江苏省人工智能会议
特邀报告

新颖机器脑的研究

史忠植
shizz@ics.ict.ac.cn

智能信息处理重点实验室
中国科学院计算技术研究所
http://www.intsci.ac.cn

INSTITUTE OF COMPUTING TECHNOLOGY
中科院计算所

Acknowledgement

INSTITUTE OF COMPUTING TECHNOLOGY

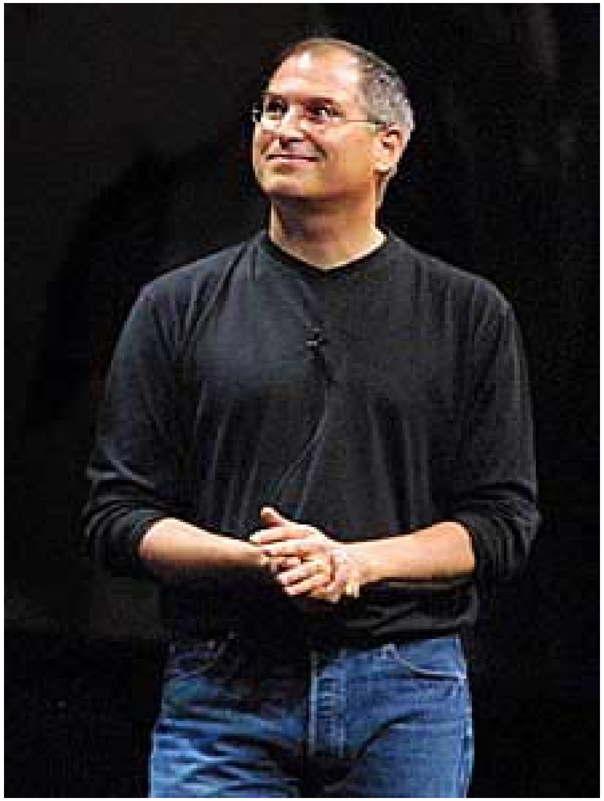
- **Key Projects of National Natural Science Foundation of China:** No.61035003, 60933004
- **National Natural Science Foundation of China:** No. 61072085, 60903141, 60970088
- **China National Basic Research Priorities Programme 973:** No. 2007CB311004

内容提要

中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY

中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY

乔布斯式创新人才



1955.2.24, 出生旧金山

1972, 高中，波兰大学修一学期.

1974， 到公司工作，设计计算机游戏

1976, 21岁， 与26岁的 Wo Zini Ike，在家车库创建Apple计算机公司.

1985, 辞职

1997.9, 返回Apple任 CEO

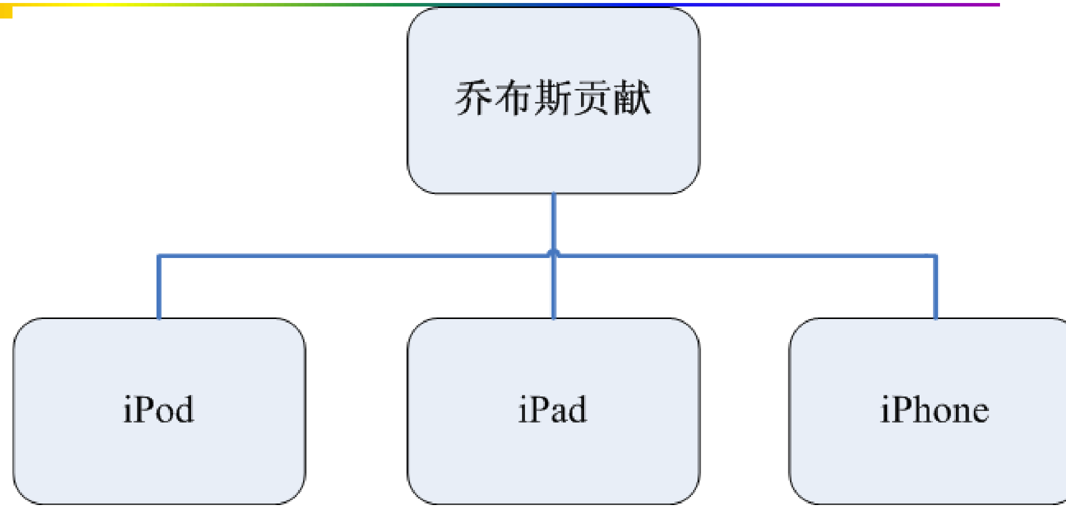
2011.8.25, 主动辞职

2011.10.5, 逝世

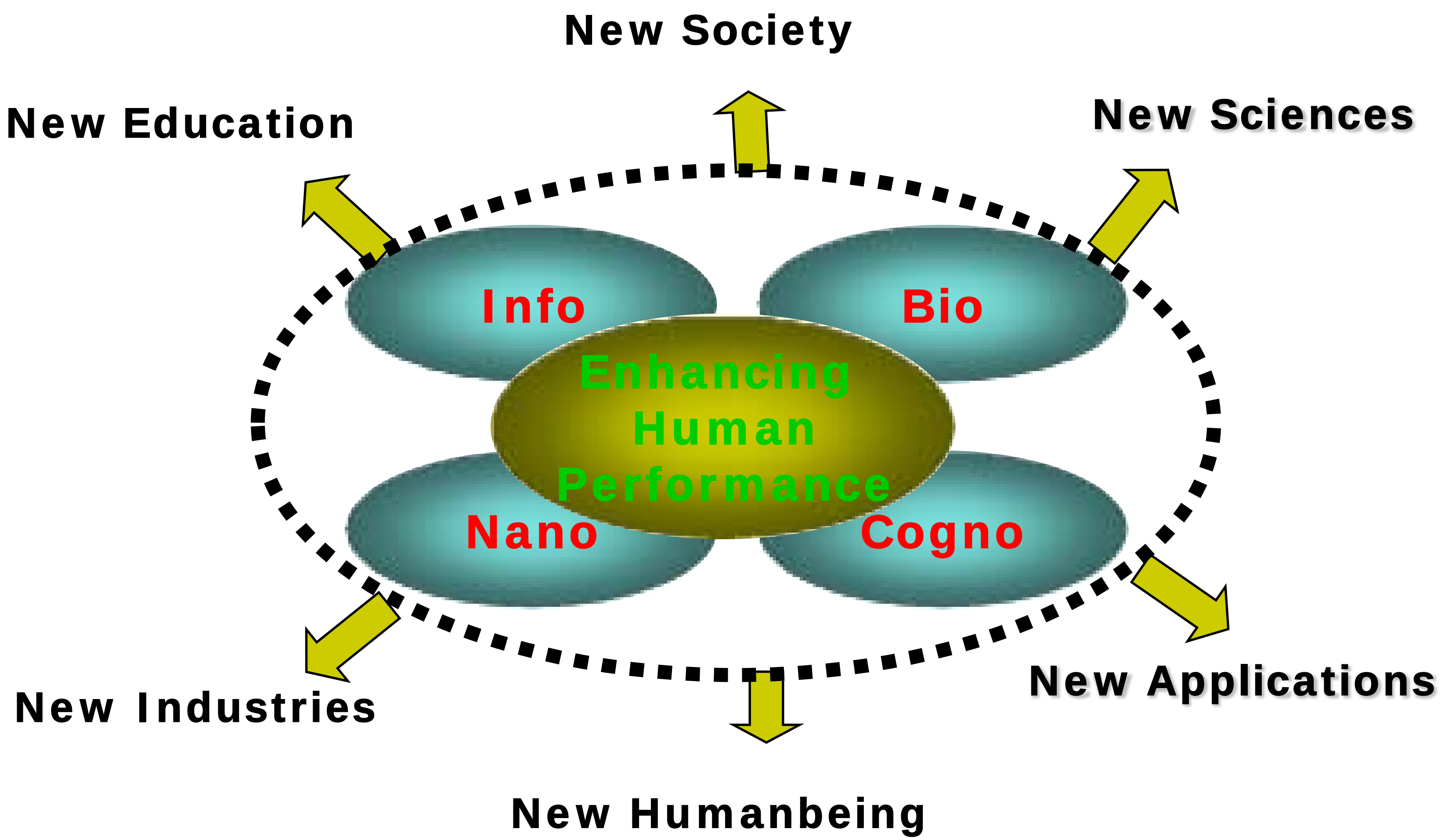
中国呼唤

中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY

乔布斯式创新人才



会聚技术



New Society
New Education
New Sciences
Info
Bio
Enhancing
Human
Performance
Nano
Cogno
New Industries
New Applications
New Humanbeing

信息技术

- 网络化
- 信息化
- 智能化

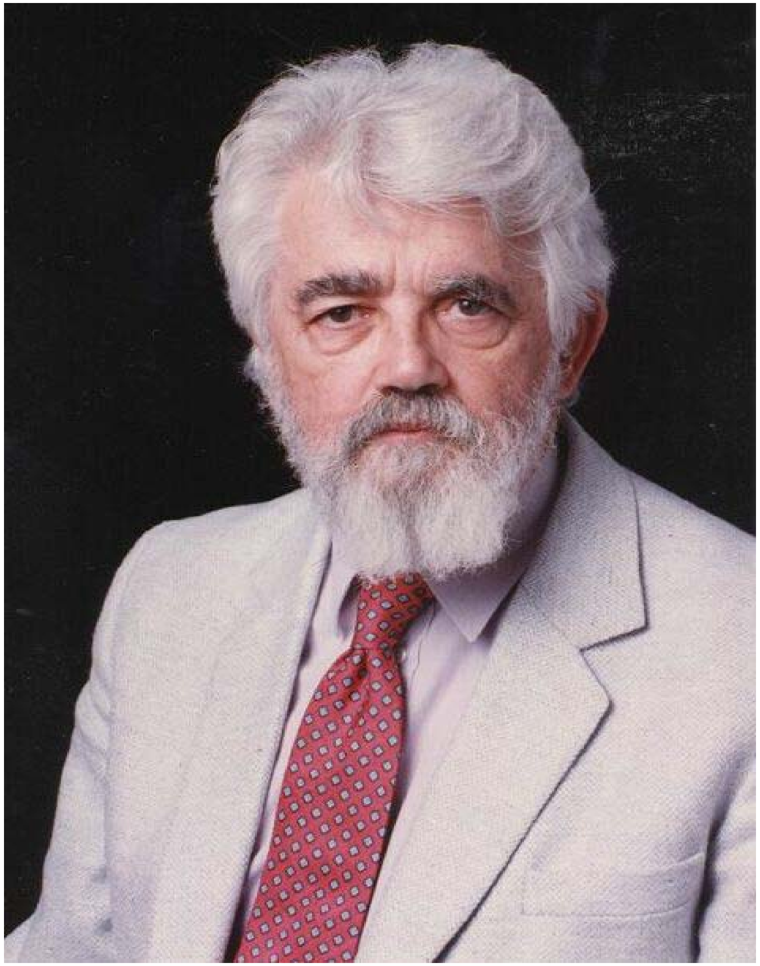
人类水平的人工智能

INSTITUTE OF COMPUTING TECHNOLOGY

- 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



通用人工智能 (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

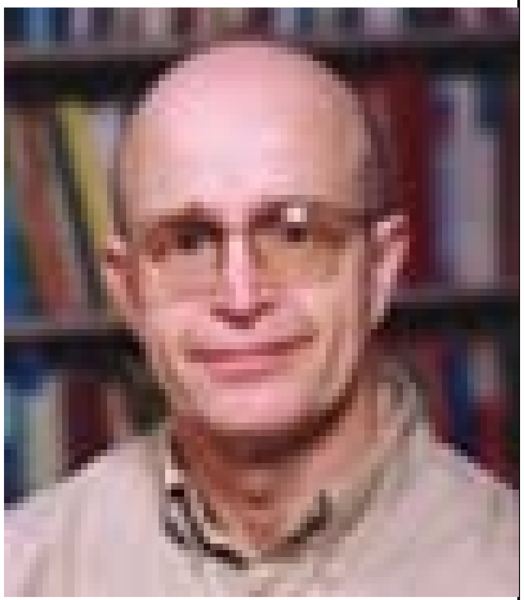
通用人工智能 (AGI)

中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY

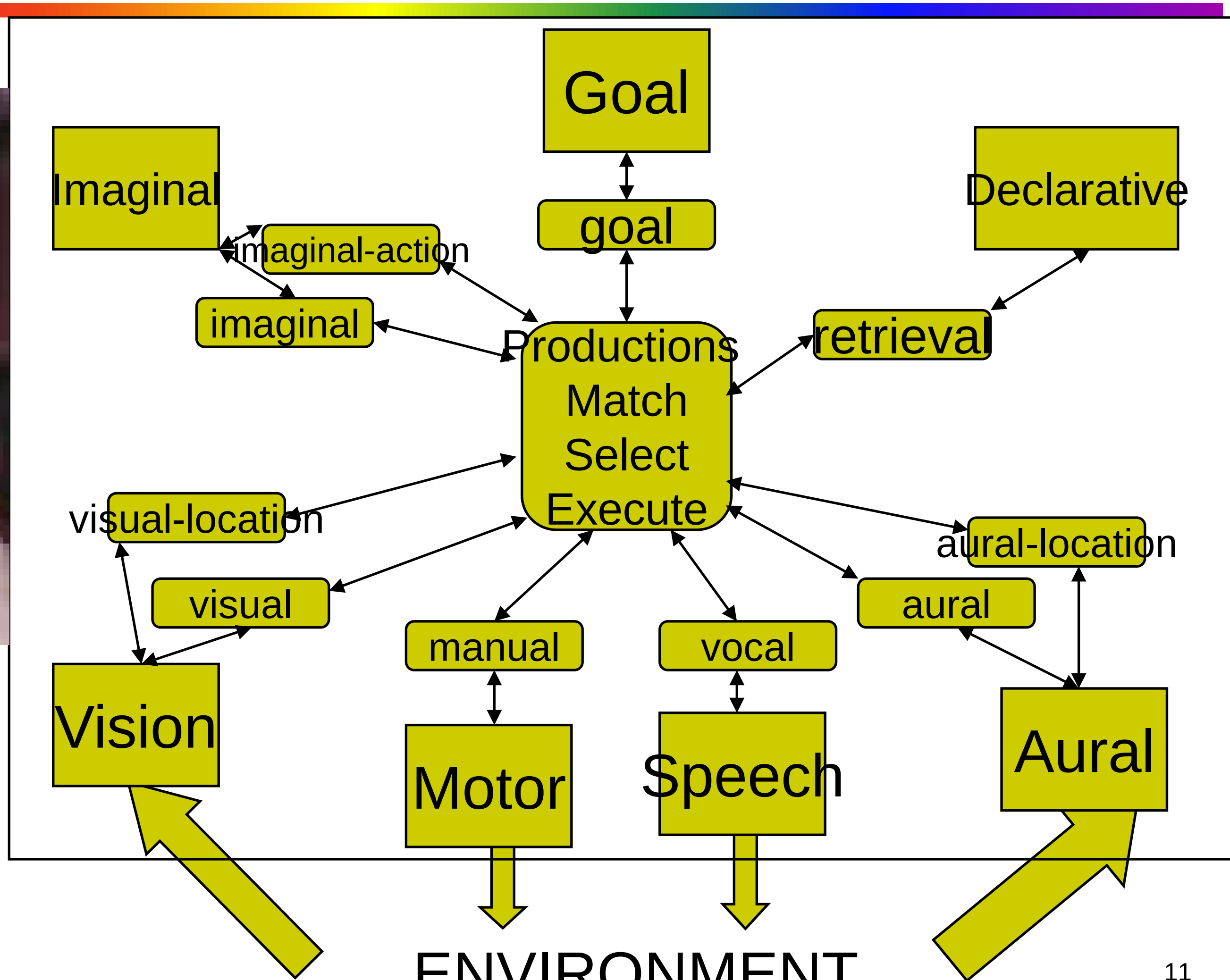
- AGI-2010: Lugano, Switzerland, March 5-8 (In Memoriam Ray Solomonoff)
- AGI-2009: Arlington, Virginia, March 6-9
- AGI-2008: University of Memphis, March 1-3

