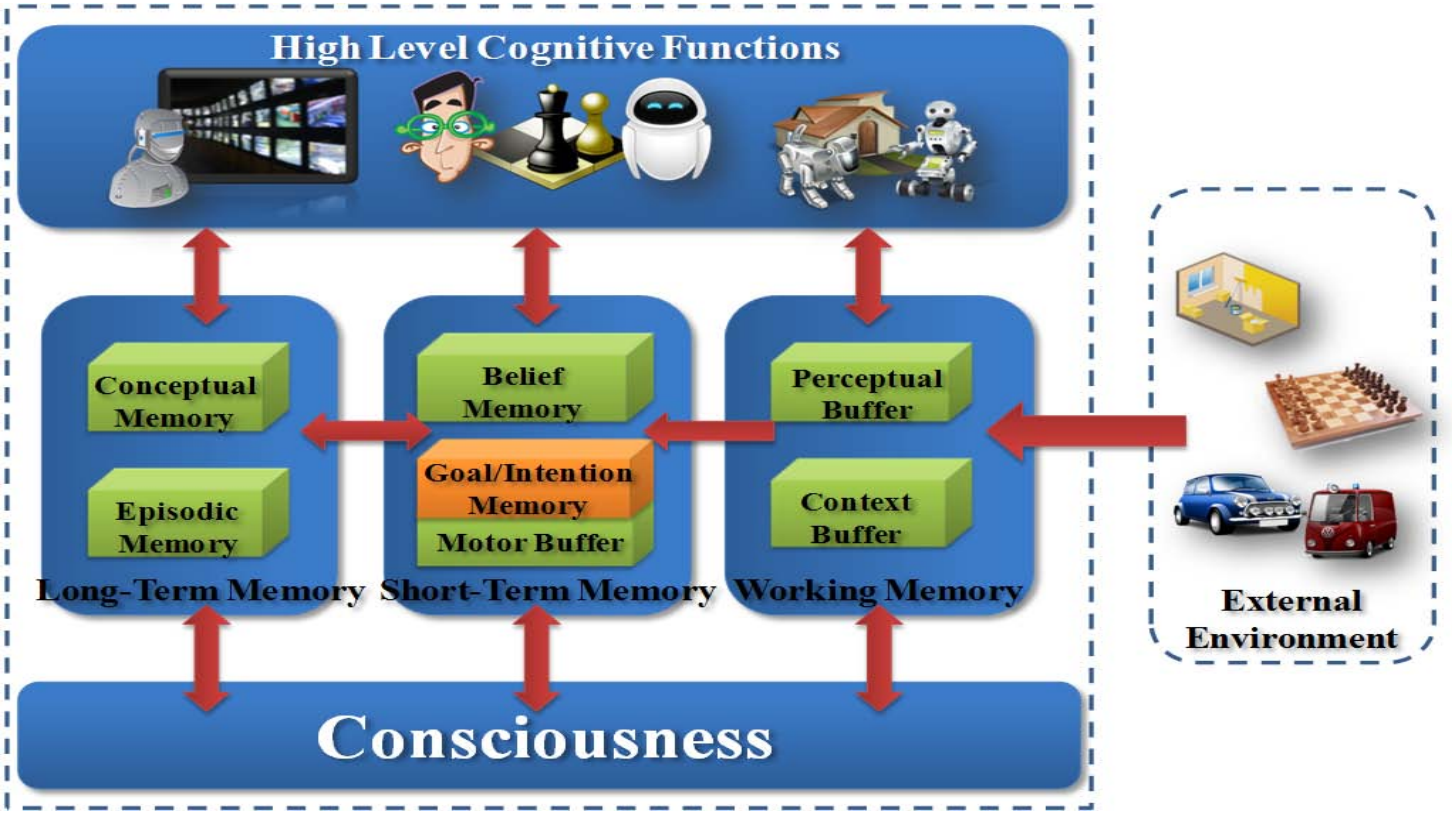
Characteristics of CAM

- Memory based Mind Model;
- Consciousness-centered Cognitive Activities;
- Action planning based way to arrange different cognitive function modules to perform cognitive activities;

CAM Architecture

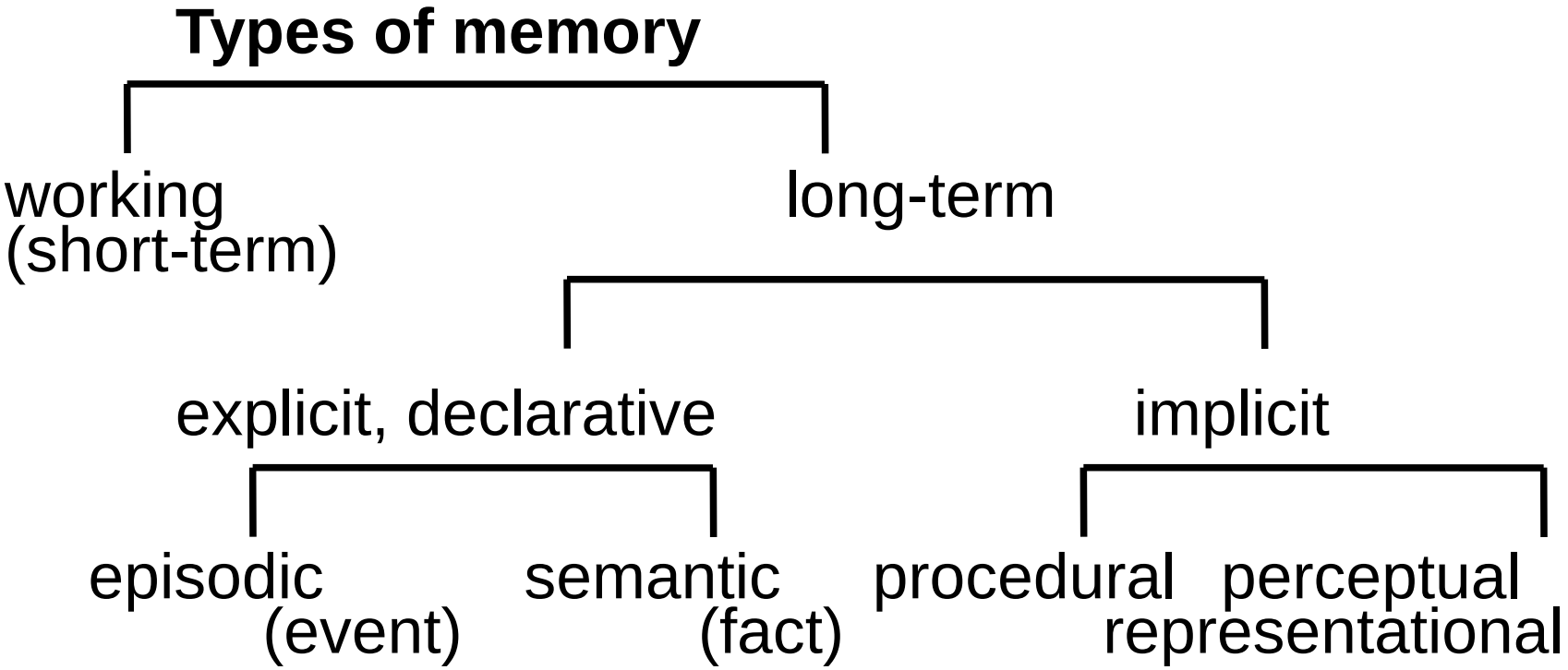
INSTITUTE OF COMPUTING TECHNOLOGY

CAM-Consciousness and Memory Model is A new mind modeling for human cognitive activities



Long-Term Memory

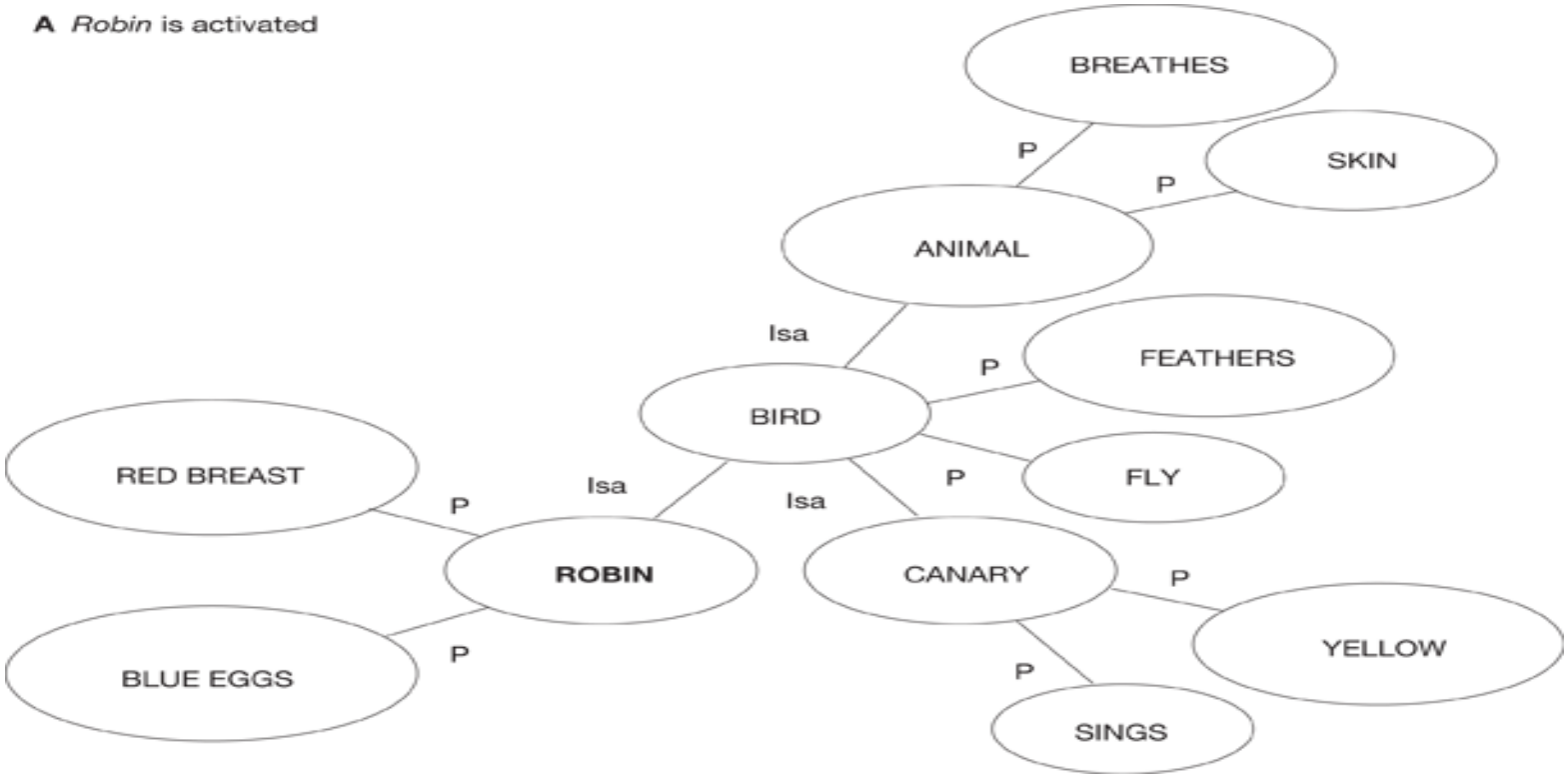
Tulving, Endel (1972). Episodic and semantic memory. Organization of memory. In E. Tulving and W. Donaldson. New York, Academic Press.