ACT-R 6.0

中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY



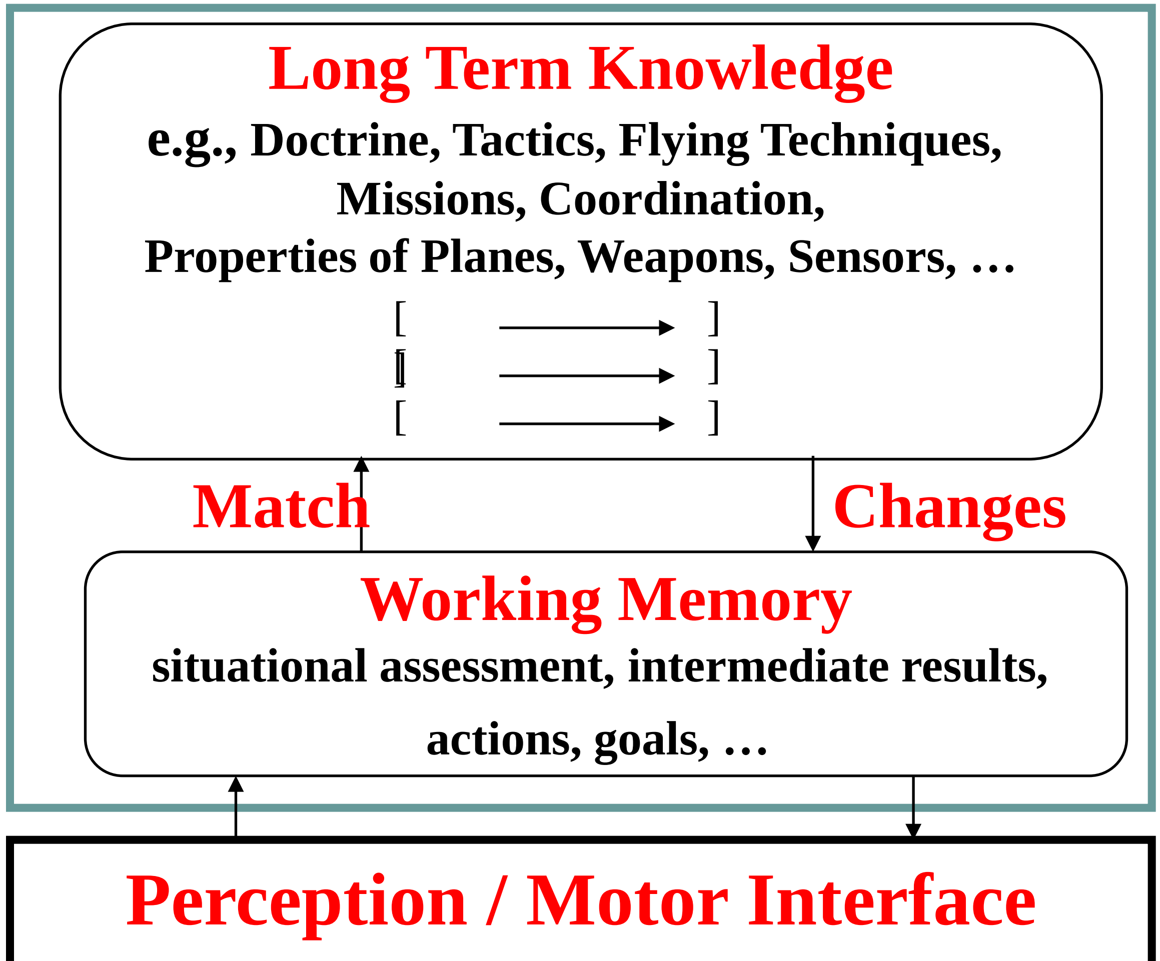
J.A. Anderson



Soar Architecture

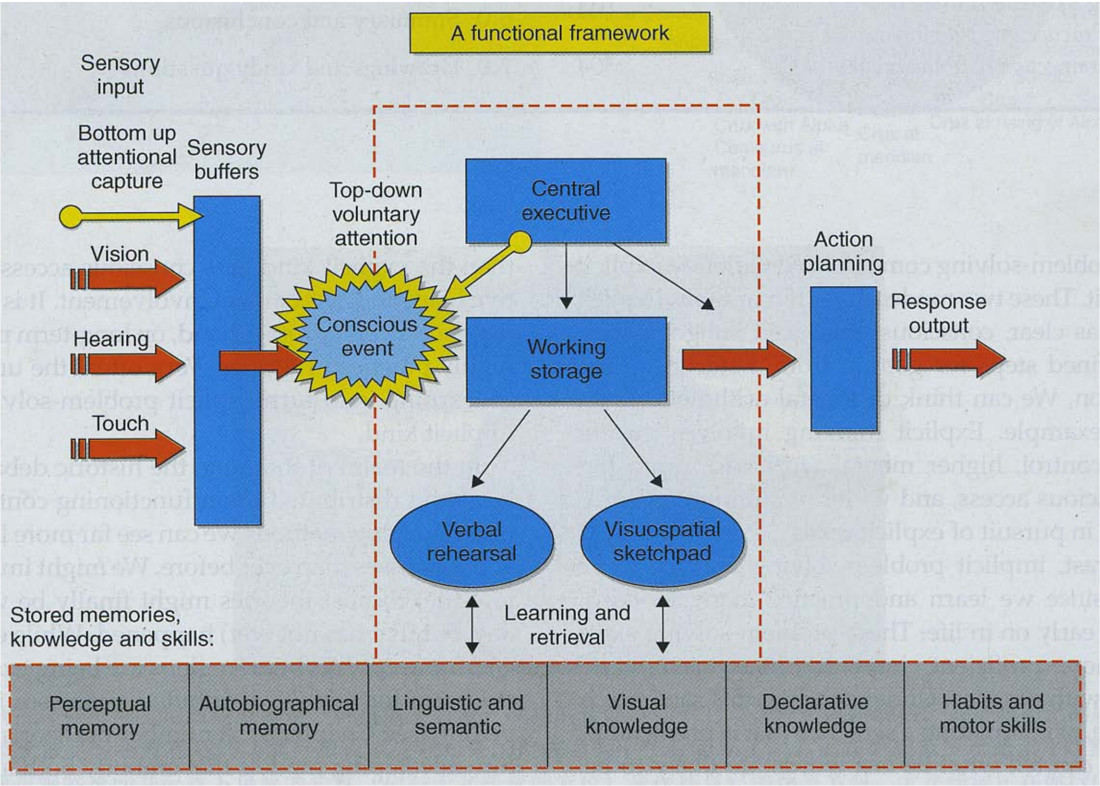
INSTITUTE OF COMPUTING TECHNOLOGY

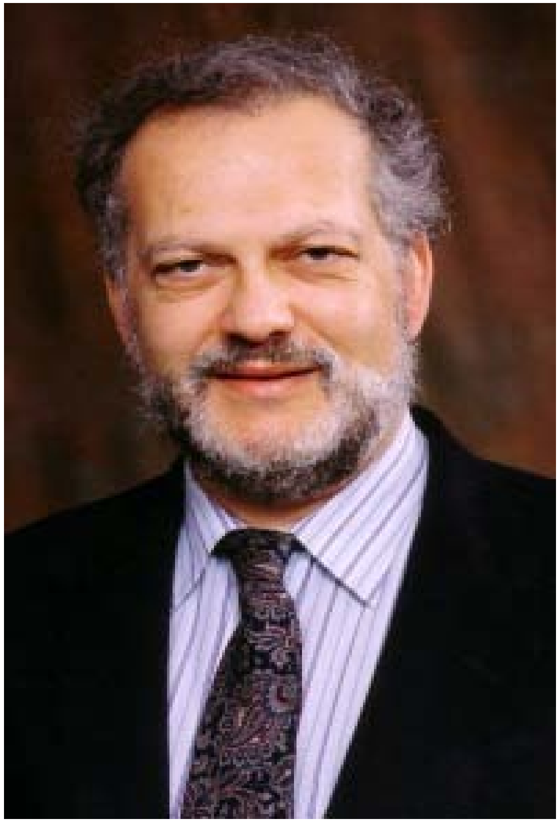
Allen Newell



认知功能框架

INSTITUTE OF COMPUTING TECHNOLOGY





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

内容提要

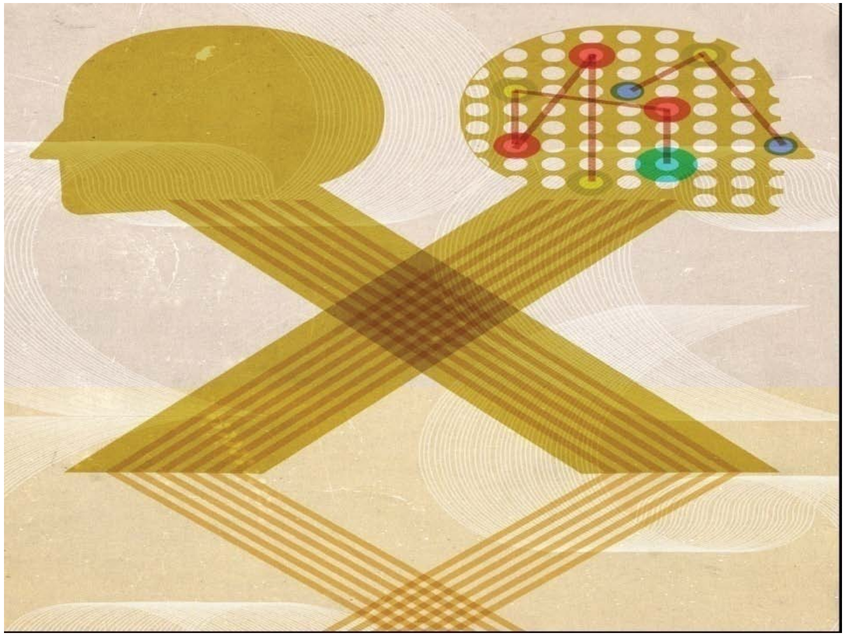
INSTITUTE OF COMPUTING TECHNOLOGY

MoNETA

INSTITUTE OF COMPUTING TECHNOLOGY

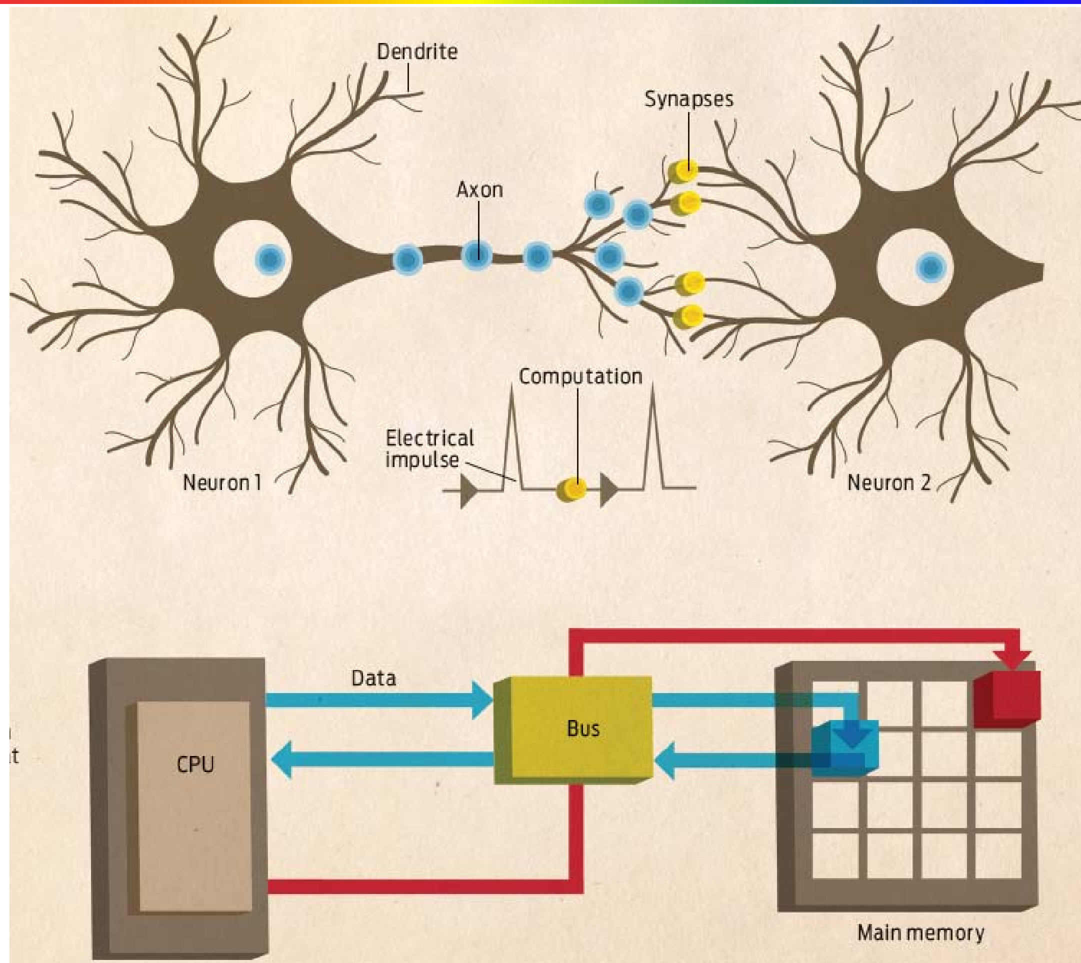
MoNETA (Modular Neural Exploring Traveling Agent)

Boston University's department of cognitive and neural systems, which will run on a brain inspired microprocessor under development at HP Labs in California

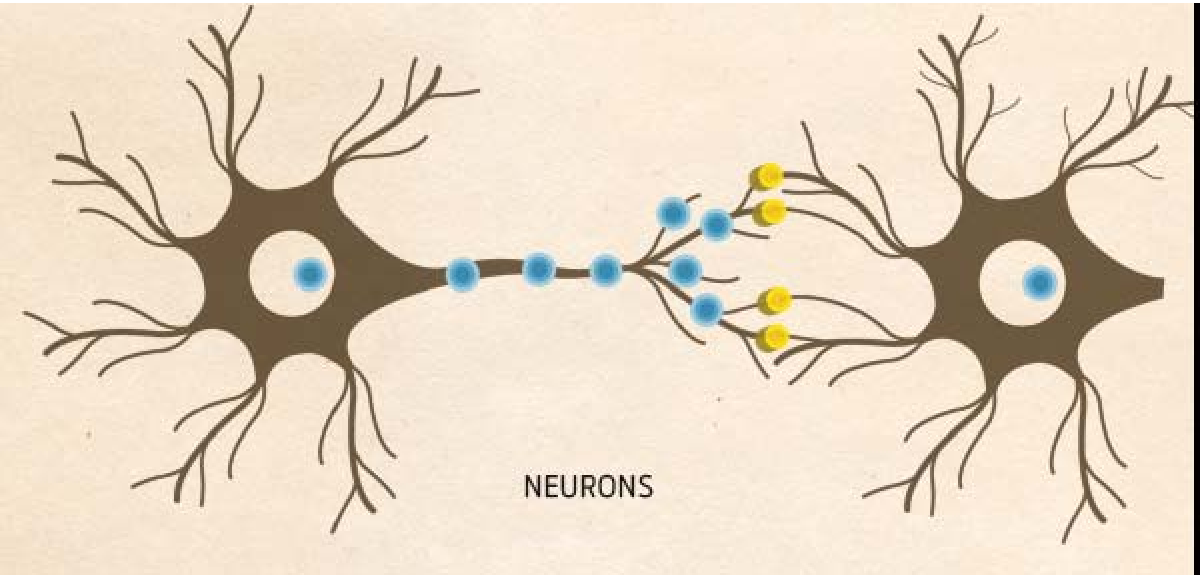


Massimiliano Versace & Ben Chandler . The Brain of a New Machine.IEEE Spectrum》, 12, 2010. 史忠植译,中国计算机学会通讯， No. 9, 2011

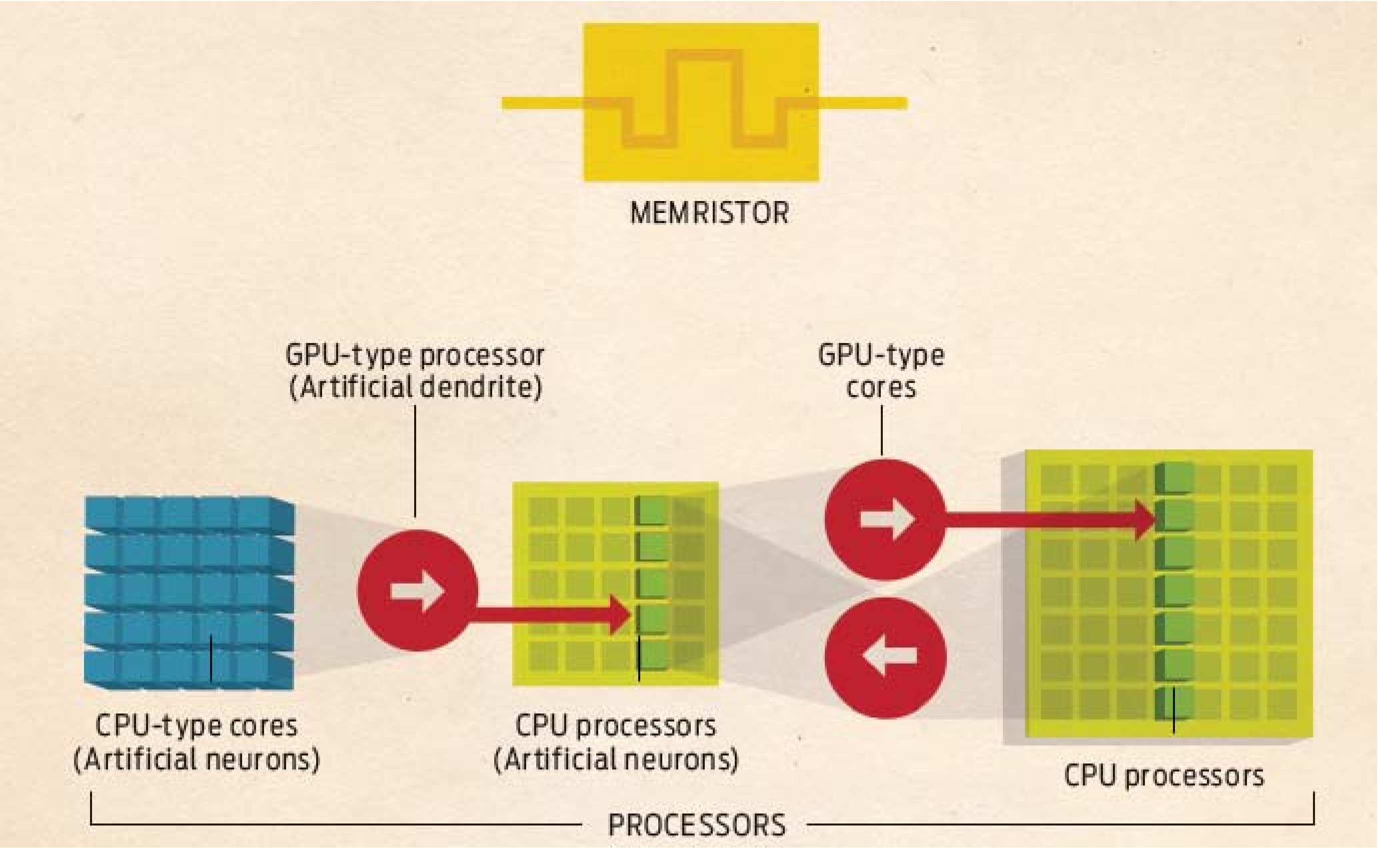
MoNETA


Dendrite
Synapses
Axon
Computation
Electrical impulse
Neuron 1
Neuron 2
Data
Bus
CPU
Main memory