Semantic Memory

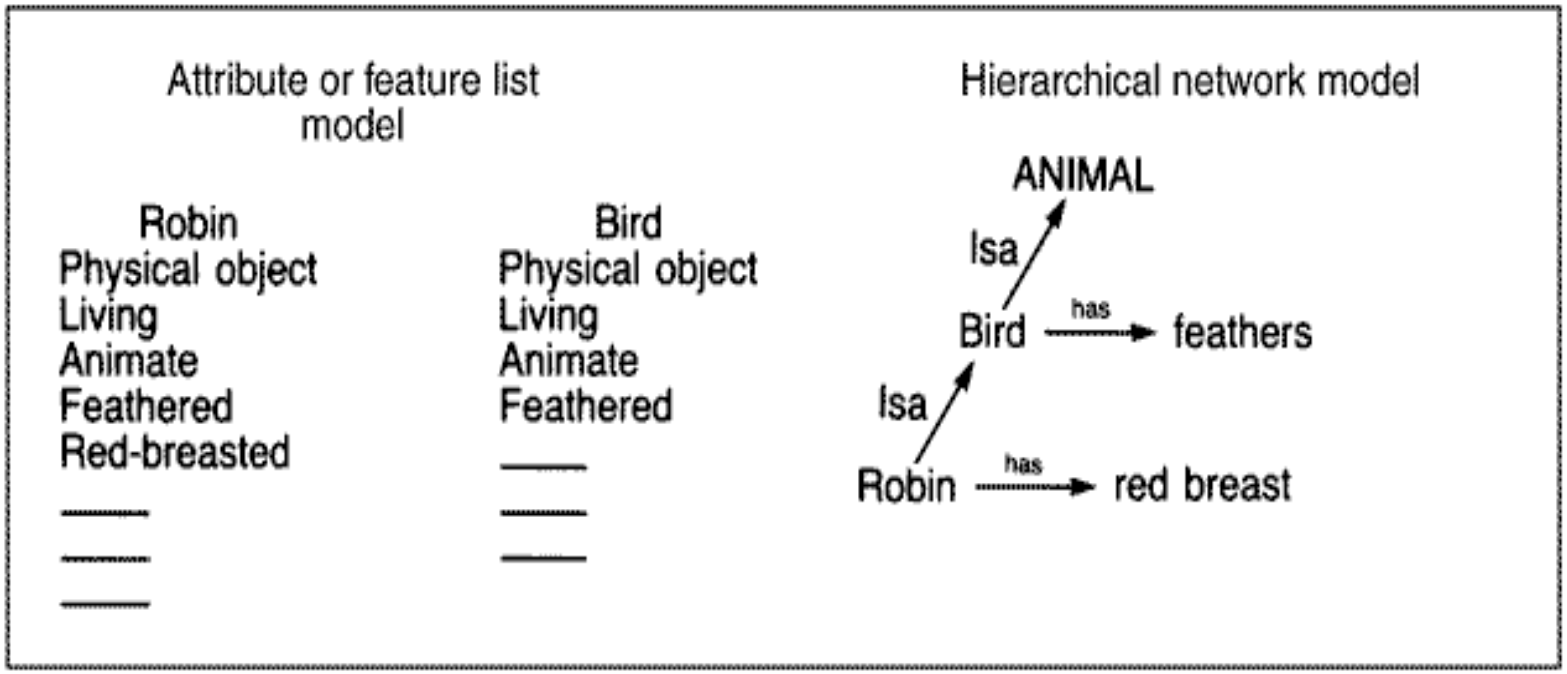
INSTITUTE OF COMPUTING TECHNOLOGY

Collin and Quillian Model



Semantic Memory

Smith's feature overlap model



Dynamic Description Logic

In order to represent agent better, we add dynamic action to DL and propose Dynamic Description Logic

$$A(x_1,\ldots,x_n)=(P_A,E_A)$$

where

(1) *A is the action name.*

(2) $x_1, \ldots, x_n$ *are individual variables, which denote the objects the action operate on.*

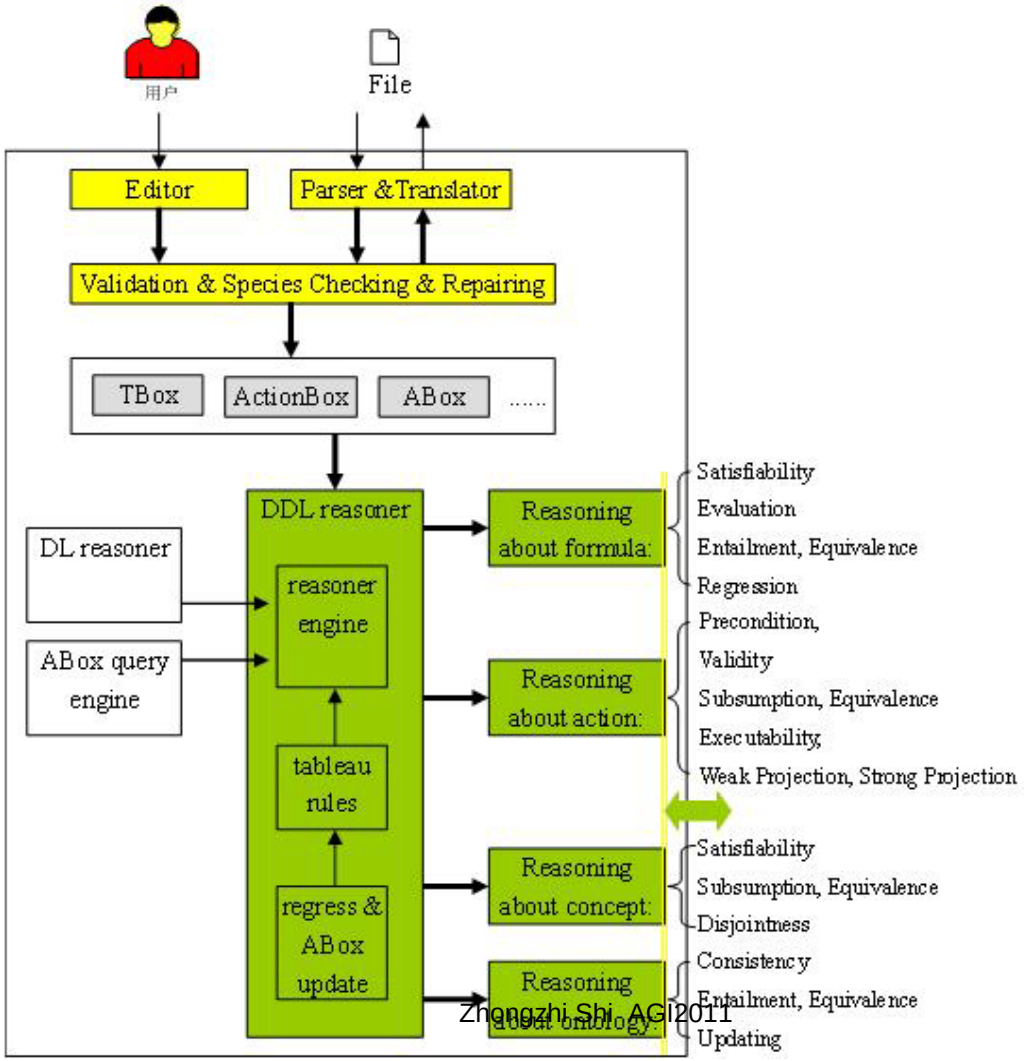
(3) P_A *is the set of preconditions, which must be satisfied before the action is executed.*

(4) E_A *is the set of results, which denote the effects of the action.*

Dynamic Description Logic

- **Concept name:** $C_1, C_2, \ldots$;
- **Role name:** $R_1, R_2, \ldots$;
- **Individual constant:** $a, b, c, \ldots$;
- **Individual variable:** $x, y, z, \ldots$;
- **Concept operation:** $\neg, \sqcap, \sqcup, \exists, \forall$;
- **Axiom operation:** $\neg, \wedge, \rightarrow\forall$;
- **Action:** $A_1, A_2, \ldots$;
- **Action-constraint : ; (composition) , ∪ (alternation), * (repeat), ? (test);**
- **Action variable:** $\alpha,\beta, \ldots$;
- **Axiom variable:** $\varphi, \psi, \Box, \ldots$;
- **State variable:** $u, v, w, \ldots$;

DDL Inference Engine



Episodic Memory

INSTITUTE OF COMPUTING TECHNOLOGY

- Episodic Memory

  The memory about past experience; contextualized store of specific events.

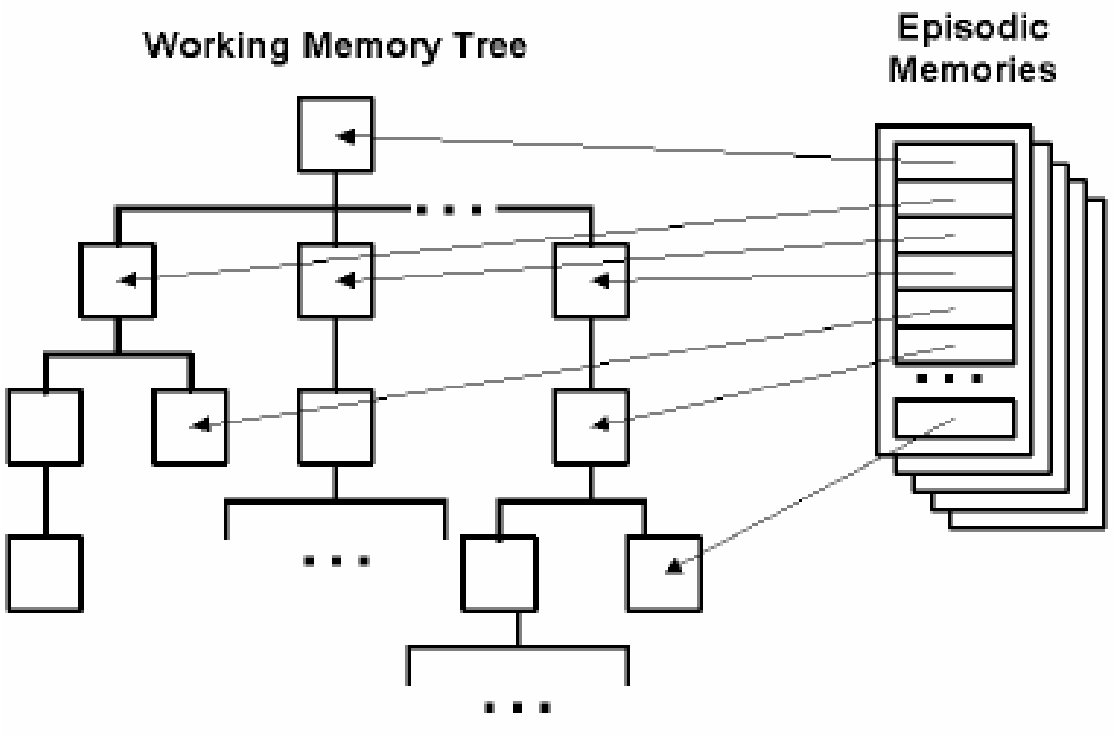
  Example: a movie seen yesterday

- Element in Episodic Memory

  episode: the stored past event;

  episode in CAM is in form of possible world sequence.