MoNETA

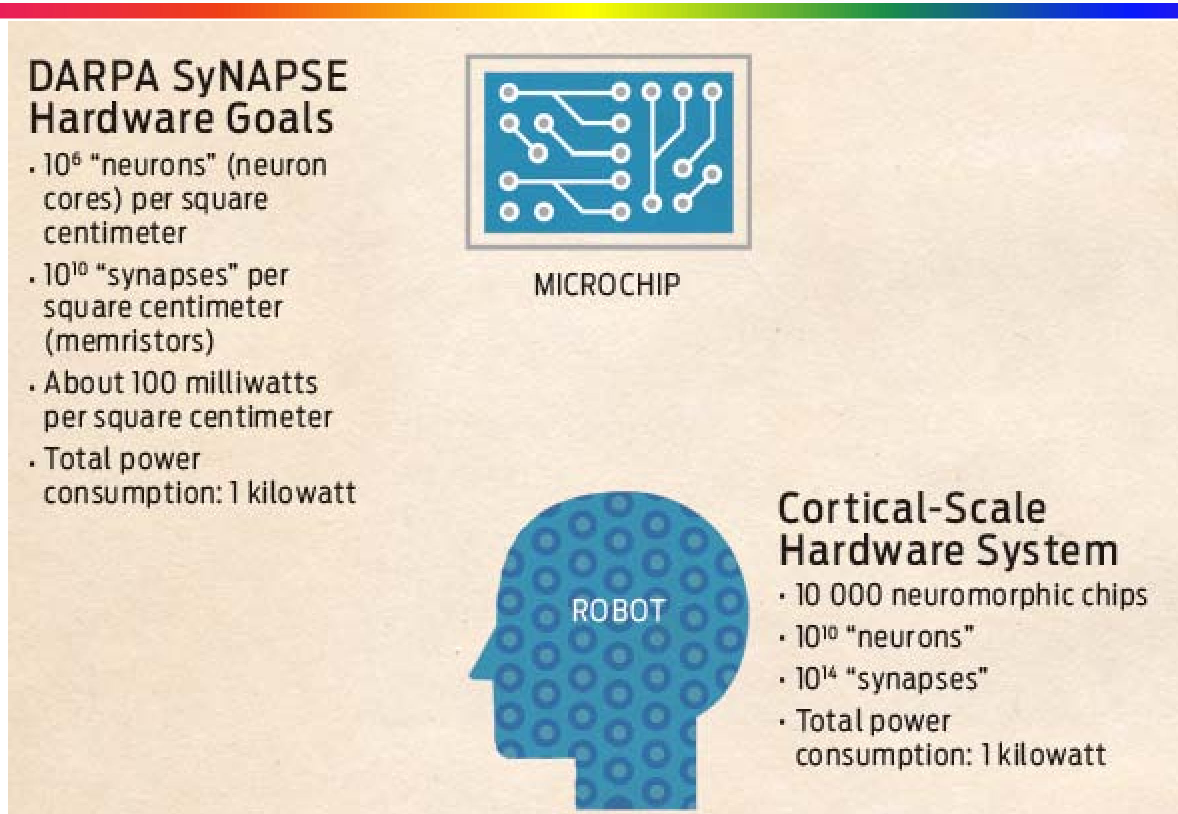


MoNETA



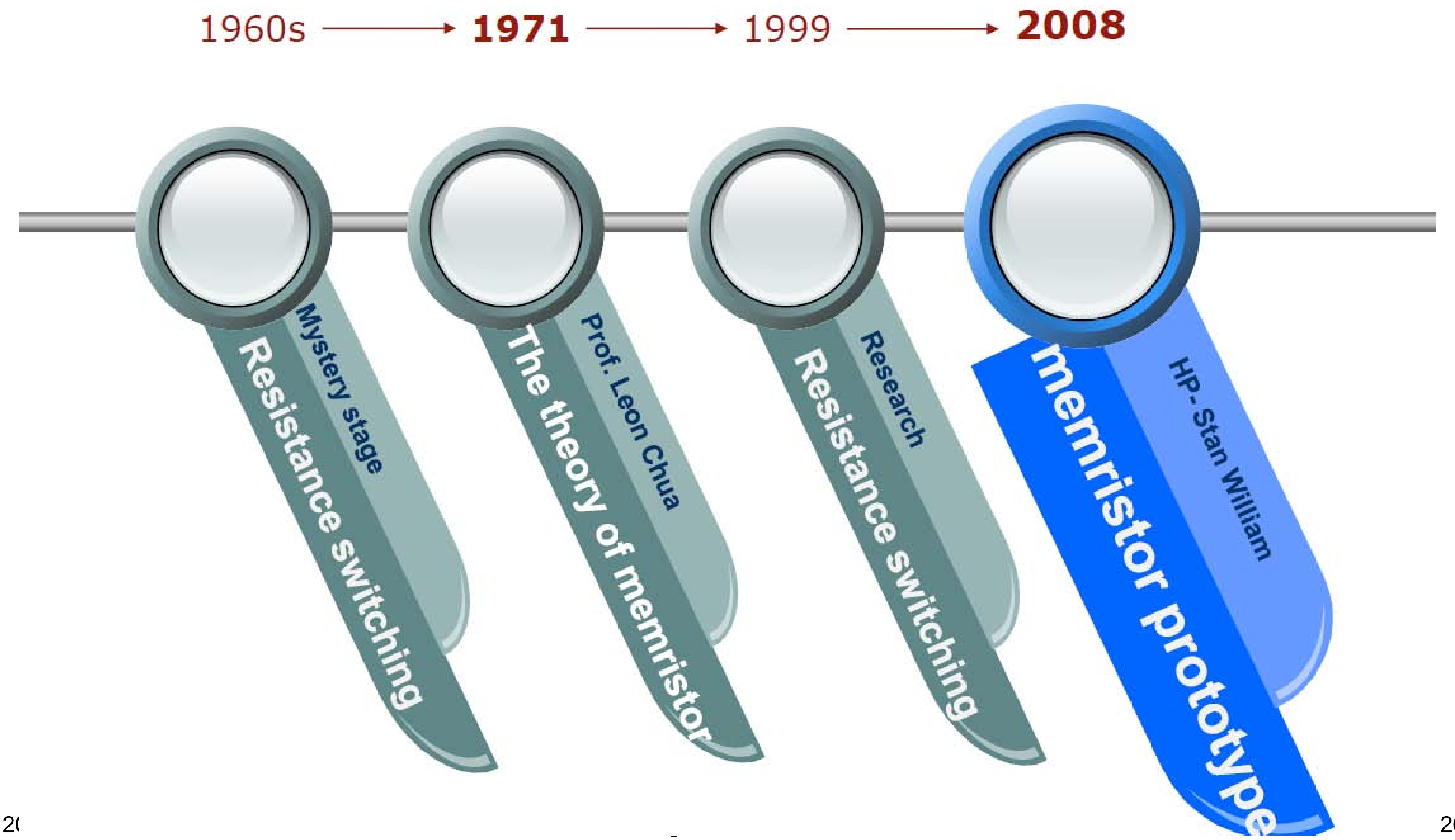
MEMRISTOR
GPU-type processor
(Artificial dendrite)
GPU-type
cores
CPU-type cores
(Artificial neurons)
CPU processors
(Artificial neurons)
CPU processors
PROCESSORS

MoNETA

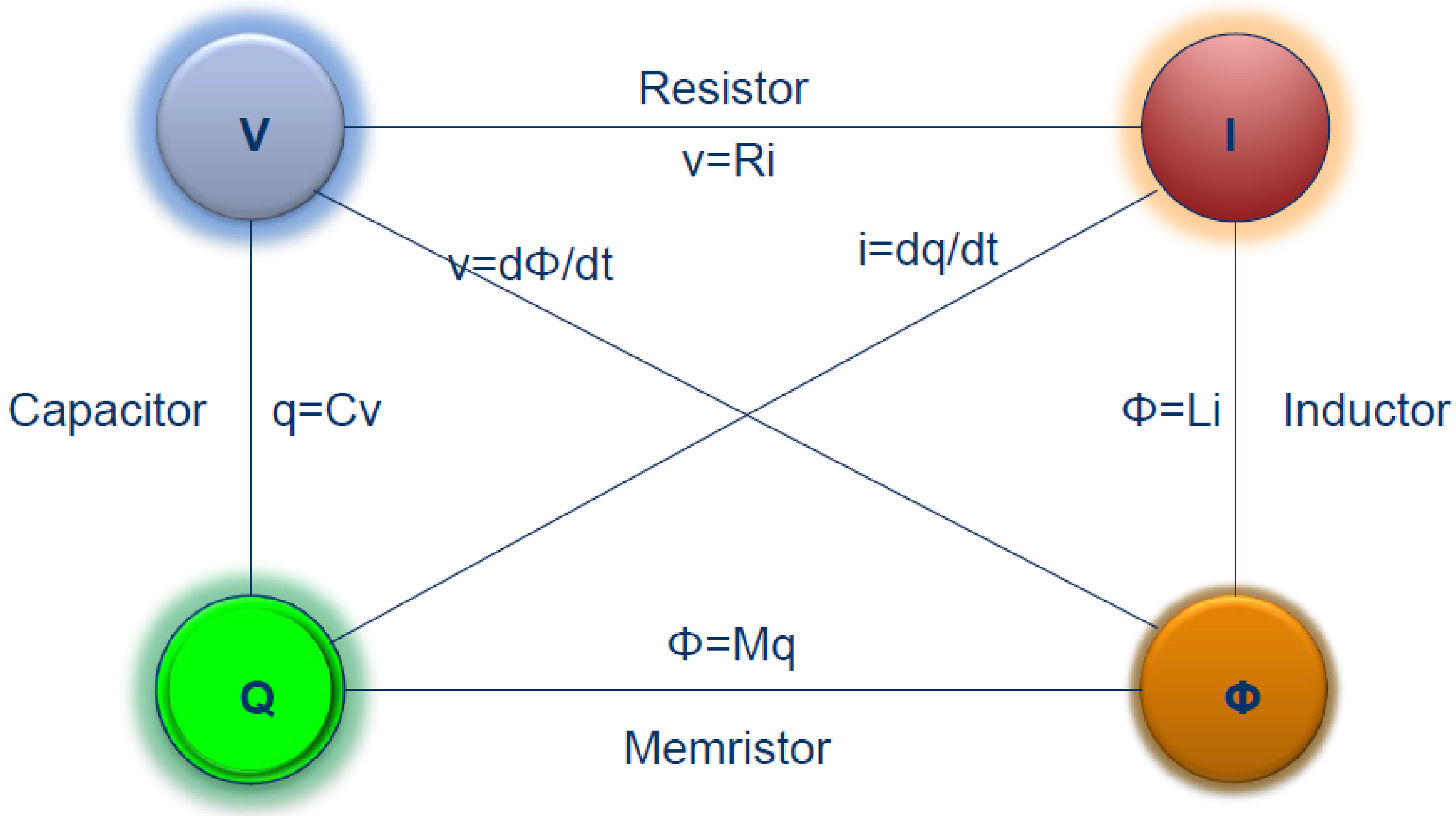

DARPA SyNAPSE Hardware Goals
• 10^6 "neurons" (neuron cores) per square centimeter
• 10^{10} "synapses" per square centimeter (memristors)
• About 100 milliwatts per square centimeter
• Total power consumption: 1 kilowatt
MICROCHIP
ROBOT
Cortical-Scale Hardware System
• 10 000 neuromorphic chips
• 10^{10} "neurons"
• 10^{14} "synapses"
• Total power consumption: 1 kilowatt

忆阻器MEMRISTOR

INSTITUTE OF COMPUTING TECHNOLOGY



忆阻器MEMRISTOR



忆阻器MEMRISTOR

中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY

足够小

足够便宜

并且效率

类似神经突触的关键特性

内容提要

INSTITUTE OF COMPUTING TECHNOLOGY

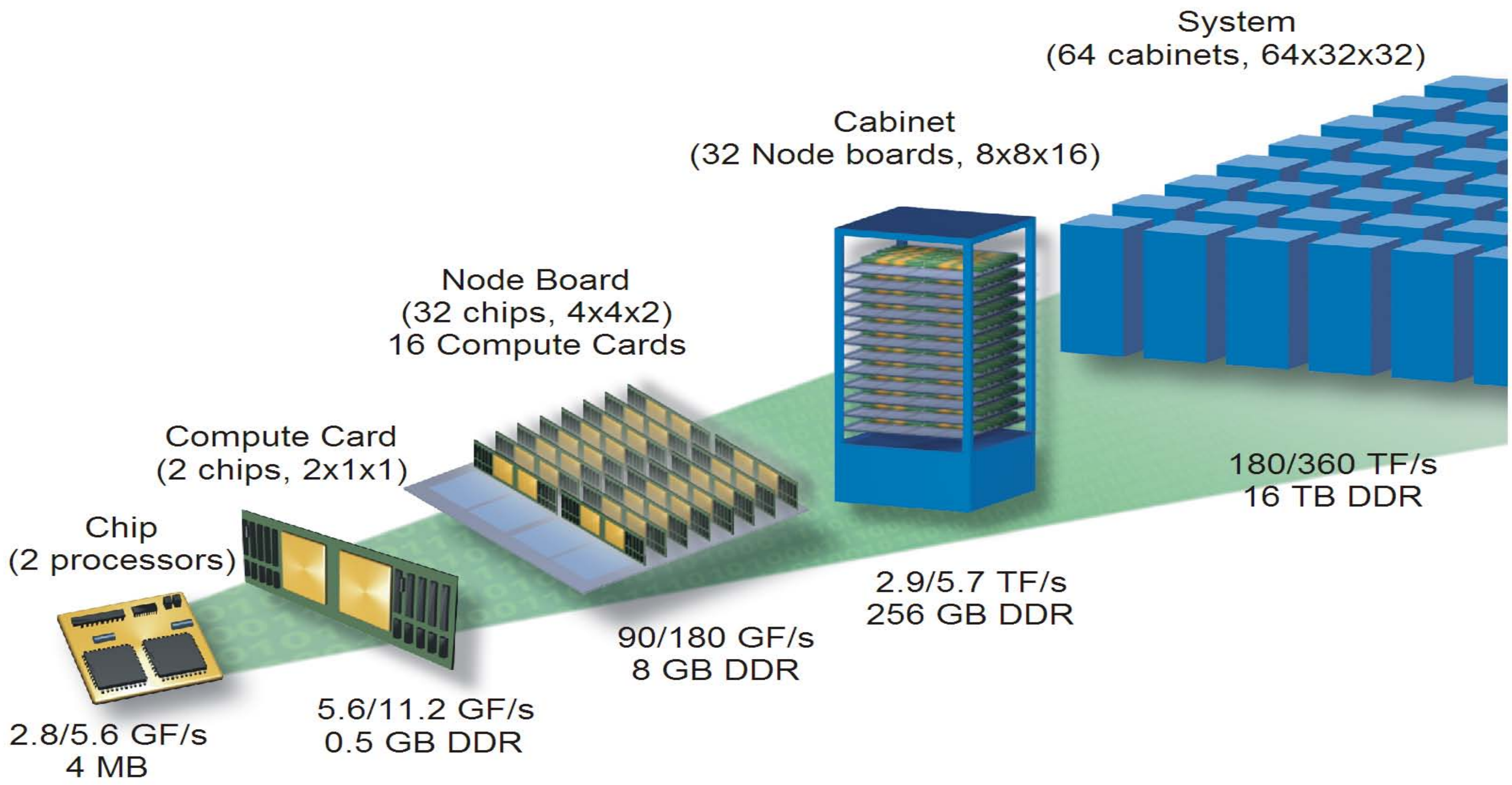
IBM类脑计算机研究

中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY

- 2005年6月19日报道，瑞士洛桑EPFL的脑心智研究所的Henry Markram和IBM公司联合开展蓝脑计划研究。
- 《ACM通讯》2009年2月份报道，IBM 从DARPA得到490万美元的资助，研制类脑计算机。 与IBM Almaden研究中心和IBM T.J. Wason研究中心一起，斯坦福大学，威斯康辛-麦迪逊大学，康奈尔大学，哥伦比亚大学医学中心和加利福尼亚Merced大学都参加该项计划。去年Modha领导的IBM团队使用蓝色基因巨型计算机，模拟具有5500万神经元和5000亿个突触的老鼠大脑。

The Blue Gene/L Architecture

INSTITUTE OF COMPUTING TECHNOLOGY



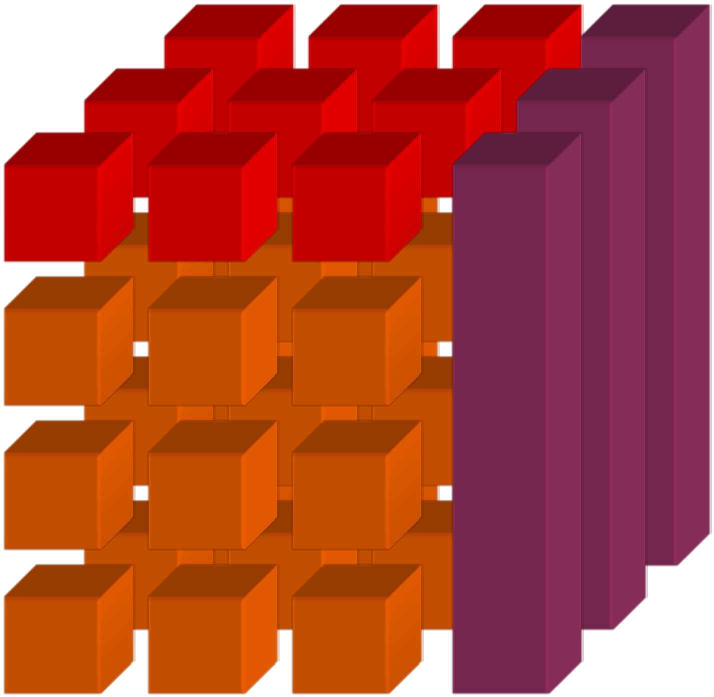
IBM Blue Brain Project

IBM Blue Brain Project(launched in 2005)