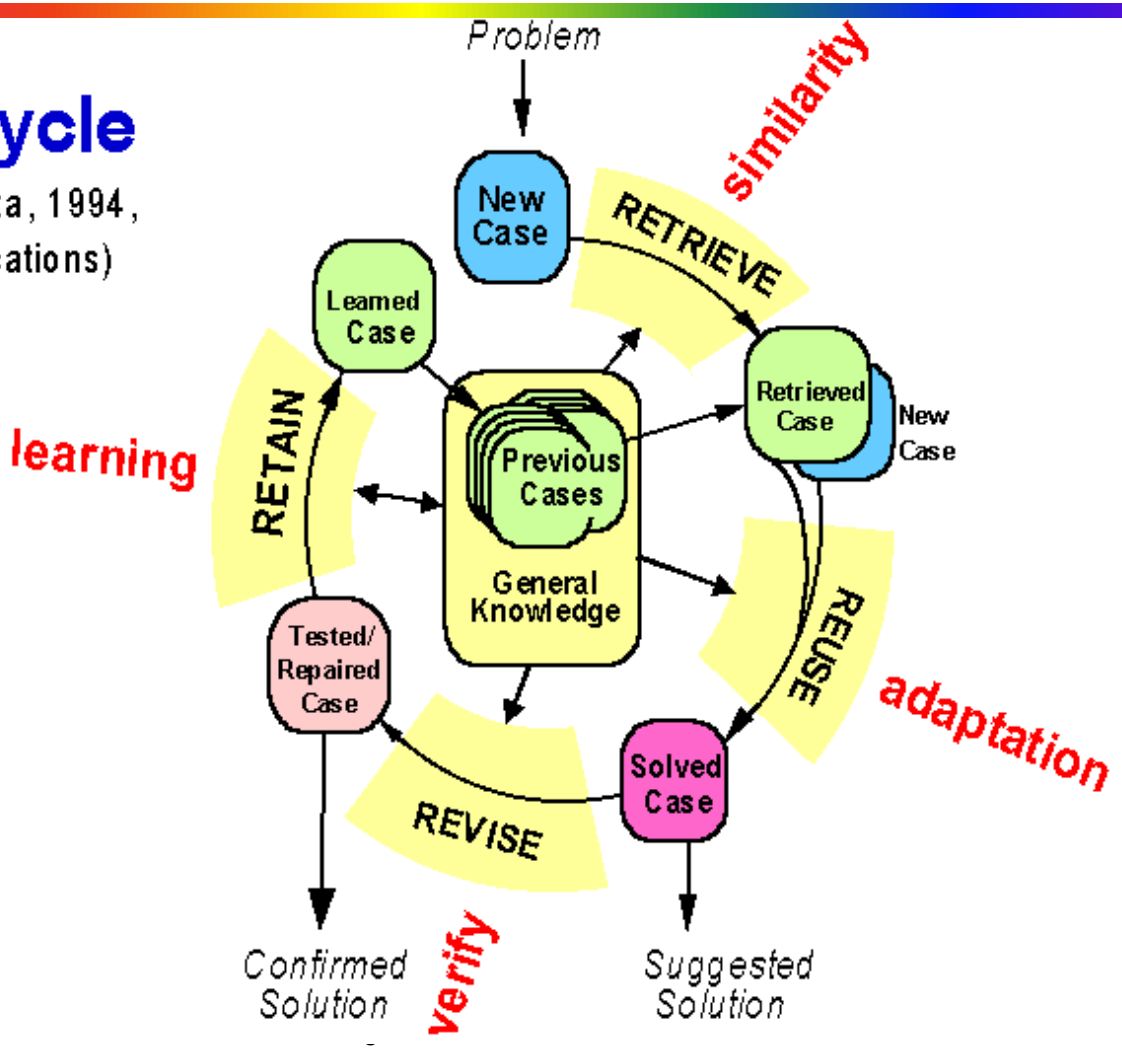
Episodic Memory



Episodic Memory

CBR Cycle

(Aamodt & Plaza, 1994, AI Communications)



Episodic Memory in CAM

INSTITUTE OF COMPUTING TECHNOLOGY

- Episode Retrieval

  ☞Retrieve episode by cue

  Cue is also in form of possible world sequence;

  Retrieve episode by matching the possible world sequence representing cue with the possible world sequence representing episode.

  ☞ Retrieve episode likes Case-based reasoning

Episodic Memory in CAM

- Matching Cue with Episode

$MatchPossibleWorld(w_1, w_2)$ is true iff $w_1 \models w_2$

$MathcAction(a_1, a_2)$ is true iff $P_{a_1} \models P_{a_2}$, $E_{a_1} \models E_{a_2}$

episode=$\{w_1, a_1, w_2, \ldots, w_n\}$, cue=$\{c_1, b_1, c_2, \ldots, c_k\}$,
$CueMatching(episode, cue)$ is true iff

exists sub sequence $\{w_{i+1}, a_{i+1}, \ldots, w_{i+k}\} \subseteq$ episode satisfying

$MatchPossibleWorld(w_{i+j}, c_j)$ is true $(j = 1, 2, \ldots, k)$
$MatchAction(a_{i+l}, b_l)$ is true $(l = 1, 2, \ldots, k-1)$

Cue Matching Algorithm

INSTITUTE OF COMPUTING TECHNOLOGY

Algorithm 1: CueMatching(e,c)

input : Episode e, Cue c
output: whether $c \preceq_p e$ hold

```
1  if length(e) < length(c) then
2  |  return false;
3  end
4  n_e:=first_node(e) ;
5  n_c:=first_node(c) ;
6  if MatchPossibleWorld(n_e, n_c) then
7  |  α_e := Null ;
8  |  α_c:=action(n_c) ;
9  |  if ¬(Pre(α_e) → Pre(α_c)) unsatisfiable according DDL tableau
   |  algorithm then
10 |  |  n_tmp:=n_e ;
11 |  |  while next_node(n_tmp) ≠ Null do
12 |  |  |  α_e := (α_e; action(n_tmp)) ;
13 |  |  |  if MatchAction(α_e, α_c) then
14 |  |  |  |  Let sub_e be the sub sequence by removing α_e from e ;
15 |  |  |  |  Let sub_c be the sub sequence by removing α_c from c ;
16 |  |  |  |  if CueMatching(sub_e, sub_c) then
17 |  |  |  |  |  return true;
18 |  |  |  |  end
19 |  |  |  end
20 |  |  |  n_tmp:=next_node(n_tmp) ;
21 |  |  end
22 |  end
23 end
24 Remove n_e from e ;
25 return CueMatching(e, c) ;
```

Match Possible World & Action

INSTITUTE OF COMPUTING TECHNOLOGY

Function `MatchPossibleWorld`(w_i, w_j)

input : possible worlds w_i, w_j
output: whether $w_i \models w_j$ hold

1 f_w:=$Conj(w_i) \rightarrow Conj(w_j)$;
2 **if** $\neg f_w$ *is unsatisfiable according to DDL tableau algorithm* **then**
3 | *return* true ;
4 **else**
5 | *return* false ;
6 **end**

Function `MatchAction`(α_i, α_j)

input : action α_i, α_j
output: whether $\alpha_i \models \alpha_j$ hold

1 **if** $\alpha_i == null$ *or* $\alpha_j == null$ **then**
2 | *return* false
3 **end**
4 f_{pre}:=$Conj(Pre(\alpha_i)) \rightarrow Conj(Pre(\alpha_j))$;
5 f_{eff}:=$Conj(Eff(\alpha_i)) \rightarrow Conj(Eff(\alpha_j))$;
6 **if** $\neg f_{pre}$ *and* $\neg f_{eff}$ *are unsatisfiable according to DDL Algorithm* **then**
7 | *return* true ;
8 **else**
9 | *return* false ;
10 **end**

What is Consciousness

INSTITUTE OF COMPUTING TECHNOLOGY

Wikipedia:

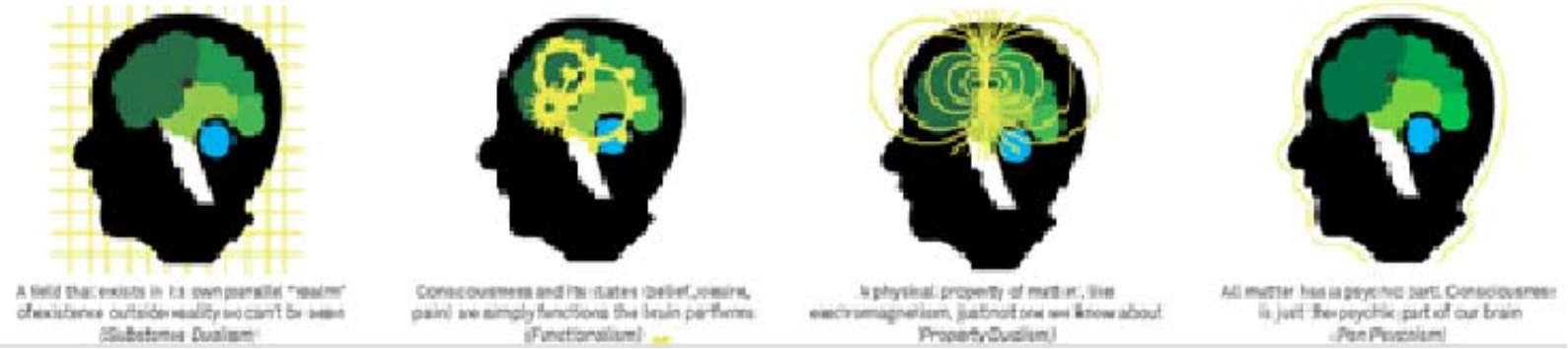
Consciousness is a term that has been used to refer to a variety of aspects of the relationship between the mind and the world with which it interacts. It has been defined, at one time or another, as: subjective experience; awareness; the ability to experience feelings; wakefulness; having a sense of selfhood; or as the executive control system of the mind.

What is Consciousness

INSTITUTE OF COMPUTING TECHNOLOGY

Stanford Encyclopedia Of Philosophy

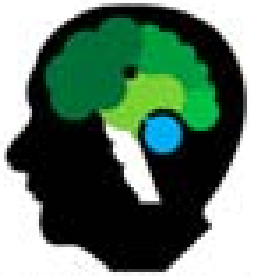
- http://www.informationisbeautiful.net/play/what-is-consciousness/



What is Consciousness

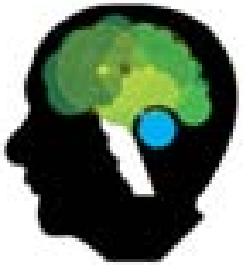
Simply, mental states are physical events that we can see in brain scans
(Identity Theory)

A sensation that "grows" inevitably out of complicated brain states
(Emergent Dualism)

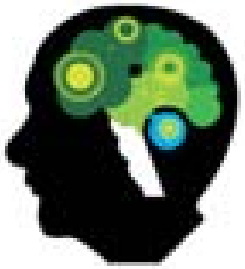


Literally just behaviour. When we behave in a certain way, we appear conscious
(Behaviourism)

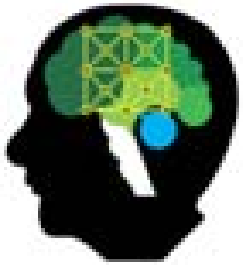
An accidental side-effect of complex physical processes in the brain
(Epiphenomenalism)



Not sure. But quantum physics, over classical physics, can better explain it
(Quantum Consciousness)