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

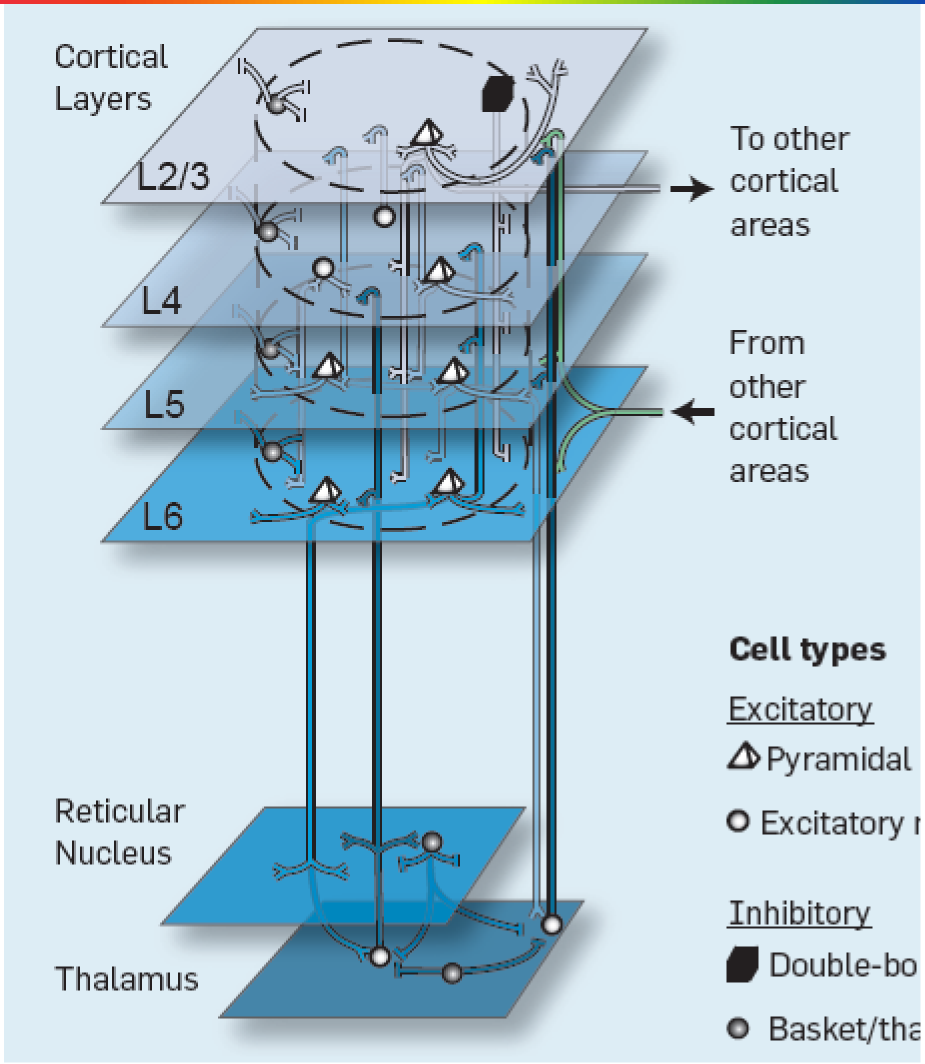
Dharmendra Modha

BlueGene/L System Software Architecture

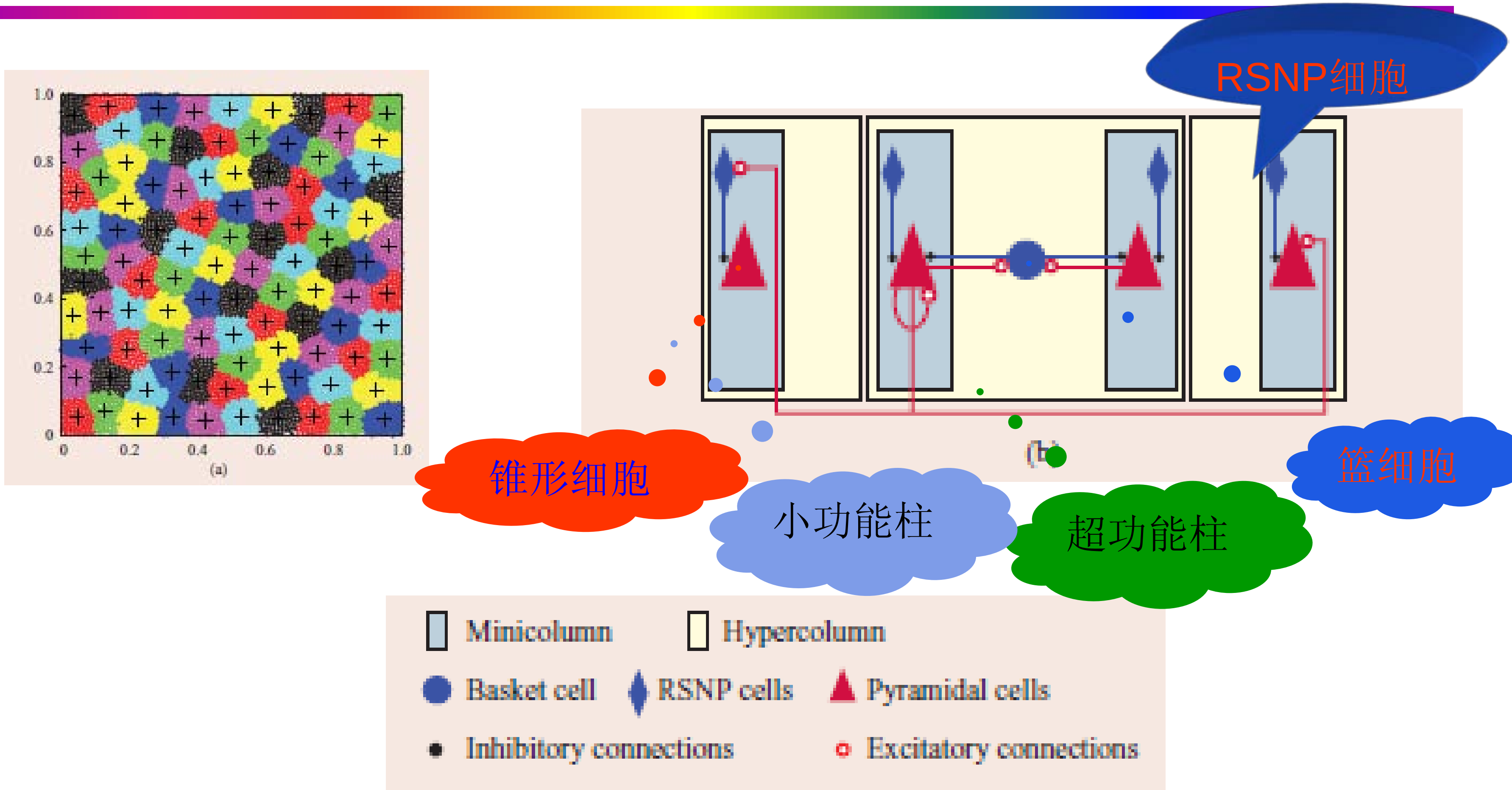


- User applications execute exclusively in the **compute nodes**
  - avoid asynchronous events (e.g., daemons, interrupts)
- The outside world interacts only with the **I/O nodes**, an offload engine
  - standard solution: Linux
- Machine monitoring and control also offloaded to **service nodes**: large SP system or Linux cluster.

IBM类脑计算机研究


Cortical Layers
L2/3
L4
L5
L6
To other cortical areas
From other cortical areas
Reticular Nucleus
Thalamus
Cell types
Excitatory
Pyramidal
Inhibitory

IBM类脑计算机研究



1.0
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1.0
(a)
RSNP细胞
锥形细胞
小功能柱
超功能柱
篮细胞
Minicolumn
Hypercolumn
Basket cell
RSNP cells
Pyramidal cells
Inhibitory connections
Excitatory connections

功能柱

20世纪60年代末，美国科学家发现，在大脑视觉皮层中，具有相同图像特征选择性和相同感受野位置的众多神经细胞，以垂直于大脑表面的方式排列成柱状结构———功能柱。30多年来，脑研究领域一直将垂直的柱状结构看作大脑功能组织的一个基本原则。但是，传统的功能柱研究还不能阐释视觉系统究竟是如何处理大范围复杂图像信息的。

功能柱

- 1972年：Wilson-Cowan方程来描述功能柱；
- 1990年：Shuster等人模拟视皮层中发现的同步振荡；
- 1993年：Jansen等人提出了耦合功能柱模型产生了类EEG波形和诱发电位；
- 1994年：Fukai设计了功能柱式的网络模型来模拟视觉图样的获取；
- 1997年：Hansel等人根据视皮层朝向柱的结构构建了一个超柱模型，研究其中的同步性和混沌特性，并对朝向选择性的功能柱机理做出解释；
- 1998年：Fransén等人把传统网络中的单细胞代换成多细胞构成的功能柱，来模拟工作记忆

功能柱

Rose-Hindmarsh方程来描述单神经元:

$$\dot{x} = y + ax^3 - bx^2 - z + I_{syn} + I_{stim}$$

$$\dot{y} = c - dx^2 - y$$

$$\dot{z} = r\left[s\left(x - x_0\right) - z\right]$$

x: 代表膜电位,
y: 表示快速回复电流,
z: 描述慢变化的调整电流,
I_{syn} 表示突触电流,
I_{stim}表示外界输入

功能柱

模型采用基于电流的突触模型，在突触前细胞的每个动作电位都将触发突触后细胞的I_{syn}输入。突触电流 I_{syn} 表示为:

$$I_{syn} = g_{syn} V_{syn} (e^{-t/\tau_1} - e^{-t/\tau_2})$$

g_{syn} 为膜电导

τ_1, τ_2 时间常数

V_{syn} 表示突触后电位

内容提要

IDA: an Intelligent Distribution Agent

Read personnel data

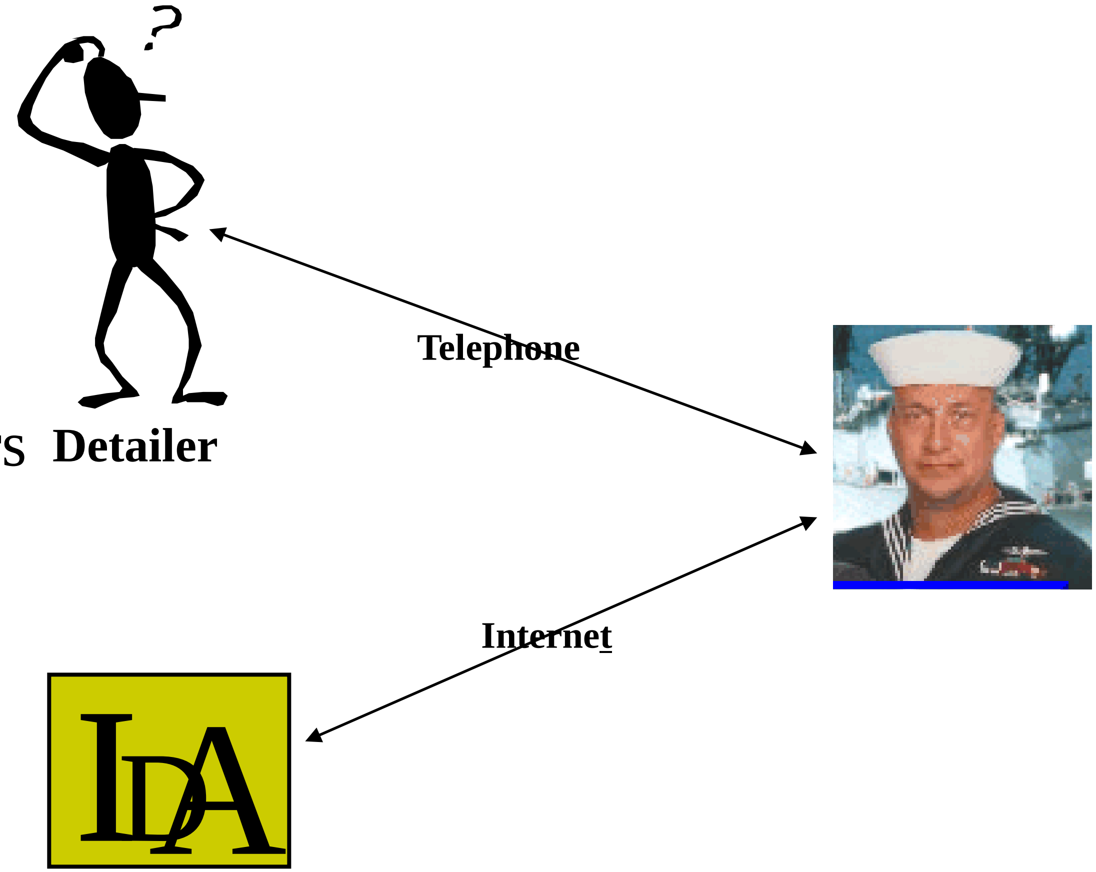
Check job requisition list

Adhere to Navy policies

Choose jobs to offer members

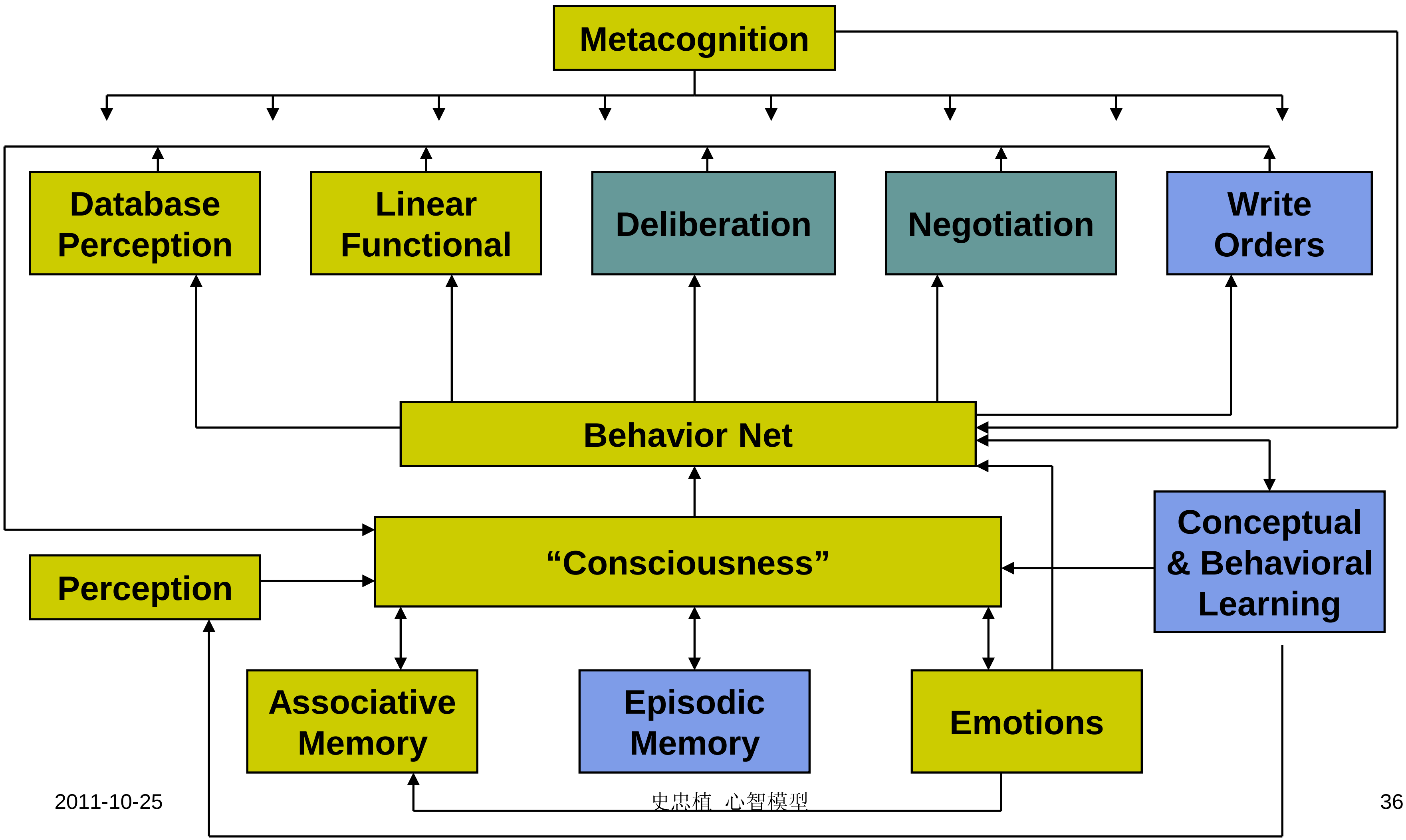
Negotiate with members

Write orders



IDA's Architecture

INSTITUTE OF COMPUTING TECHNOLOGY



Modules and Mechanisms

- Perception—Copycat Architecture—Hofstadter
- Action Selection—Behavior Net—Maes
- Associative Memory—Sparse Distributed Memory—Kanerva
- Episodic Memory—Case-based Memory
- Emotions—Pandemonium Theory—Jackson
- Metacognition—Fuzzy Classifier Systems—Holland, Zadeh
- Learning—Copycat Architecture, Case-based Reasoning
- Constraint Satisfaction—Linear Functional
- Language Generation—Pandemonium Theory
- Deliberation—Pandemonium Theory
- "Consciousness" —Pandemonium Theory

Levels of abstraction

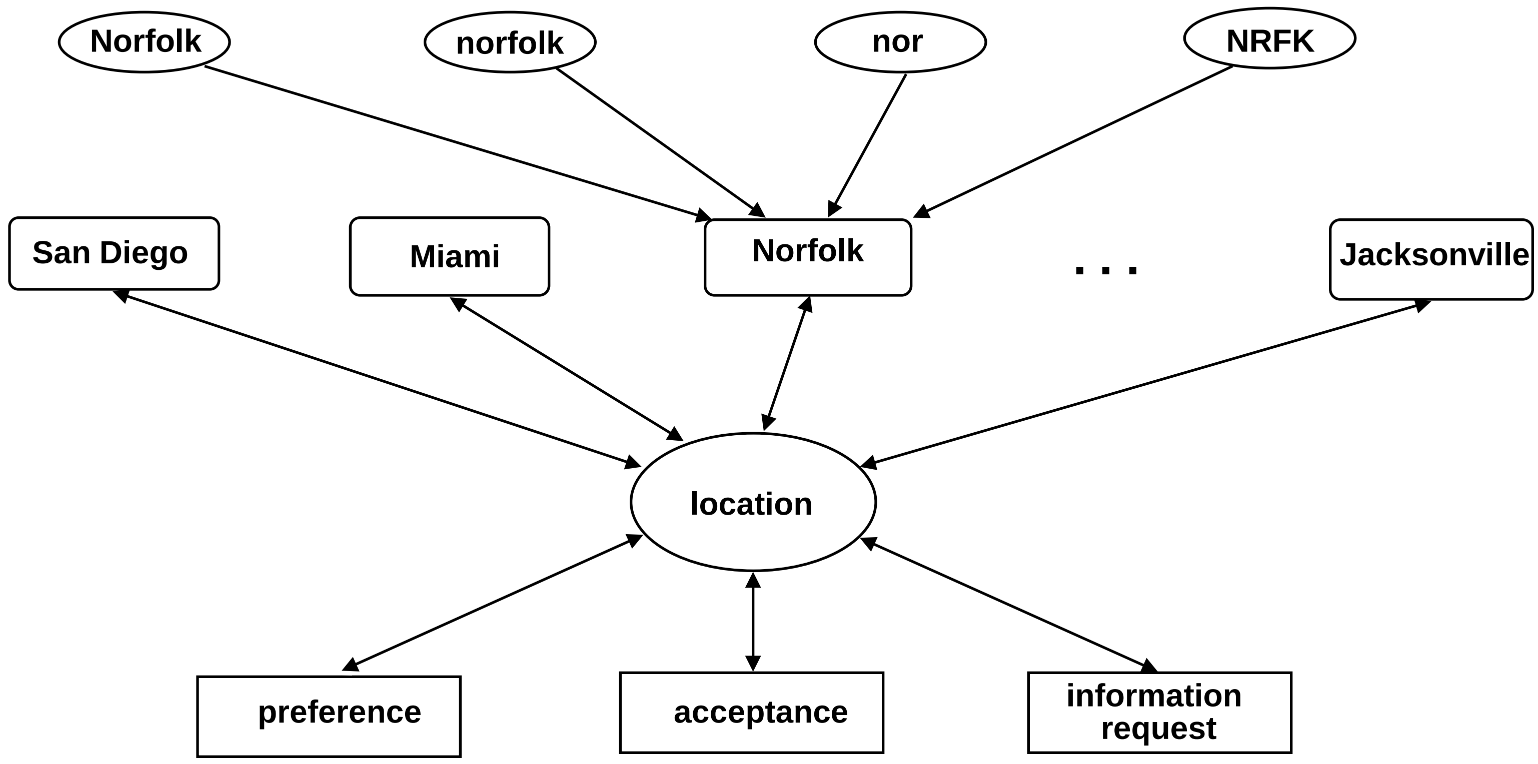
- High level
  - behaviors
  - message type nodes
  - emotions
  - metacognitive actions
  - etc.
- Low level
  - codelets

Codelets

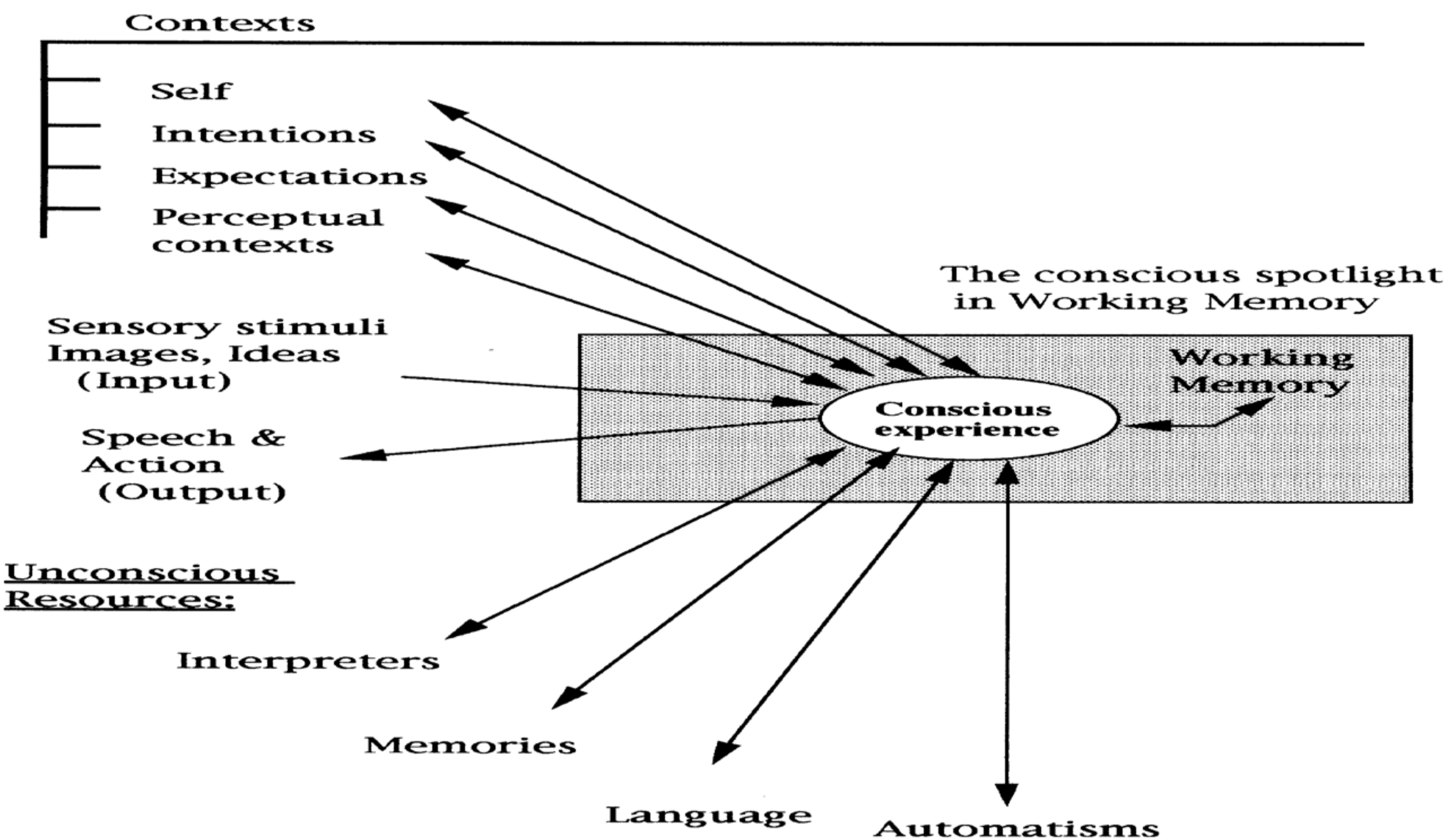
中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY

- Small pieces of code each performing a simple, specialized task
- Acts as a demon, always watching for a chance to act
- Most subserve some high level entity, e.g.
  - behavior
  - slipnet node
  - metacognitive action
- Some codelets work on their own, e.g.
  - watching for incoming mail
  - checking for time and place conflicts
- Codelets do almost all the work
- IDA is a multi-agent system

Perception via a Slipnet

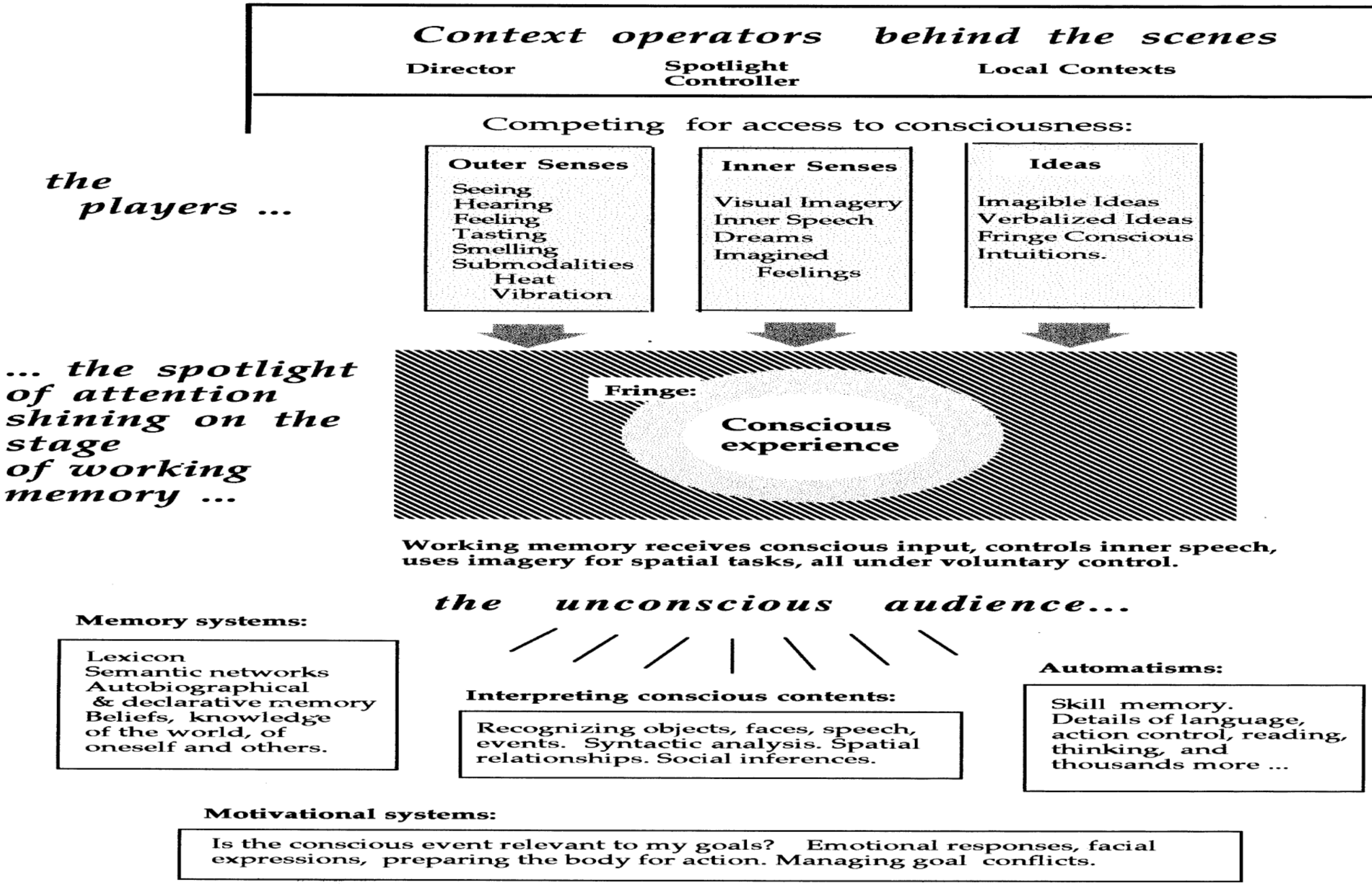


意识GWT



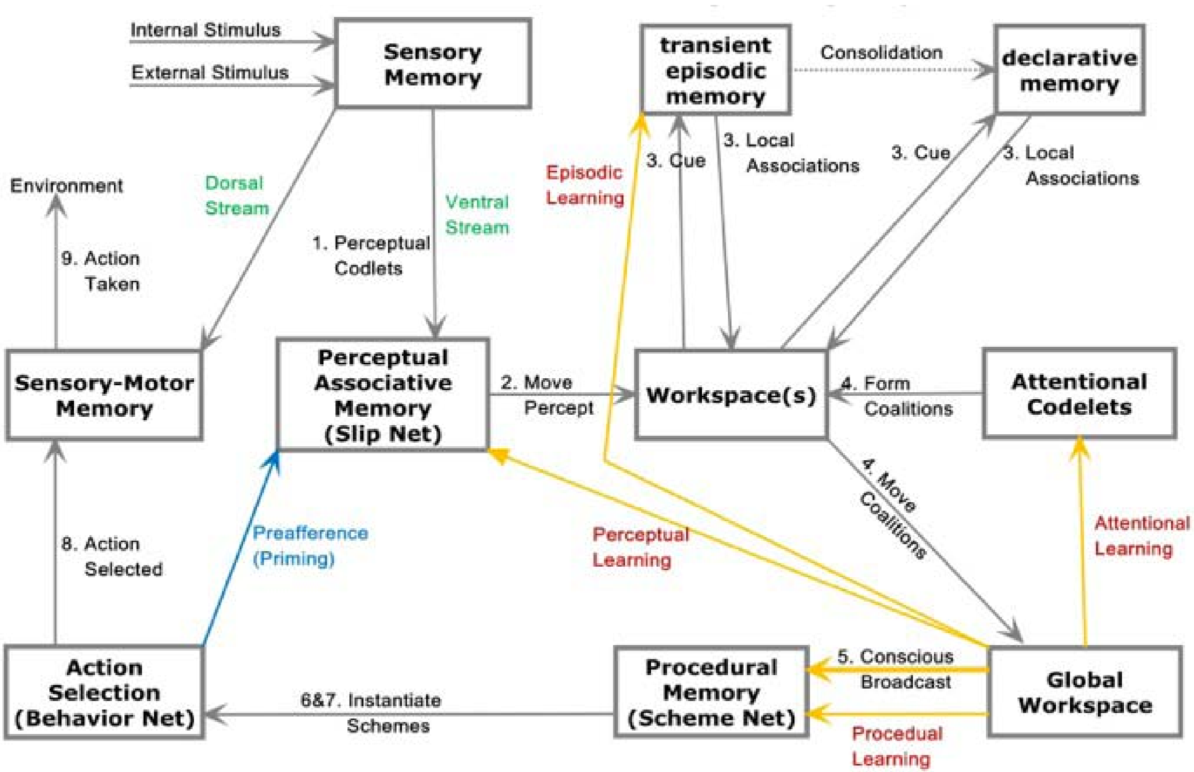
Contexts
Self
Intentions
Expectations
Perceptual contexts
The conscious spotlight in Working Memory
Sensory stimuli Images, Ideas (Input)
Speech & Action (Output)
Working Memory
Conscious experience
Unconscious Resources:
Interpreters
Memories
Language
Automatisms

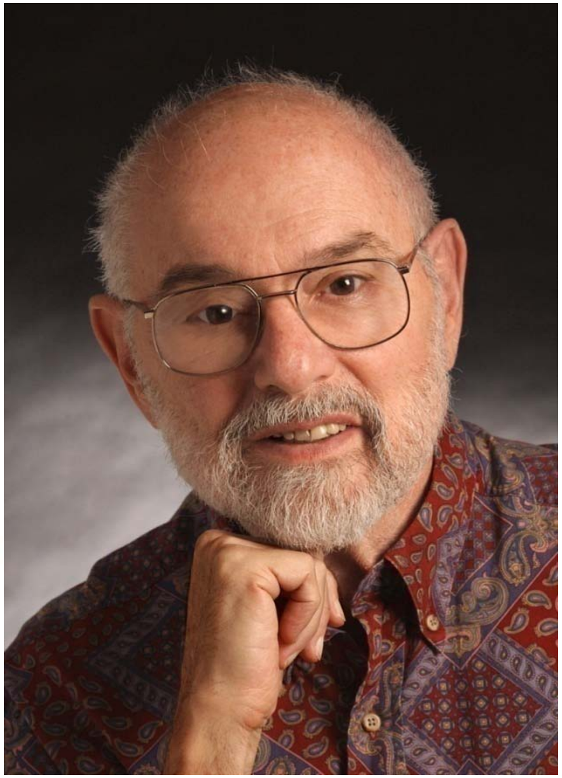
意识GWT


Context operators behind the scenes
Director
Spotlight Controller
Local Contexts
Competing for access to consciousness:
the players ...
Outer Senses
Seeing
Hearing
Feeling
Tasting
Smelling
Submodalities
Heat
Vibration
Inner Senses
Visual Imagery
Inner Speech
Dreams
Imagined Feelings
Ideas
Imagible Ideas
Verbalized Ideas
Fringe Conscious
Intuitions.
... the spotlight of attention shining on the stage of working memory ...
Fringe:
Conscious experience
Working memory receives conscious input, controls inner speech, uses imagery for spatial tasks, all under voluntary control.
the unconscious audience...
Memory systems:
Lexicon
Semantic networks
Autobiographical & declarative memory
Beliefs, knowledge of the world, of oneself and others.
Interpreting conscious contents:
Recognizing objects, faces, speech, events. Syntactic analysis. Spatial relationships. Social inferences.
Automatisms:
Skill memory.
Details of language, action control, reading, thinking, and thousands more ...
Motivational systems:
Is the conscious event relevant to my goals? Emotional responses, facial expressions, preparing the body for action. Managing goal conflicts.