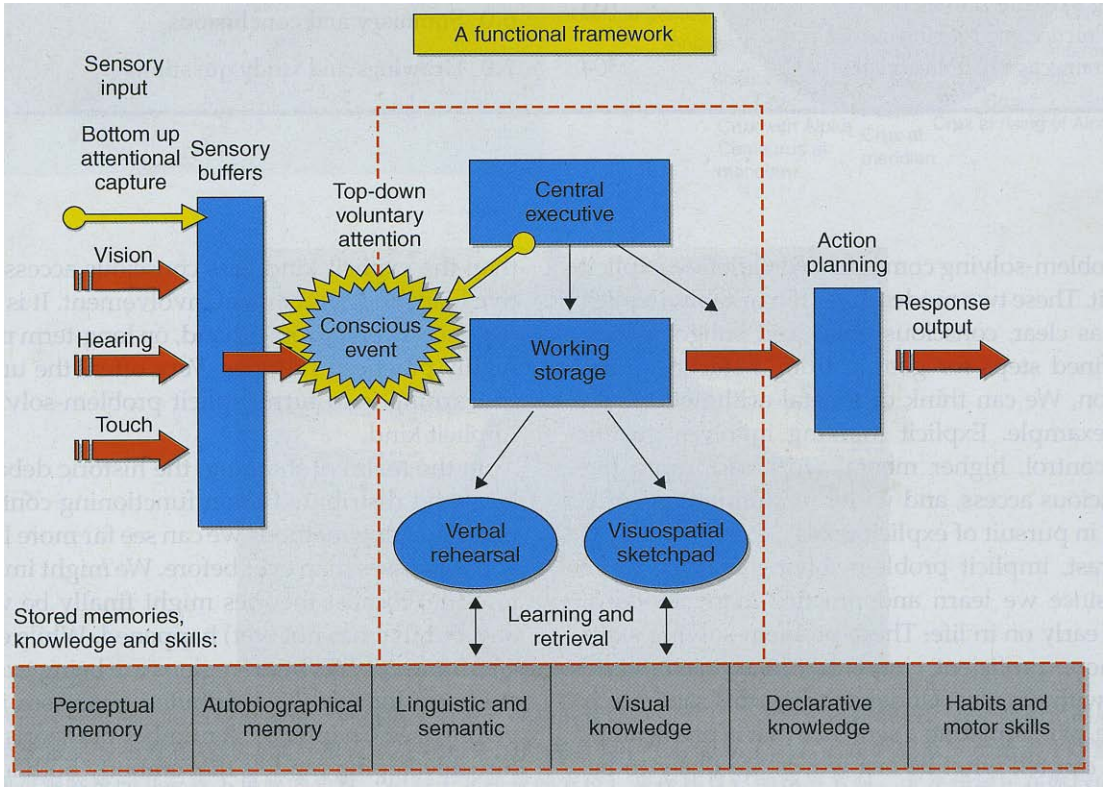
The sensation of your most significant thoughts being highlighted
(Cognitivism)

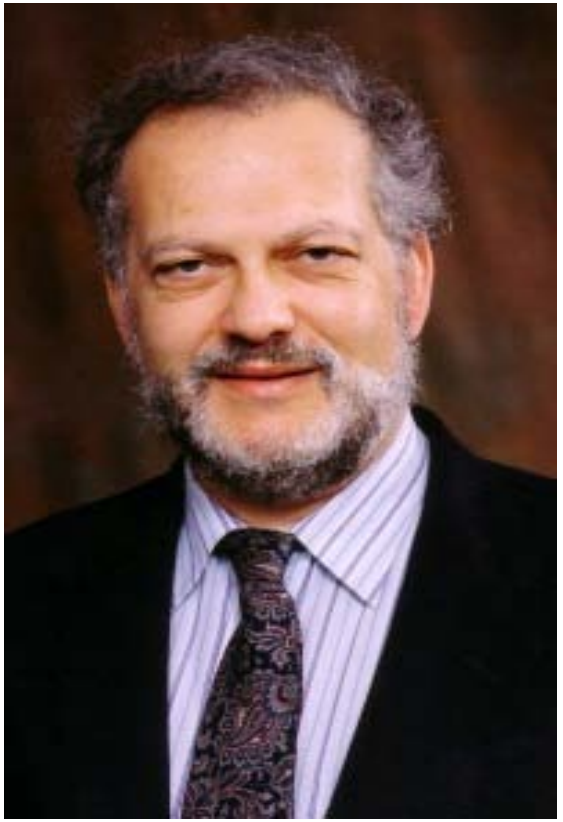


Consciousness is just higher order thoughts (thoughts about other thoughts)
(Higher Order Theory)

A continuous stream of ever-recurring phenomena, pinched like eddies into isolated minds
(Buddhism)

Framework of Cognitive Functions





B. Baars, N. Gage. Cognition, Brain, and Consciousness: Introduction to Cognitive Neuroscience. Elsevier Ltd 2007

Computational Model of Consciousness

INSTITUTE OF COMPUTING TECHNOLOGY

- Executive control system of the mind;

  Farthing, G. W. (1992). *The Psychology of Consciousness.* Upper Saddle River, NJ: Prentice Hall.

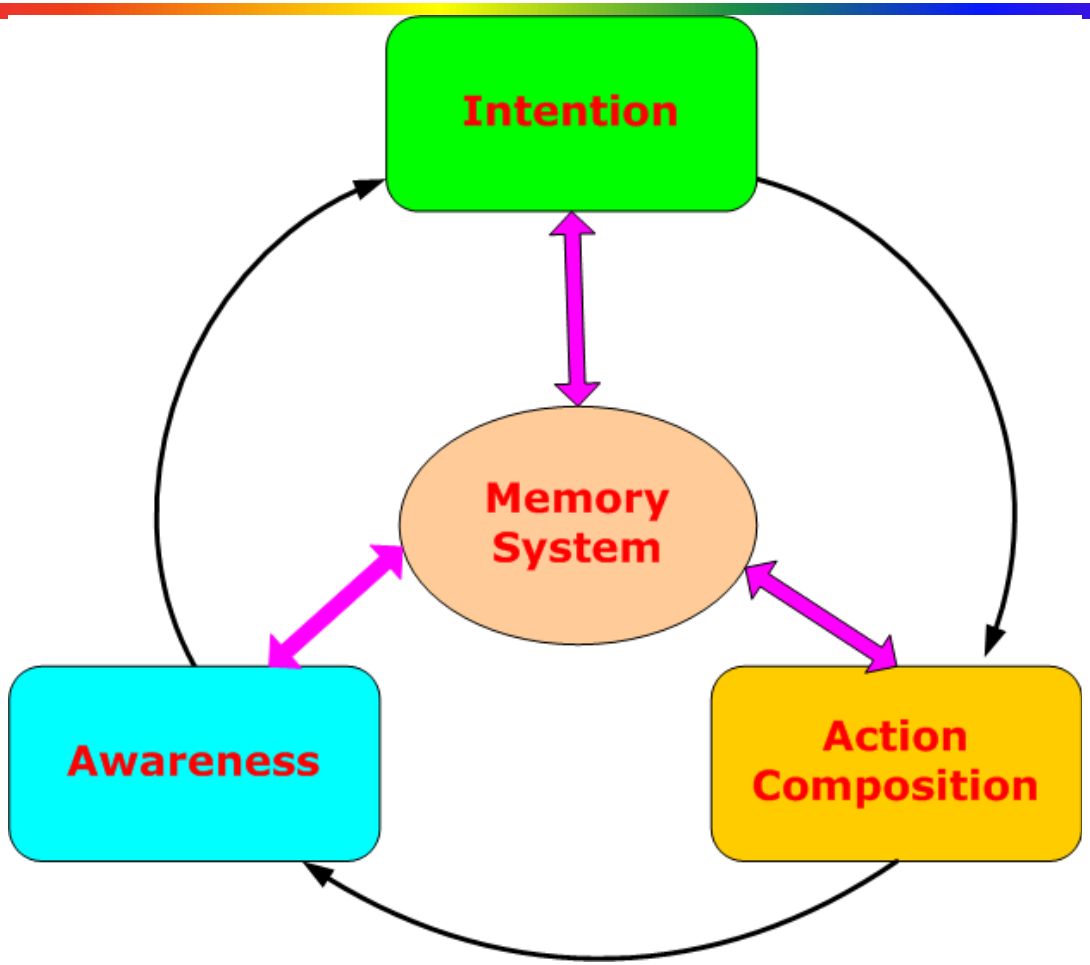
- **Intentionality**

  Intentionality refers to the feature of a desire to switch from one mental state to another mental state.

- Global Workspace Theory

  GWT was developed based on robust evidence regarding conscious processes, combined with the artificial intelligence concept of a "blackboard architecture" that combined multiple sources of knowledge in order to identify an acoustical signal in a complex, noisy, and ambiguous environment

Cognitive Cycle of CAM

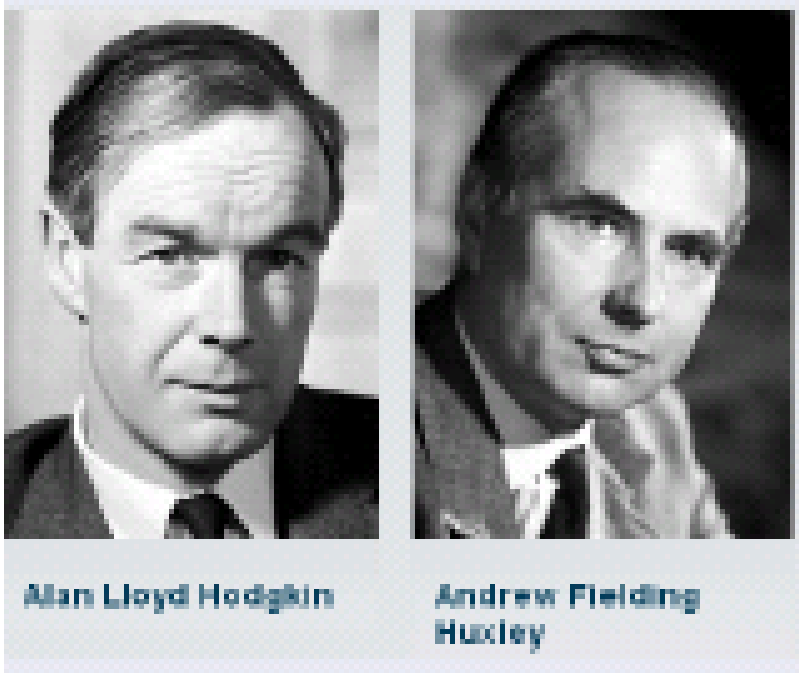


Cognitive cycle of CAM

Contents Outline

INSTITUTE OF COMPUTING TECHNOLOGY

Nobel Prize in Physiology or Medicine in 1963



- Combination of experiments, theoretical hypotheses, data fitting and model prediction
- Empirical model to describe generation of action potentials
- Published in the Journal of Physiology in 1952 in a series of 5 articles (with Bernard Katz)

The Hodgkin-Huxley Model

$$c_m \frac{dV}{dt} = -i_m + \frac{I_e}{A}$$

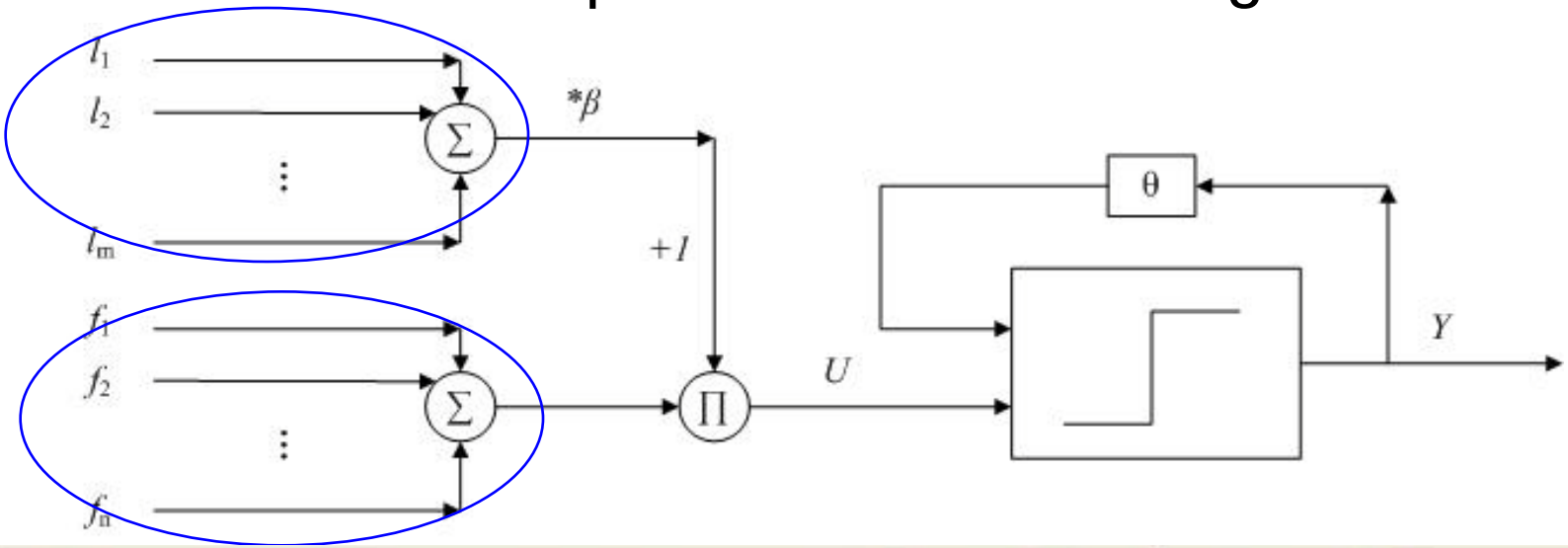
$$i_{\mathrm{m}} = \bar{g}_{\mathrm{L}}(V - E_{\mathrm{L}}) + \bar{g}_{\mathrm{K}} n^4 (V - E_{\mathrm{K}}) + \bar{g}_{\mathrm{Na}} m^3 h (V - E_{\mathrm{Na}})$$

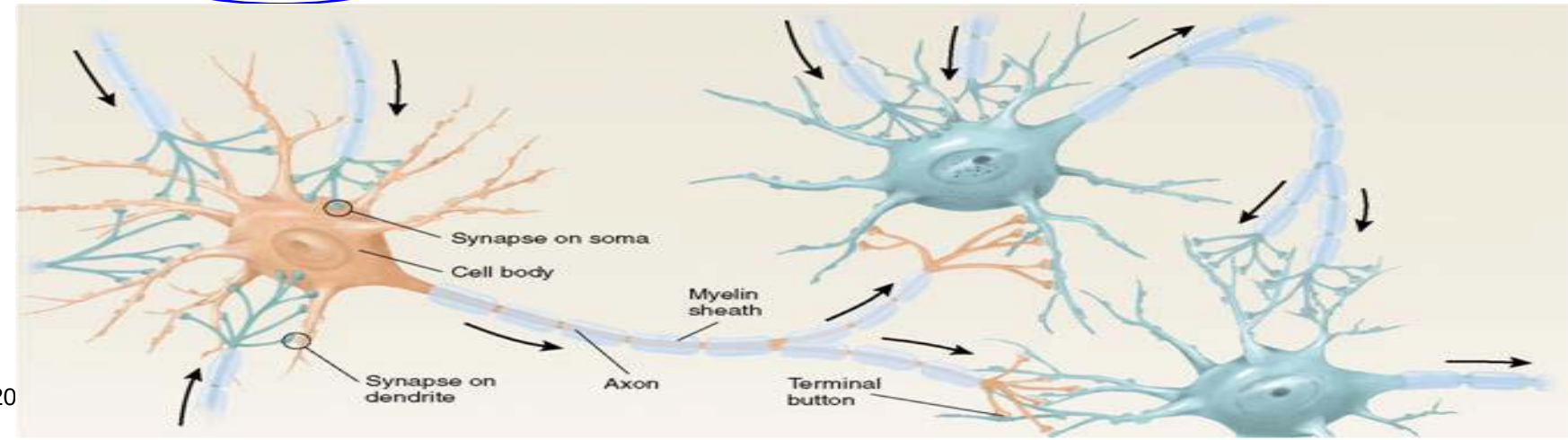
$$\tau_z(V)\frac{dz}{dt} = z_\infty(V) - z$$

Gating equation

Feature Linking Network

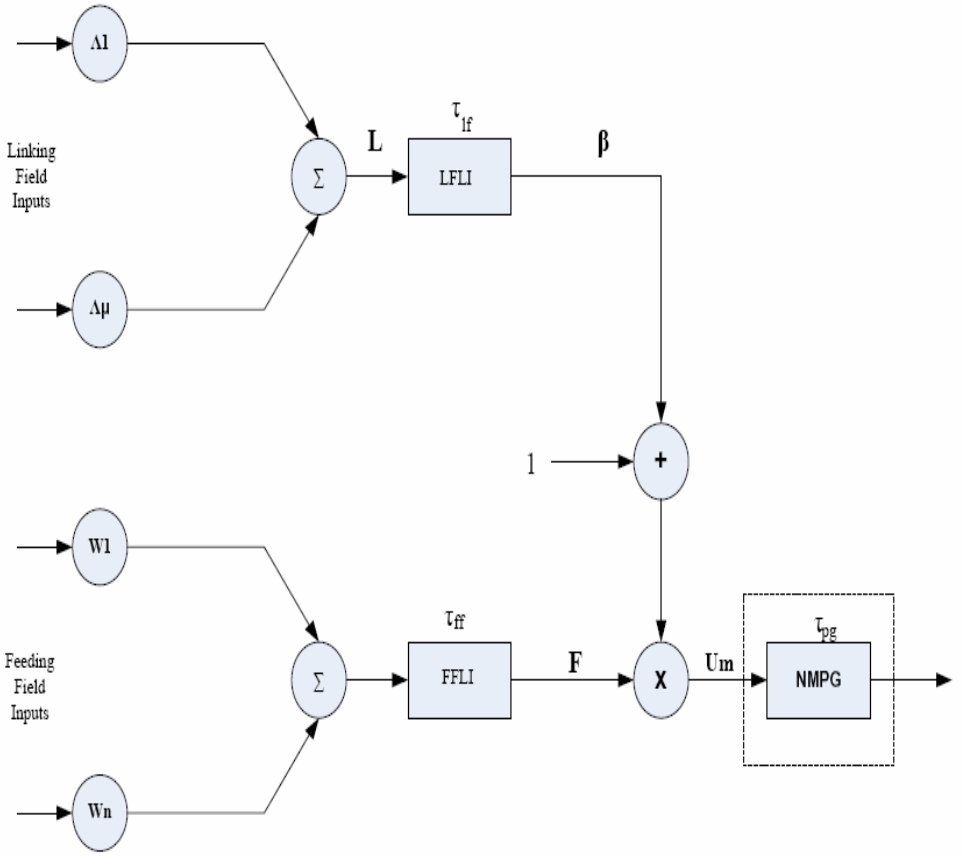
- 1990 Eckhorn Proposed Feature Linking Network





Pulse-Coupled Neural Network

The PCNN was originally presented by Eckhorn et al.
in order to explain the synchronous neuronal burst phenomena in the cat visual cortex.
The model neuron consists of three parts: the dendritic tree, the linking modulation, and the pulse generator.



Bayesian Linking Field Model

INSTITUTE OF COMPUTING TECHNOLOGY

Principle ideas

- It utilizes noisy neural model, where the inputs and outputs of neurons are firing probabilities, but not pulse
- Each neuron has two parts of inputs, namely feeding inputs and linking inputs.
- Weights of connections between neurons reflect the statistical relation between them, and can be learnt via learning.
- The outputs of neurons are determined by both their own inputs and the outputs of their neighbors.

Neuron Model

- Dynamic activity

$$P(X) = \sum_i w_i P(f_i)$$

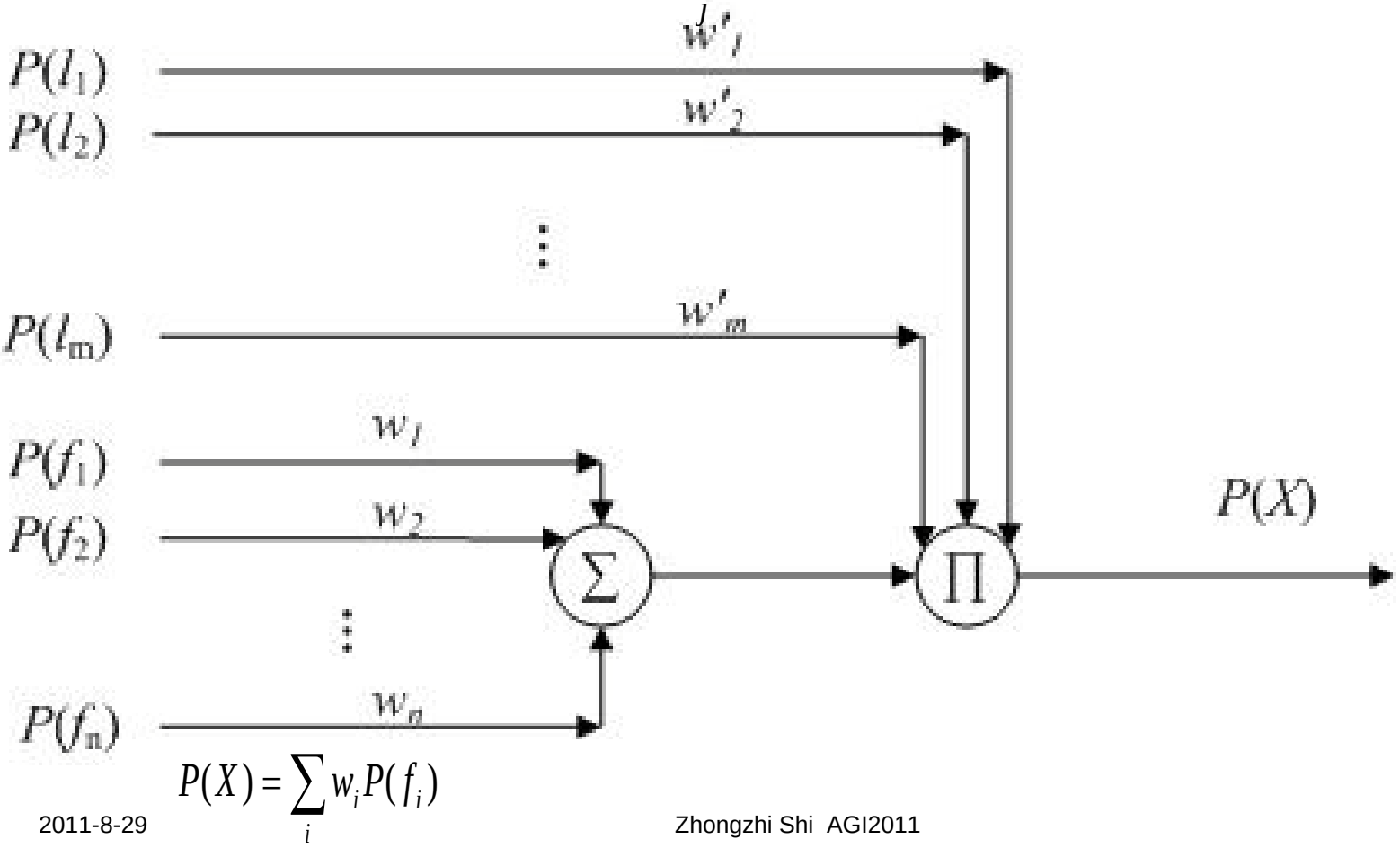
Feeding input

$$P(X \mid l_1, l_2, \ldots) = P(X) \cdot \prod_j w_j' P(l_j)$$

Linking input

Bayesian Linking Field Model

$$P(X \mid l_1, l_2, \ldots) = P(X) \cdot \prod_j w_j' P(l_j)$$



Competition Among Neurons

INSTITUTE OF COMPUTING TECHNOLOGY

- **Let *X*1 and *X*2 be two different neurons; *F*1 and *F*2 be the set of their feeding pre-synaptic neurons respectively. Then there exists competitive relation between *X*1 and *X*2 if and only if at least one of the two conditions below holds.**
  - $F_1 \cap F_2 \neq \emptyset$
  - Exist $f1 \in F1$ and $f2 \in F2$, and $f1$ and $f2$ are competitive