Framework of LIDA

INSTITUTE OF COMPUTING TECHNOLOGY

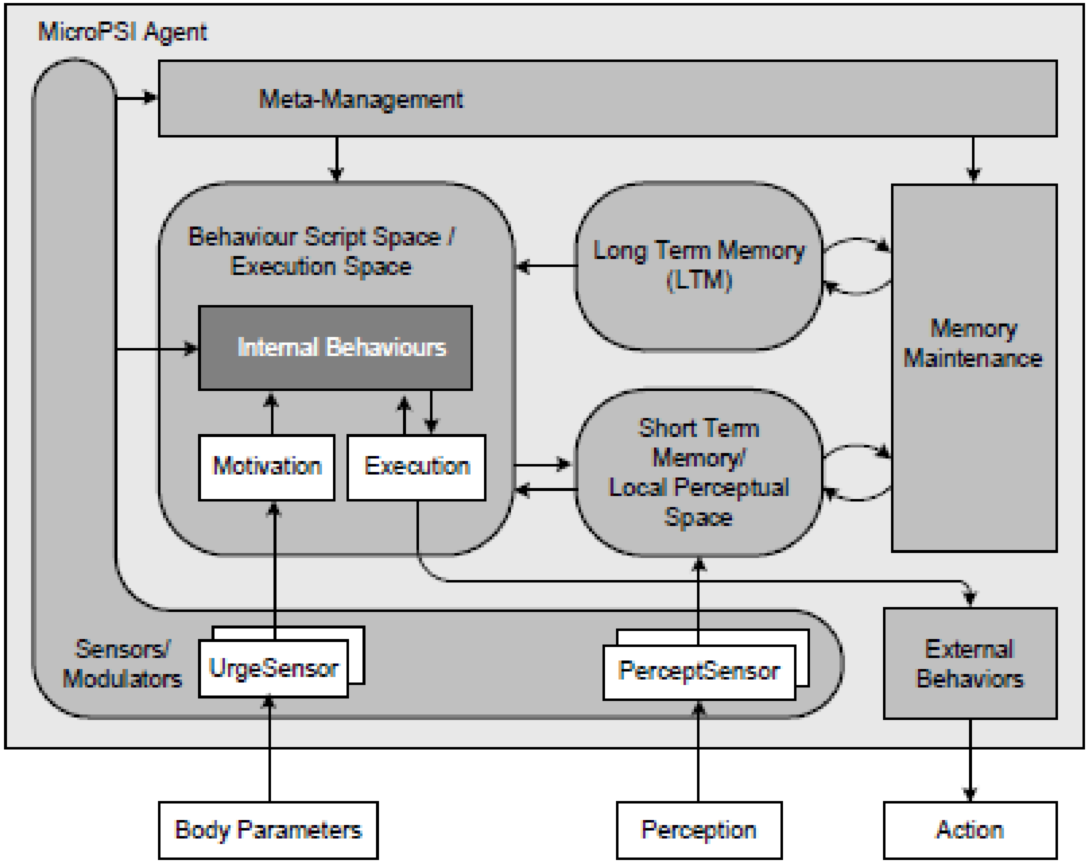




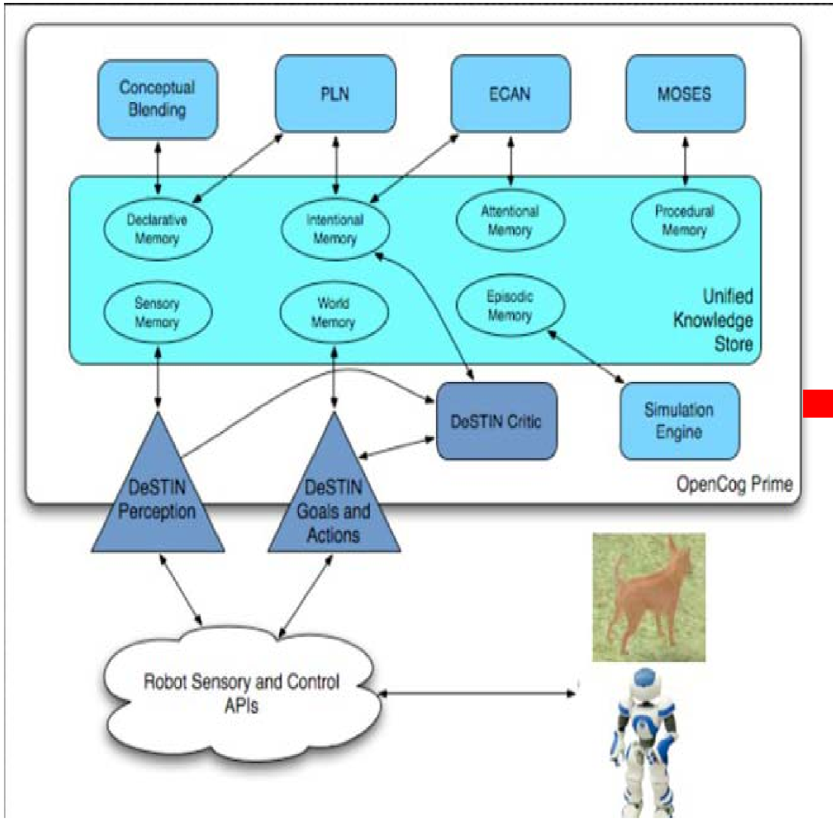
Javier Snaider, Ryan Mccall and Stan Franklin. The LIDA Framework as a General Tool for AGI. AGI2011

MicroPsi Architecture

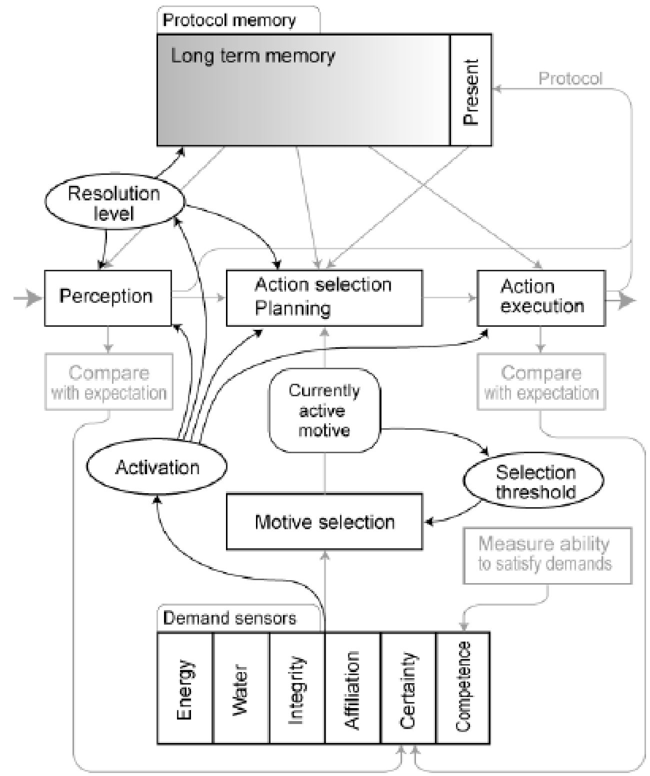
Dr. Joscha Bach



OpenPsi



OpenCog



OpenPsi

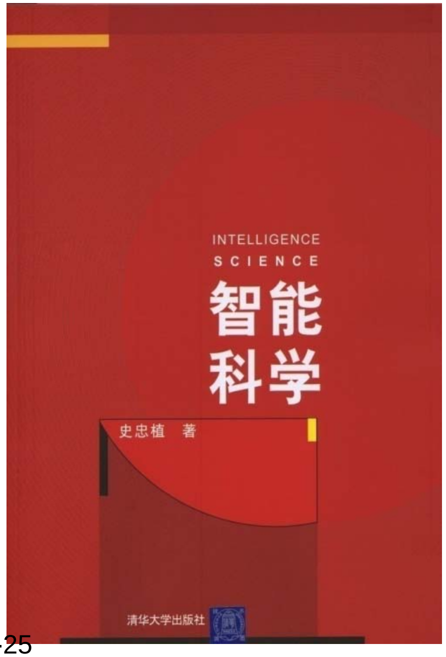
Dr. Ben Goertzel et al.

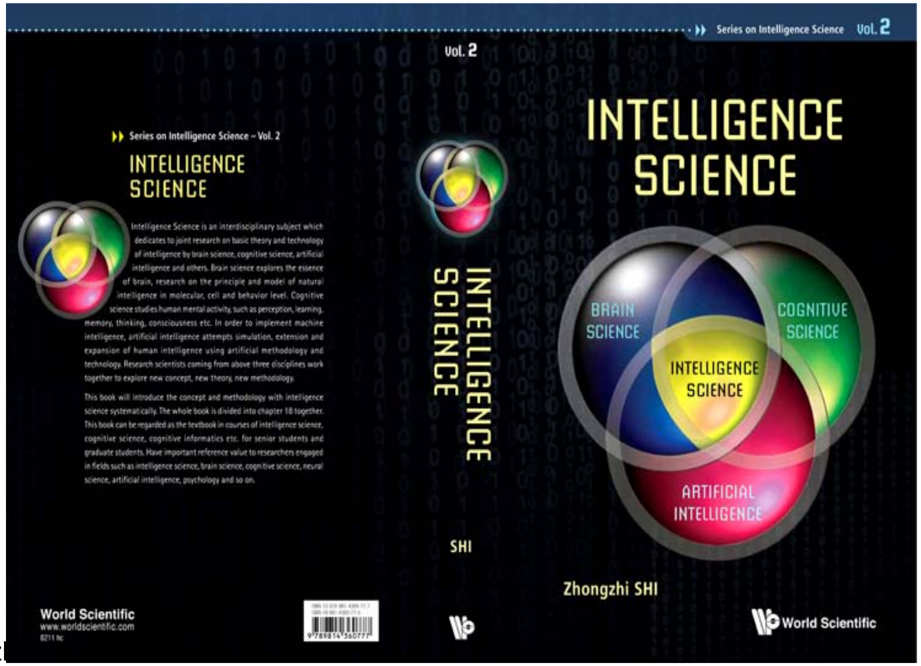
内容提要

- 引言
- MoNETA
- 大型机模拟脑功能
- LIDA
- 心智模型CAM
- 展望

Intelligence Science

- 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.

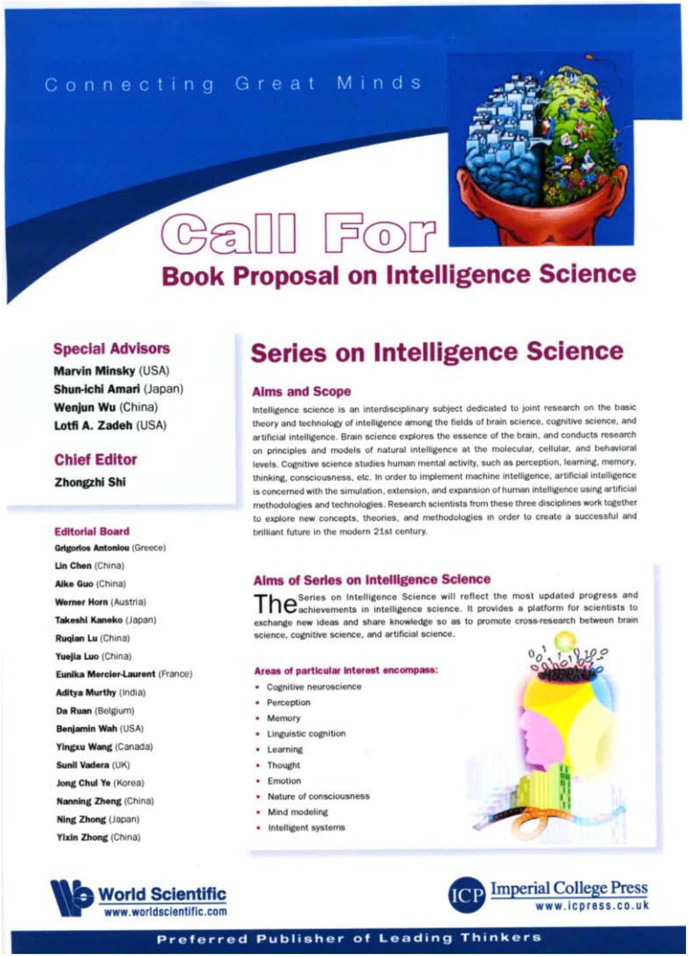
Special Advisors：

Marvin Minsky (USA)

Shun-ichi Amari (Japan)

Wenjun Wu (China)

Lotfi A. Zadeh (USA)



Basic Issues of Intelligence Science

中科院计算所
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*

Mind Model CAM

INSTITUTE OF COMPUTING TECHNOLOGY

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

A new mind modeling for human cognitive activities

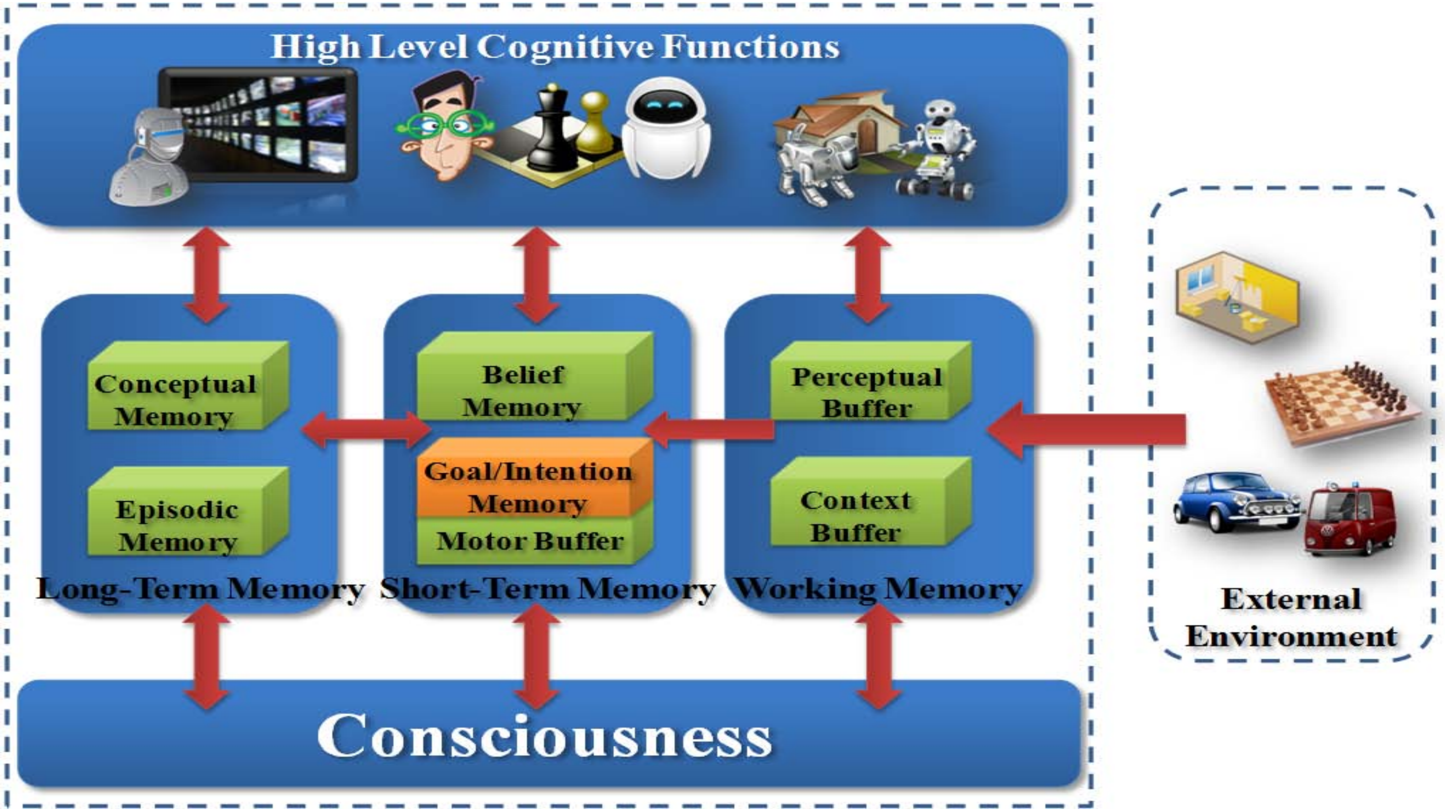
Characteristics of CAM

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

CAM Architecture

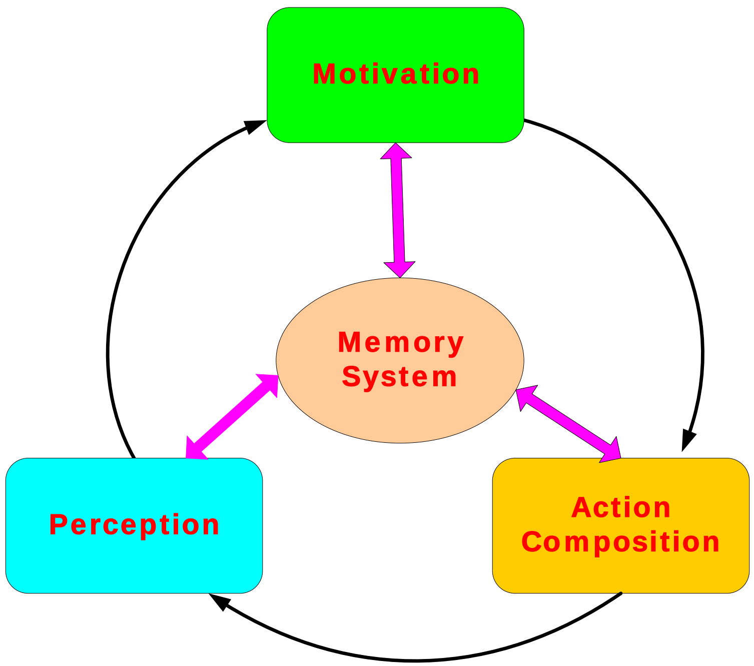
INSTITUTE OF COMPUTING TECHNOLOGY

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



Cognitive Cycle of CAM

中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY



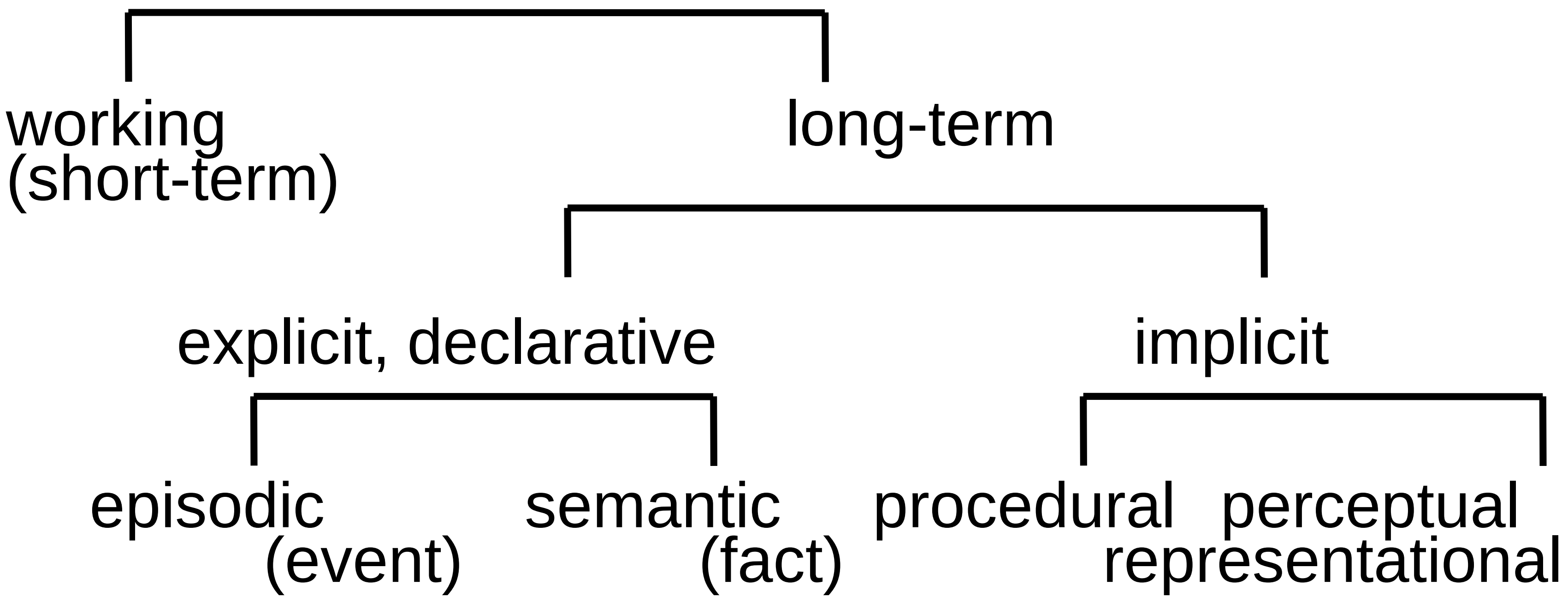
Cognitive cycle of CAM

Long-Term Memory

INSTITUTE OF COMPUTING TECHNOLOGY

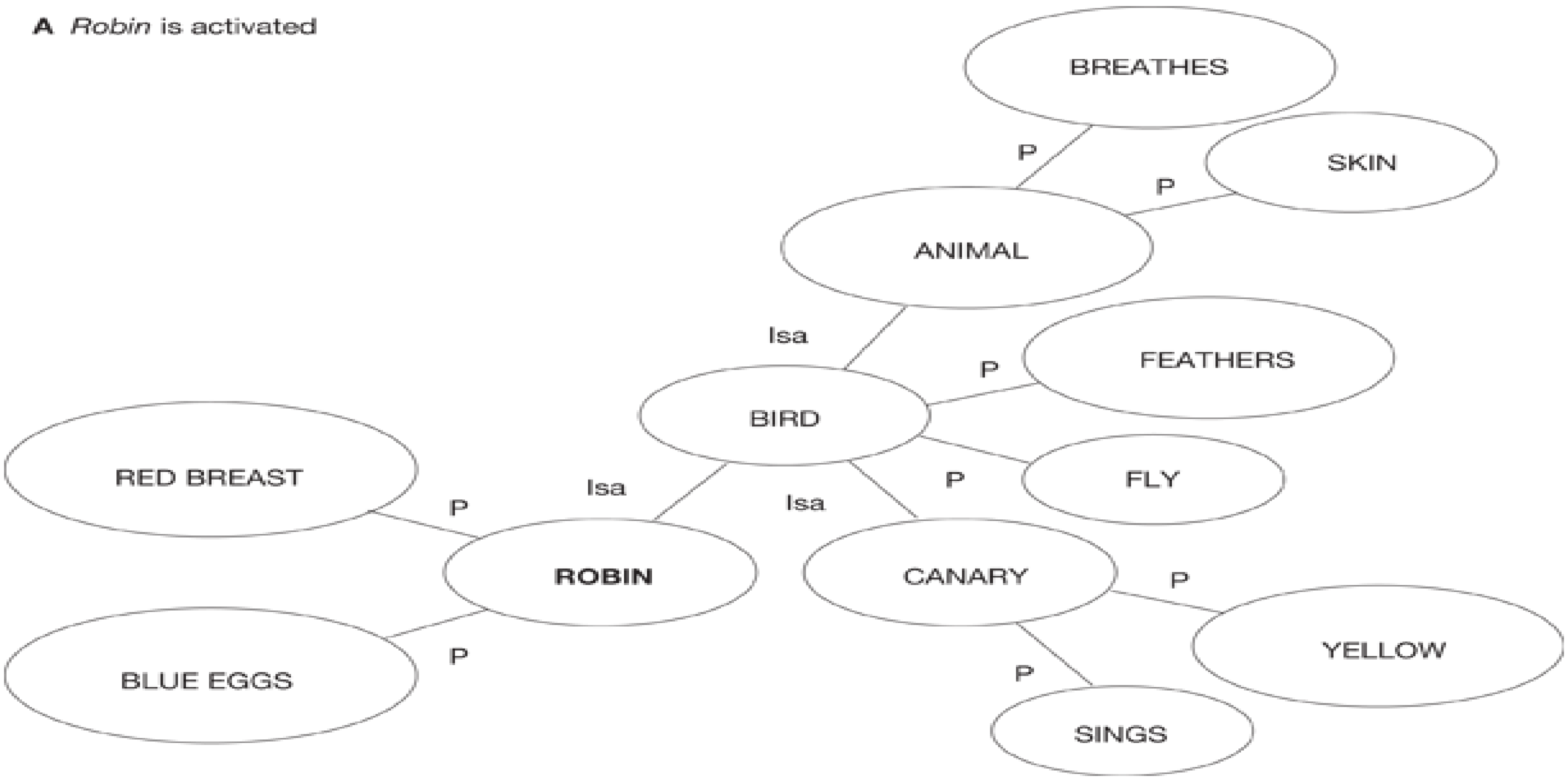
Tulving, Endel (1972). Episodic and semantic memory. Organization of memory. In E. Tulving and W. Donaldson. New York, Academic Press.

Types of memory



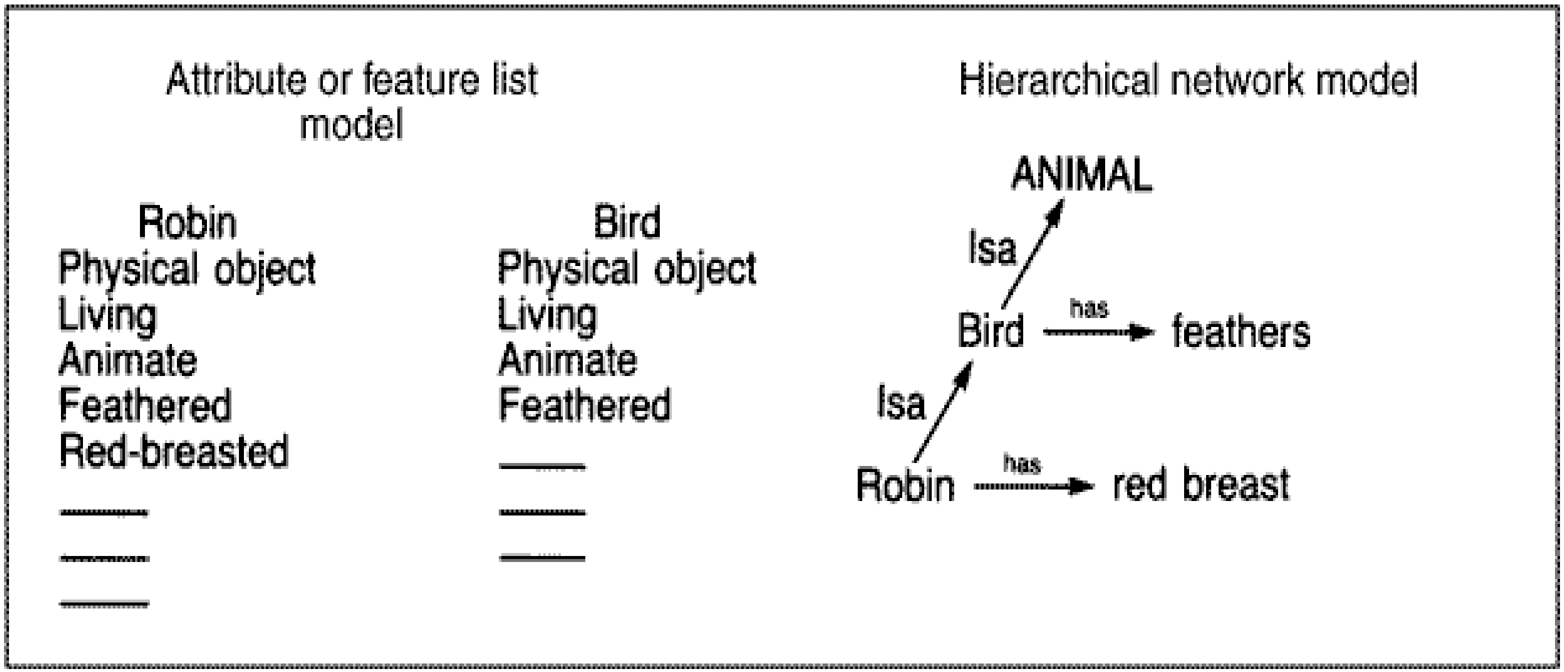
Semantic Memory

Collin and Quillian Model



Semantic Memory

Smith’s feature overlap model



Dynamic Description Logic

INSTITUTE OF COMPUTING TECHNOLOGY

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.*

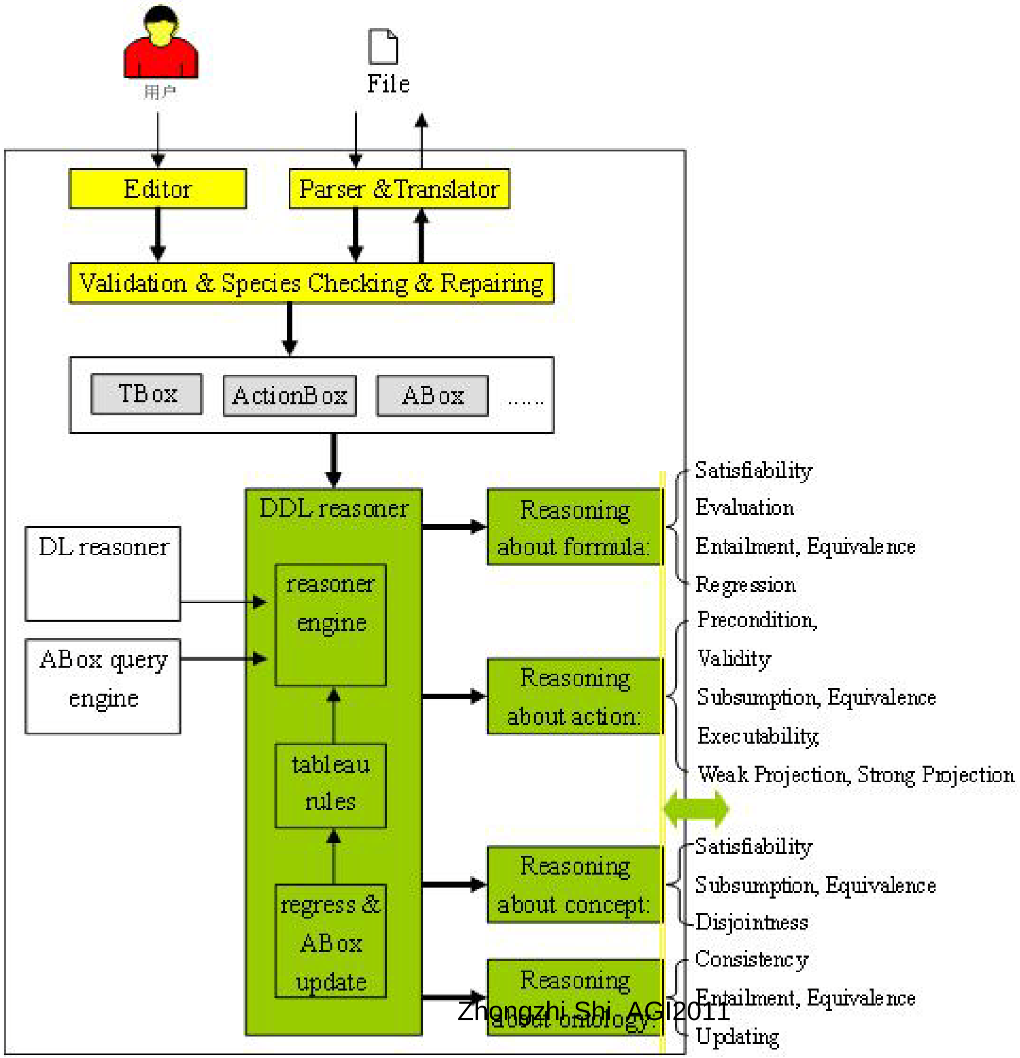
(4) E_A *is the set of results, which denote the effects of the action.*

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: φ, ψ, □, …;**
- **State variable: $u, v, w, \ldots$;**