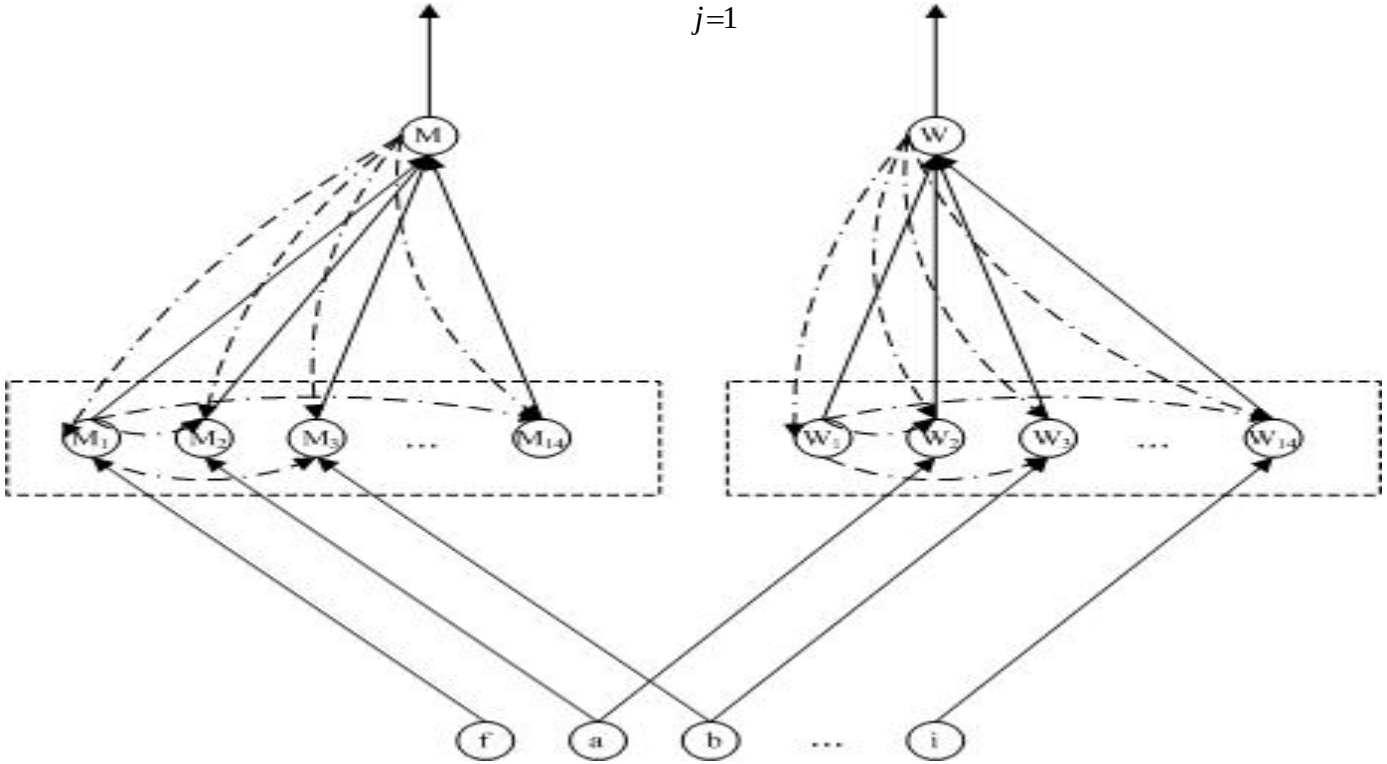
Competition Among Neurons

- To implement competitive relations, we normalize the firing probabilities of the neurons that are competitive each other.
- Let *X*1, *X*2, …, *X*n be *n* neurons that are competitive each other; *P*before(*X*i) is the firing probability of *X*i before competition. Then the firing probability of *X*i after competition is:

$$P_{after}(X_i) = \frac{P_{before}(X_i)}{\sum_{j=1}^{n} P_{before}(X_j)}$$

Bayesian Linking Field Model

$$P_{after}(X_i) = \frac{P_{before}(X_i)}{\sum_{j=1}^{n} P_{before}(X_j)}$$

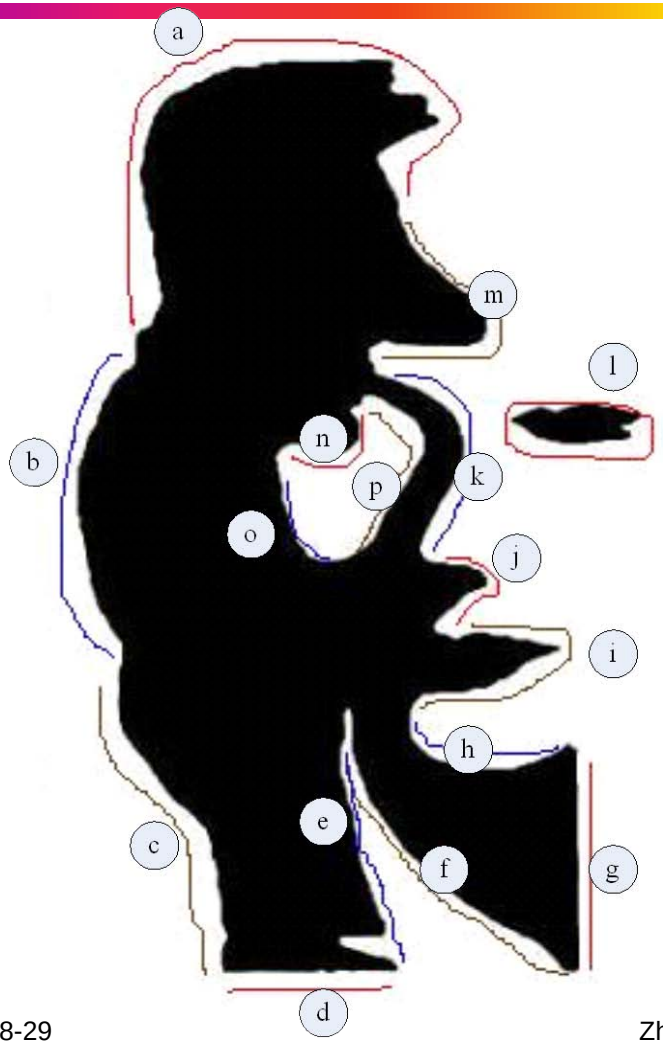


Perceptual Problem Description

The picture can be interpreted as

- a man playing a saxophone or
- a woman face

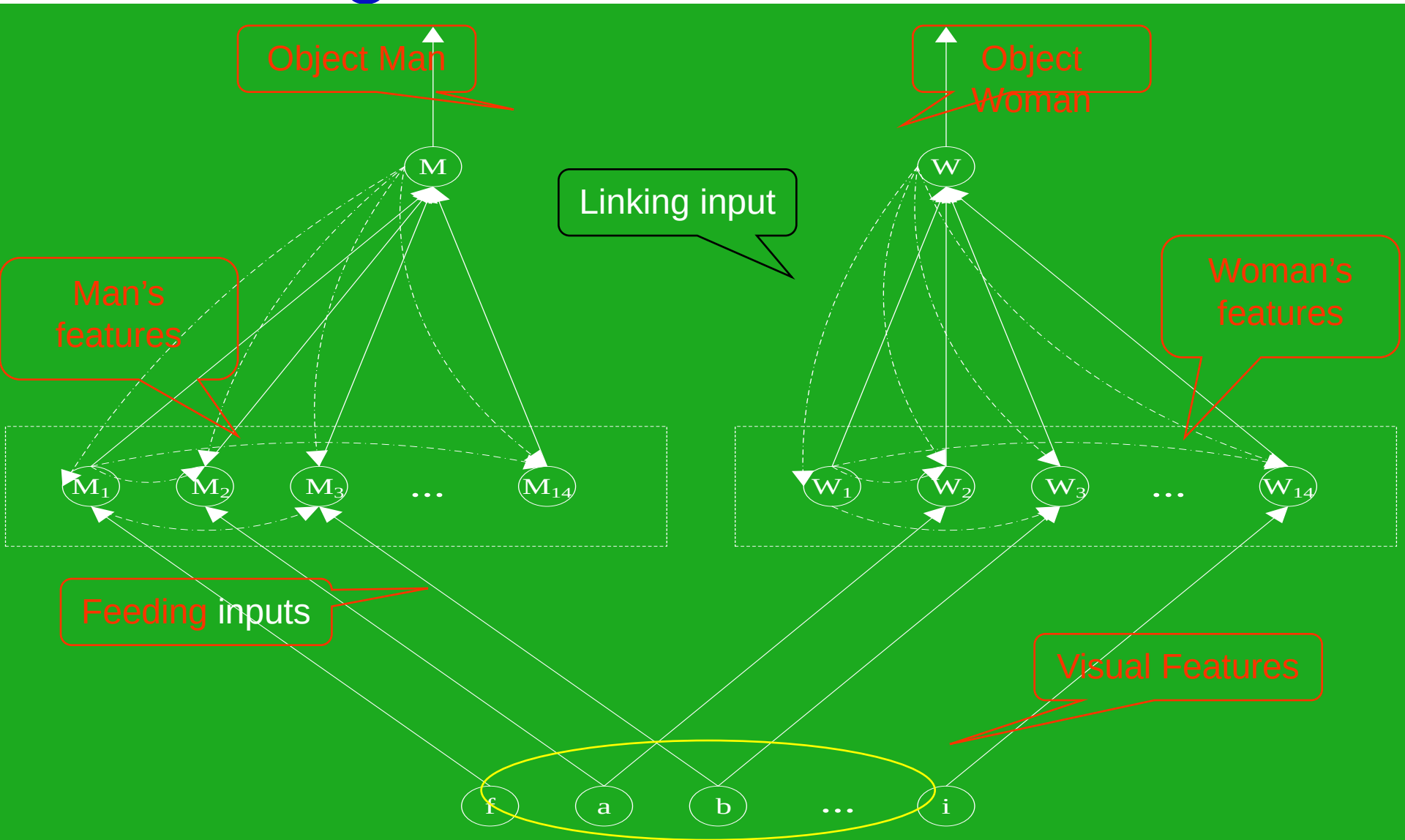
Simplification



Entire picture is partition into 16 compartments

The architecture of Bayesian Linking Field Model

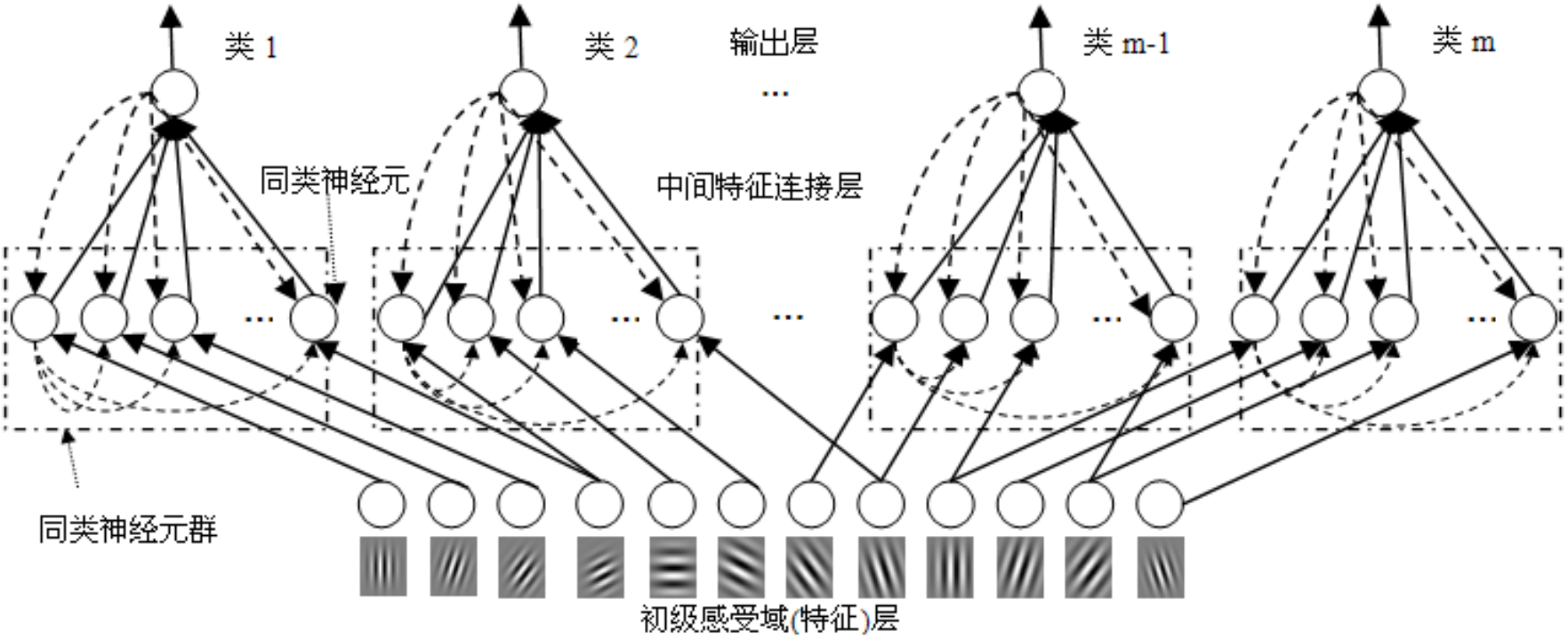
INSTITUTE OF COMPUTING TECHNOLOGY



Computational Model for Feature Binding

INSTITUTE OF COMPUTING TECHNOLOGY

Bayesian Link field Networks Model



Contents Outline

INSTITUTE OF COMPUTING TECHNOLOGY

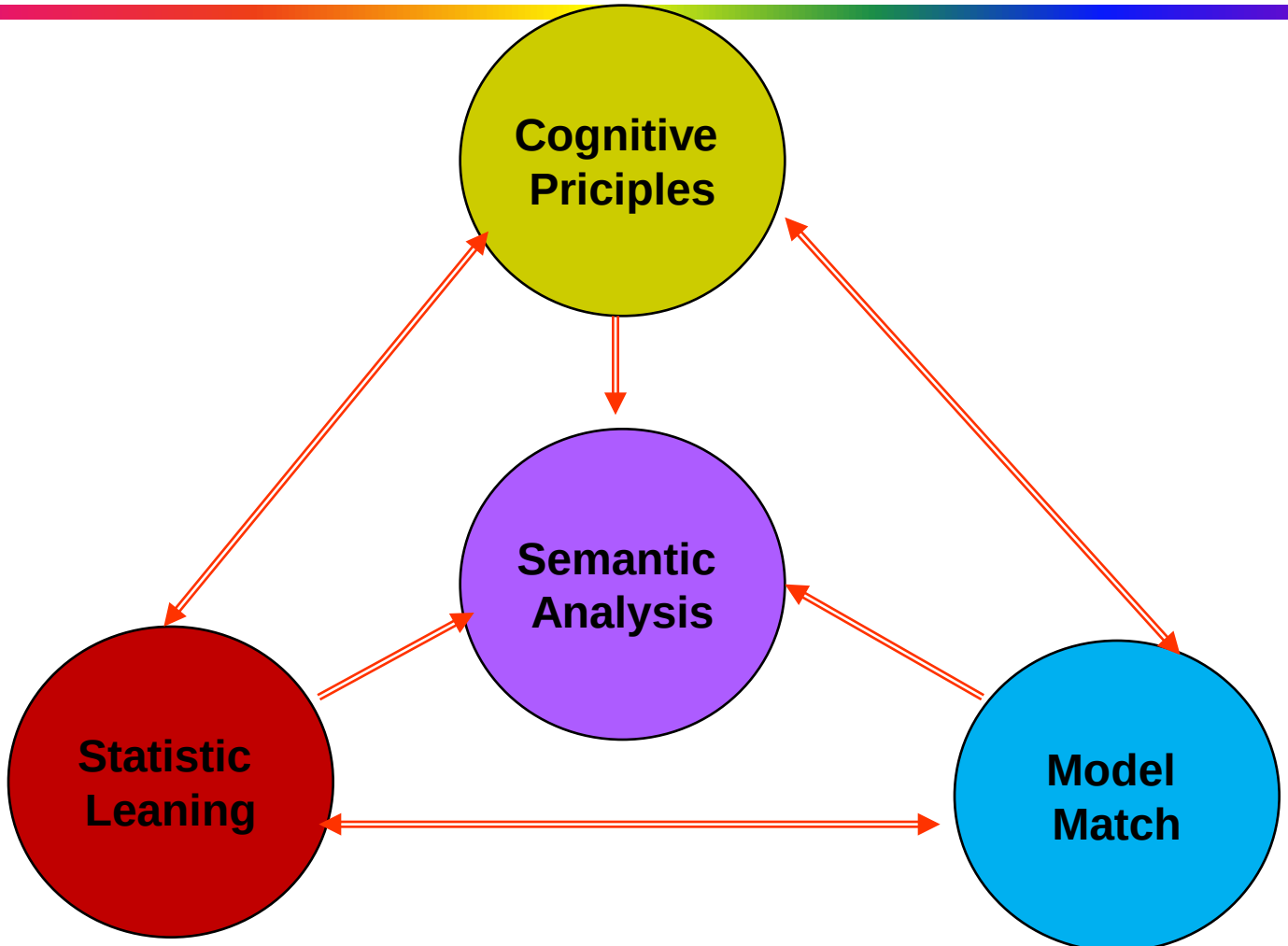
- Introduction
- Mind Model CAM
- Feature Binding
- Image Understanding
- World Wide Brain
- Perspectives

Image Understanding

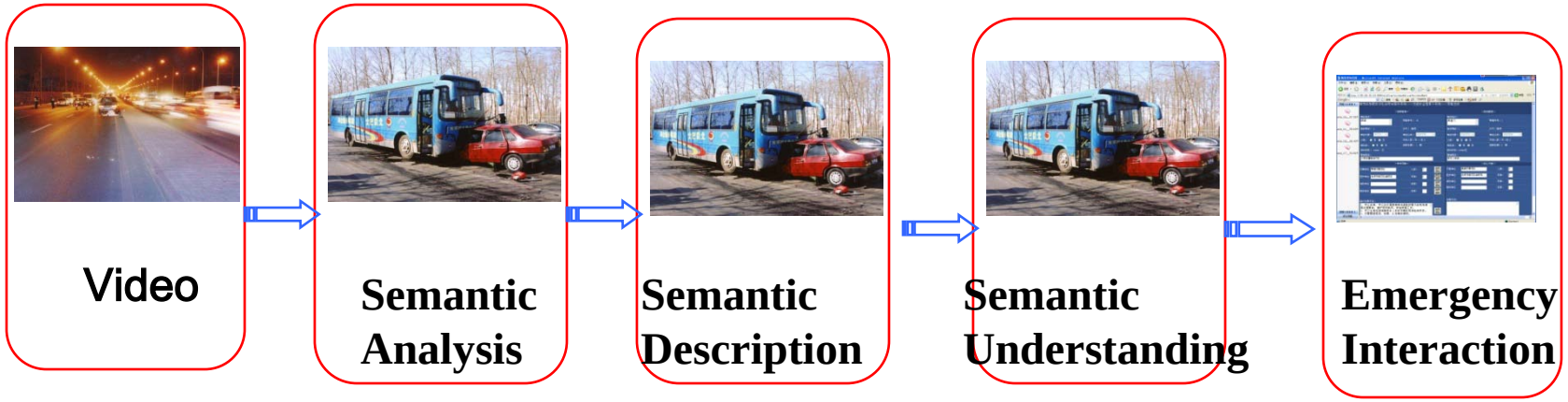
Suzhou Garden

Suzhou Building

Semantic Analysis



Traffic Surveillance

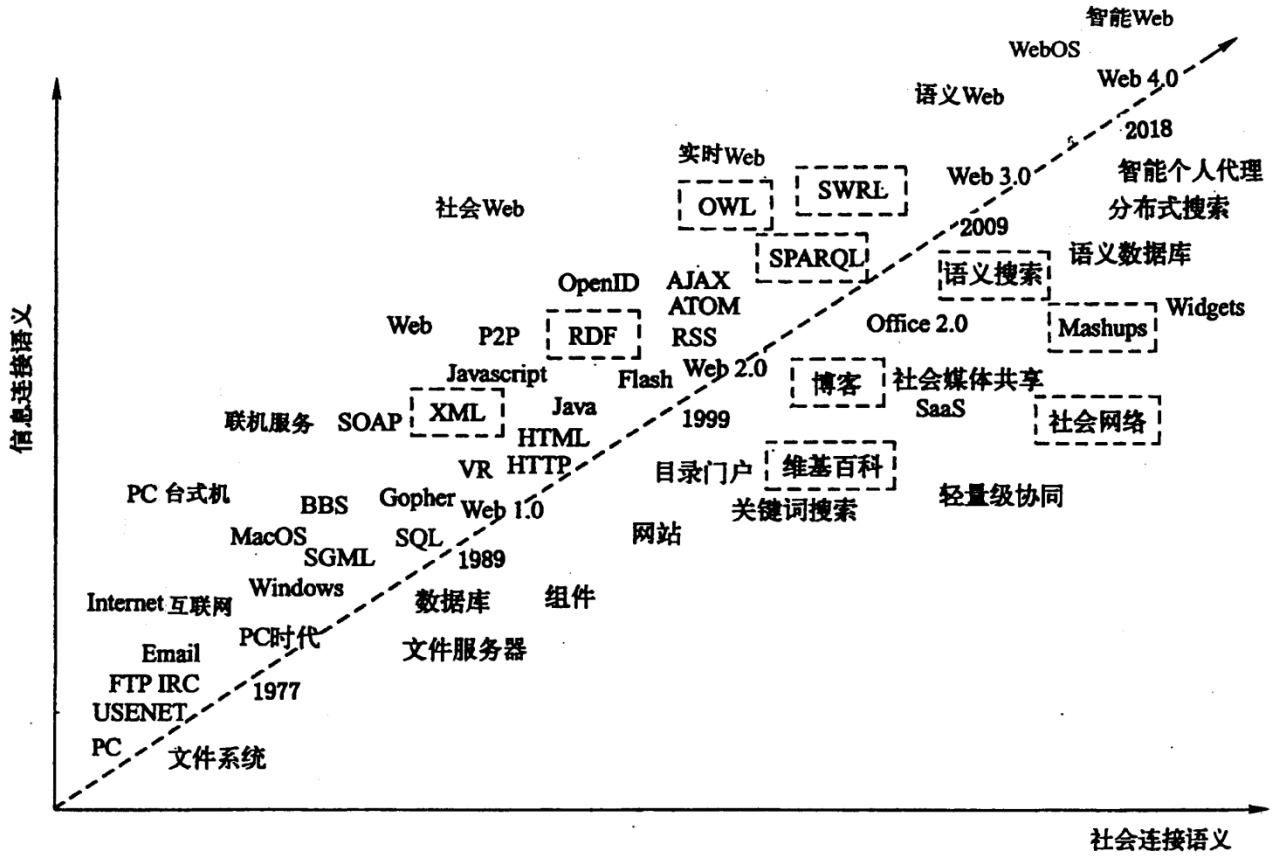


Contents Outline

INSTITUTE OF COMPUTING TECHNOLOGY

- Introduction
- Mind Model CAM
- Feature Binding
- Image Understanding
- World Wide Brain
- Perspectives

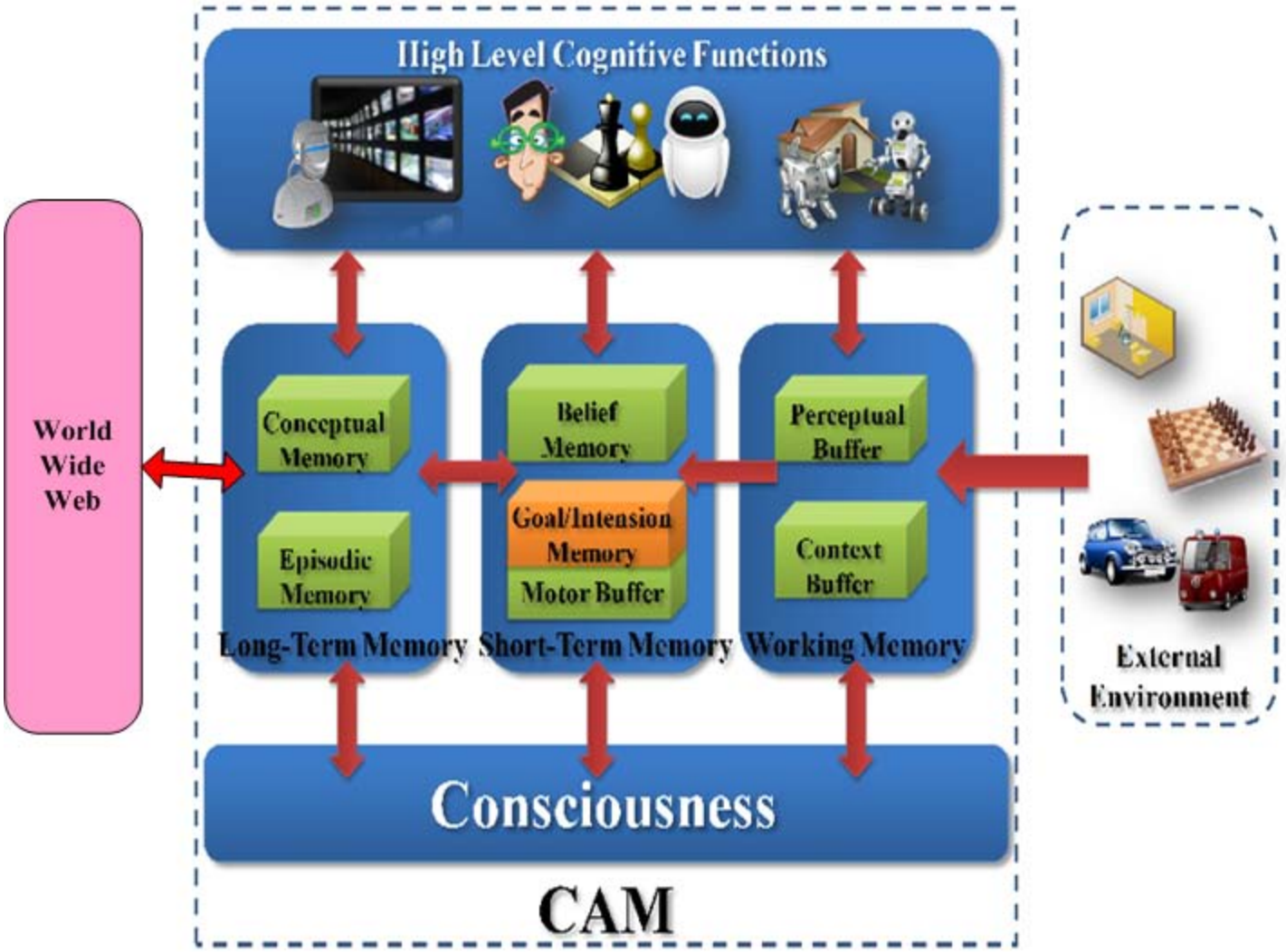
Internet Development Roadmap



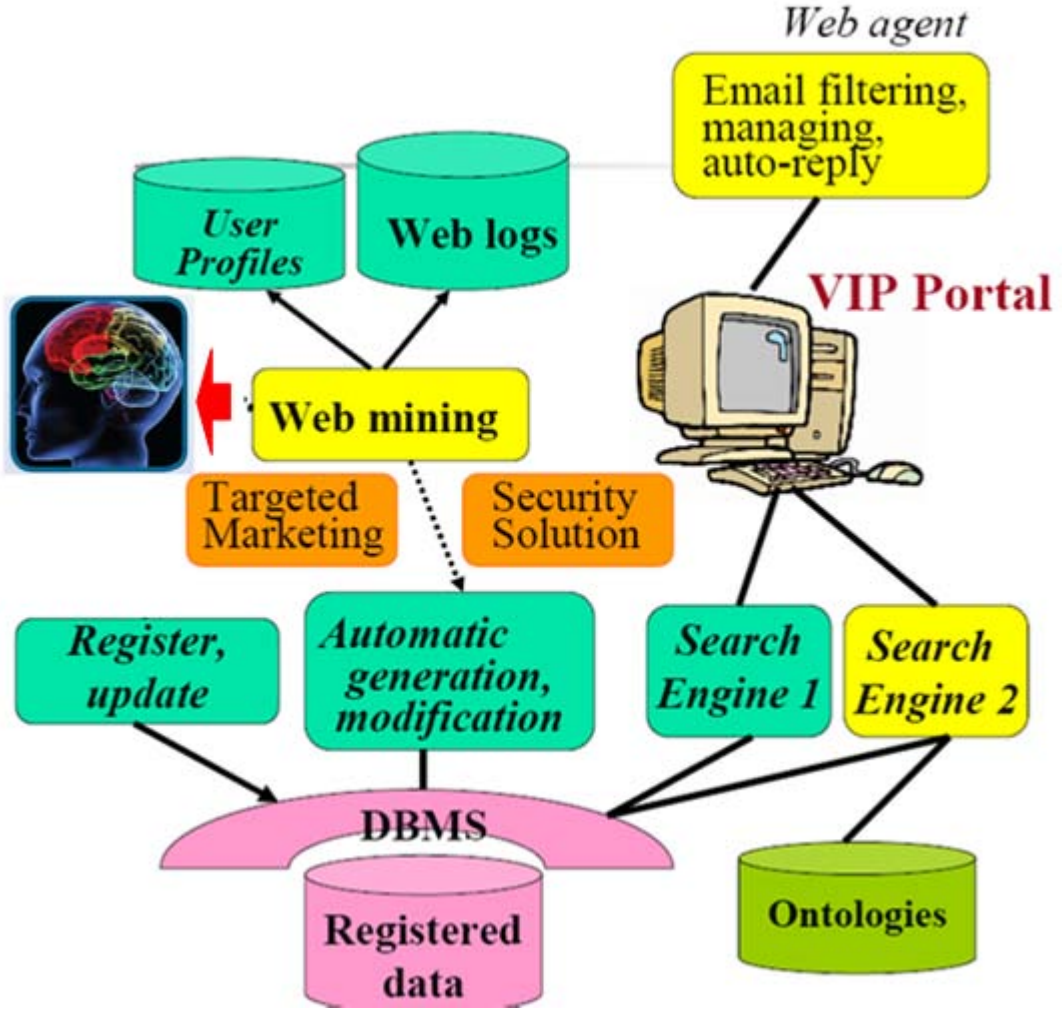
Spivack 2008

World Wide Brain

INSTITUTE OF COMPUTING TECHNOLOGY



World Wide Brain



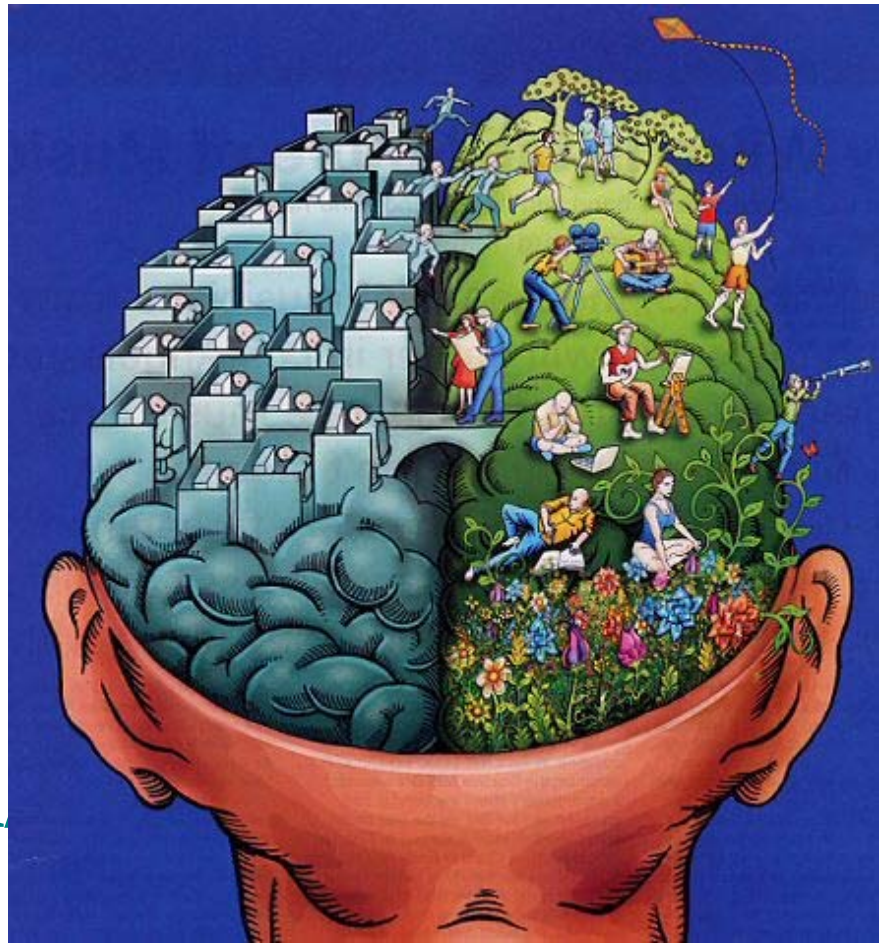
Conclusions

INSTITUTE OF COMPUTING TECHNOLOGY

- Human level intelligence needs new research paradigm.
- Intelligence Science is a new research paradigm
- The mind model CAM has obvious features and potential applications

Thank You

INSTITUTE OF COMPUTING TECHNOLOGY



Intelligence Science

http://www.intsci.ac.cn/