DDL Inference Engine

INSTITUTE OF COMPUTING TECHNOLOGY



Episodic Memory

- Episodic Memory

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

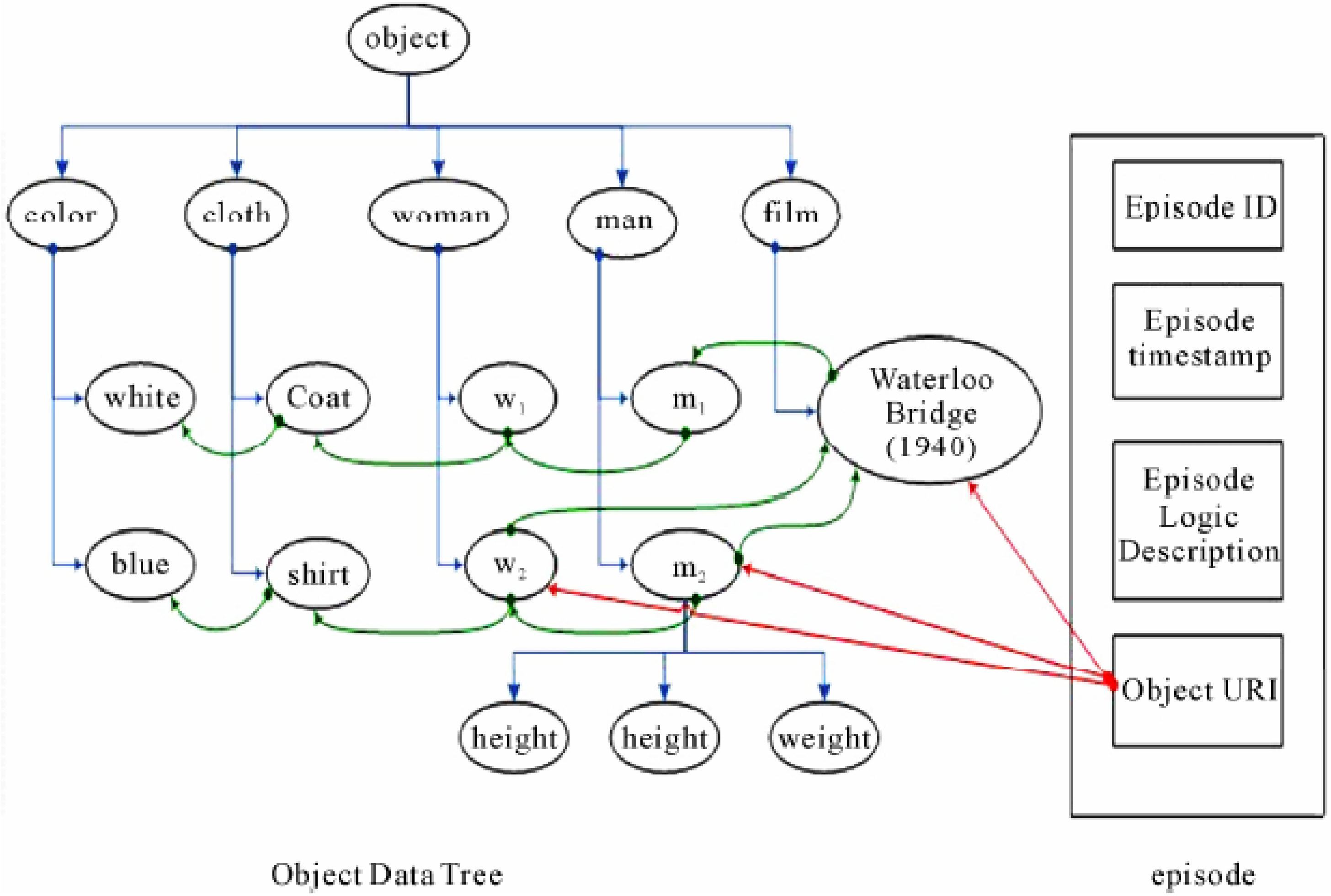
  Example: a movie seen yesterday

- Element in Episodic Memory

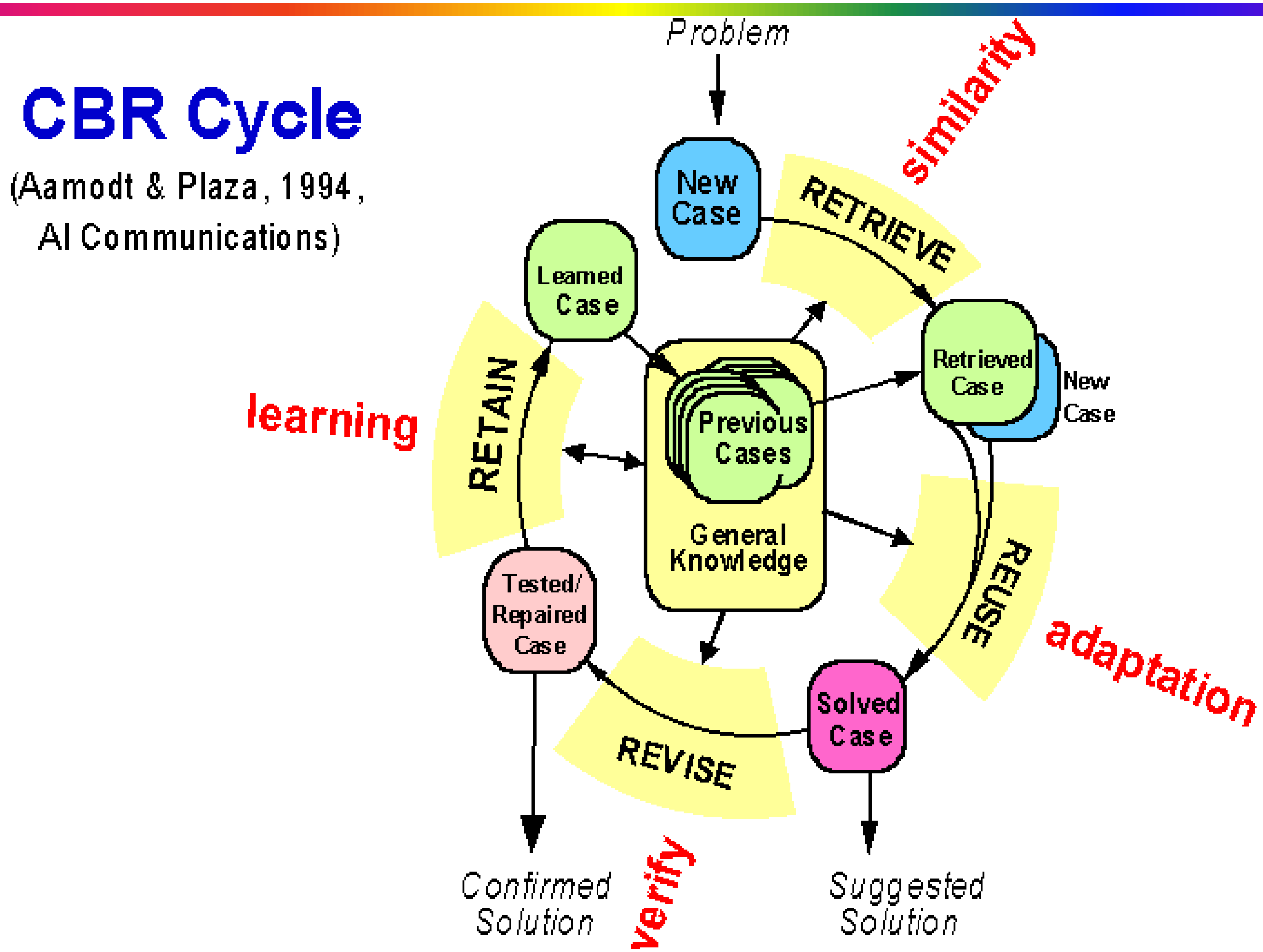
  episode: the stored past event;

  episode in CAM is in form of possible world sequence.

Episodic Memory



Episodic Memory



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

INSTITUTE OF COMPUTING TECHNOLOGY

- Matching Cue with Episode

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

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

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

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

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

Cue Matching Algorithm

Algorithm 1: CueMatching(e,c)

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

if $length(e) < length(c)$ **then**
  return false;
end
$n_e := first\_node(e)$;
$n_c := first\_node(c)$;
if $MatchPossibleWorld(n_e, n_c)$ **then**
  $\alpha_e := Null$;
  $\alpha_c := action(n_c)$;
  if $\neg(Pre(\alpha_e) \longrightarrow Pre(\alpha_c))$ *unsatisfiable according DDL tableau algorithm* **then**
    $n_{tmp} := n_e$;
    while $next\_node(n_{tmp}) \neq Null$ **do**
      $\alpha_e := (\alpha_e; action(n_{tmp}))$;
      if $MatchAction(\alpha_e, \alpha_c)$ **then**
        Let sub_e be the sub sequence by removing α_e from e ;
        Let sub_c be the sub sequence by removing α_c from c ;
        if $CueMatching(sub_e, sub_c)$ **then**
          return true;
        end
      end
      $n_{tmp} := next\_node(n_{tmp})$;
    end
  end
end
Remove n_e from e ;
return $CueMatching(e, c)$;

Match Possible World & Action

Function MatchPossibleWorld(w_i,w_j)

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

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

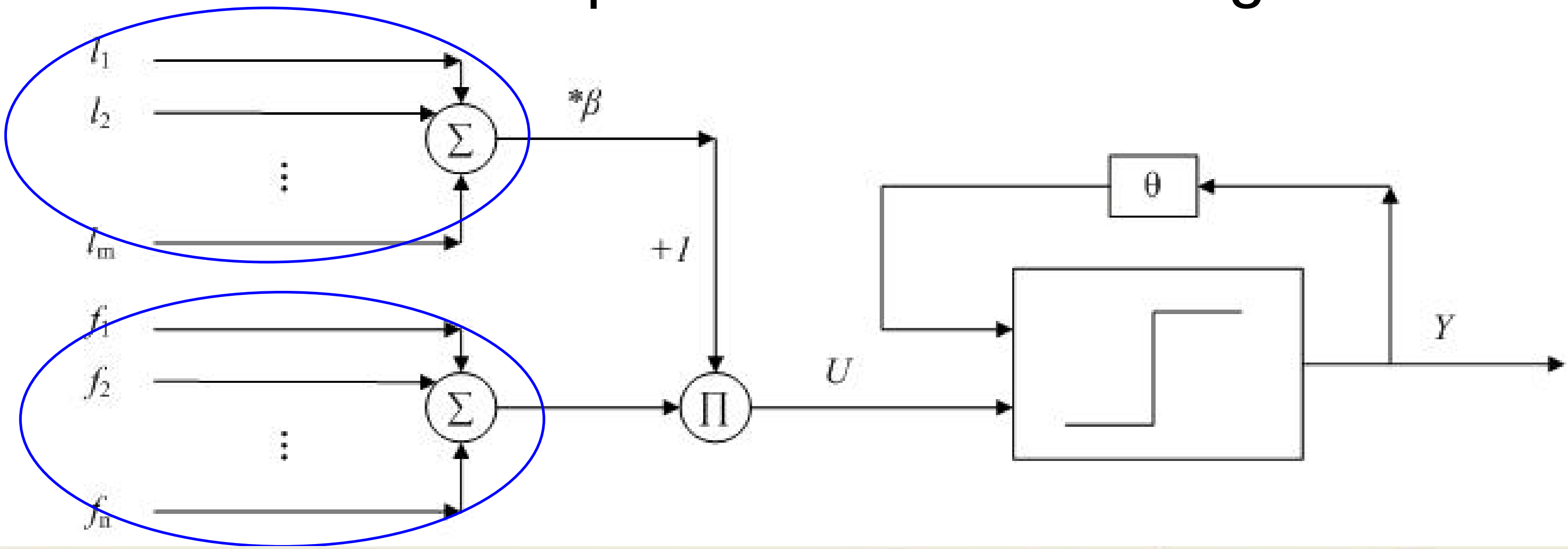
Function MatchAction(α_i,α_j)

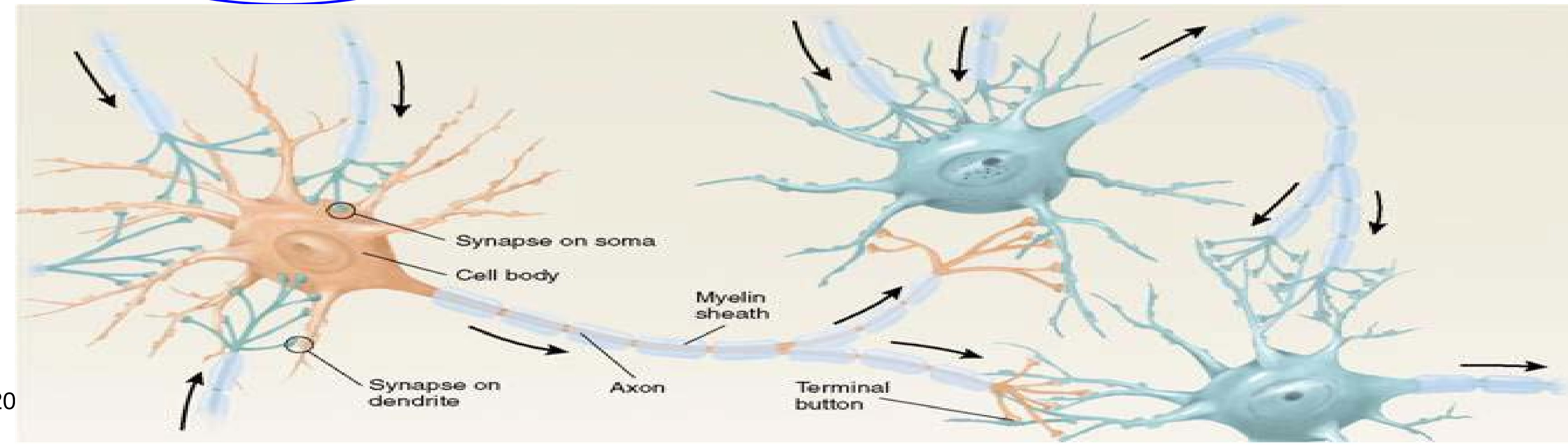
input : action α_i, α_j
output: whether $\alpha_i \models \alpha_j$ hold

if $\alpha_i == null$ *or* $\alpha_j == null$ **then**
    return false
end
f_{pre}:=$Conj(Pre(\alpha_i)) \rightarrow Conj(Pre(\alpha_j))$;
f_{eff}:=$Conj(Eff(\alpha_i)) \rightarrow Conj(Eff(\alpha_j))$;
if $\neg f_{pre}$ *and* $\neg f_{eff}$ *are unsatisfiable according to DDL Algorithm* **then**
    return true ;
else
    return false ;
end

Feature Linking Network

- 1990 Eckhorn Proposed Feature Linking Network

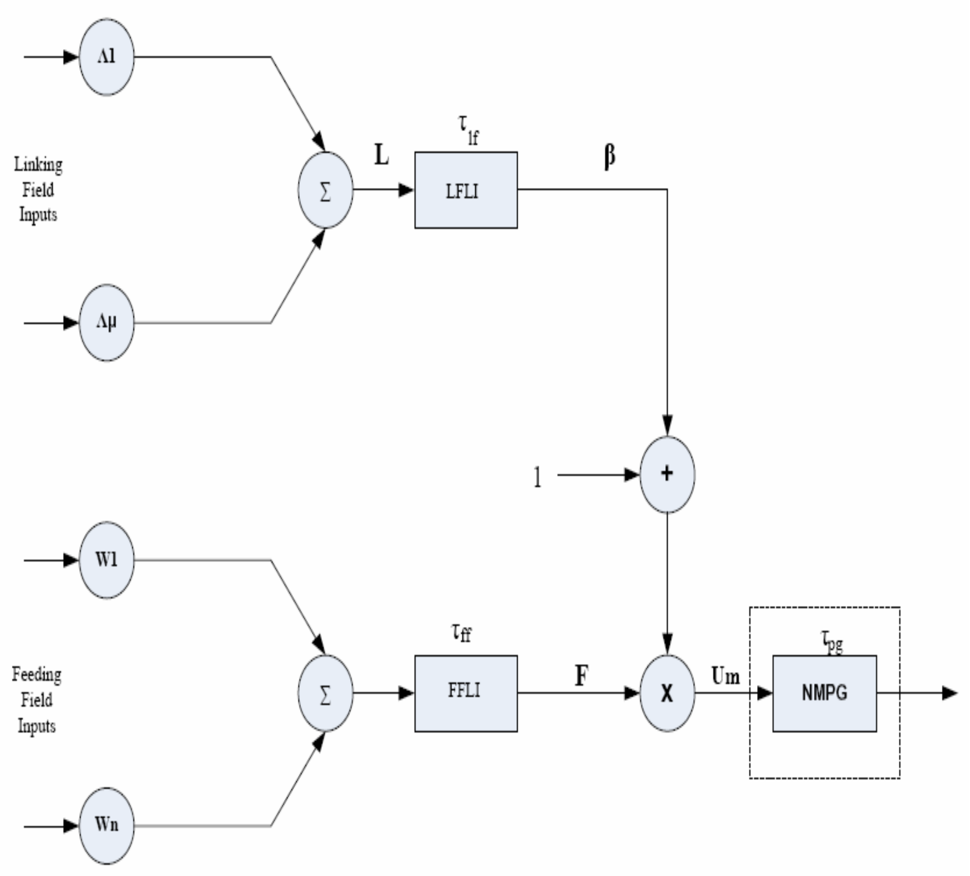




Pulse-Coupled Neural Network

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

INSTITUTE OF COMPUTING TECHNOLOGY

- Dynamic activity

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

Feeding input

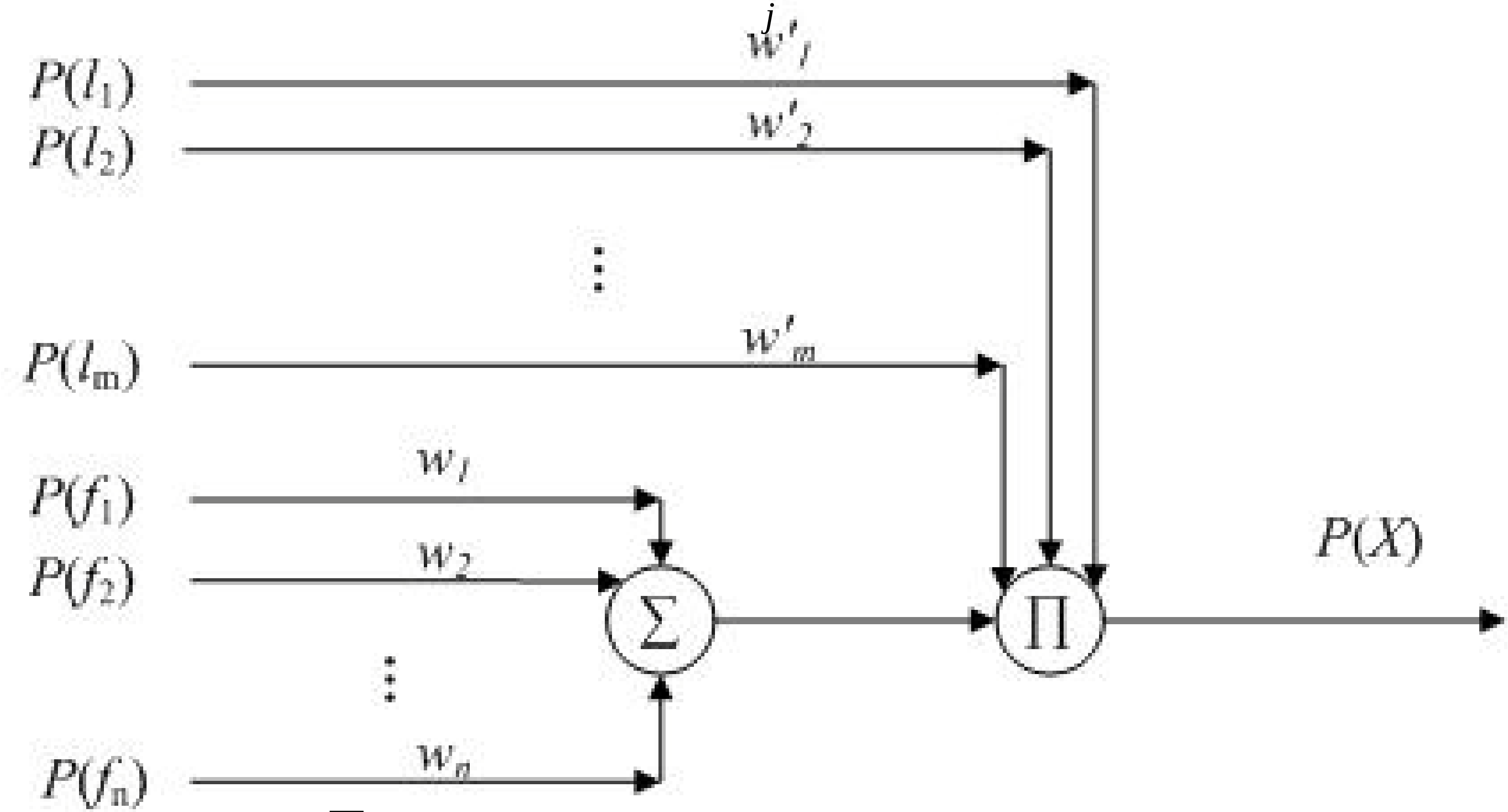
$$P(X \mid l_1, l_2, \ldots) = P(X) \cdot \prod_j w_j^{'} P(l_j)$$

Linking input

Bayesian Linking Field Model

中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY

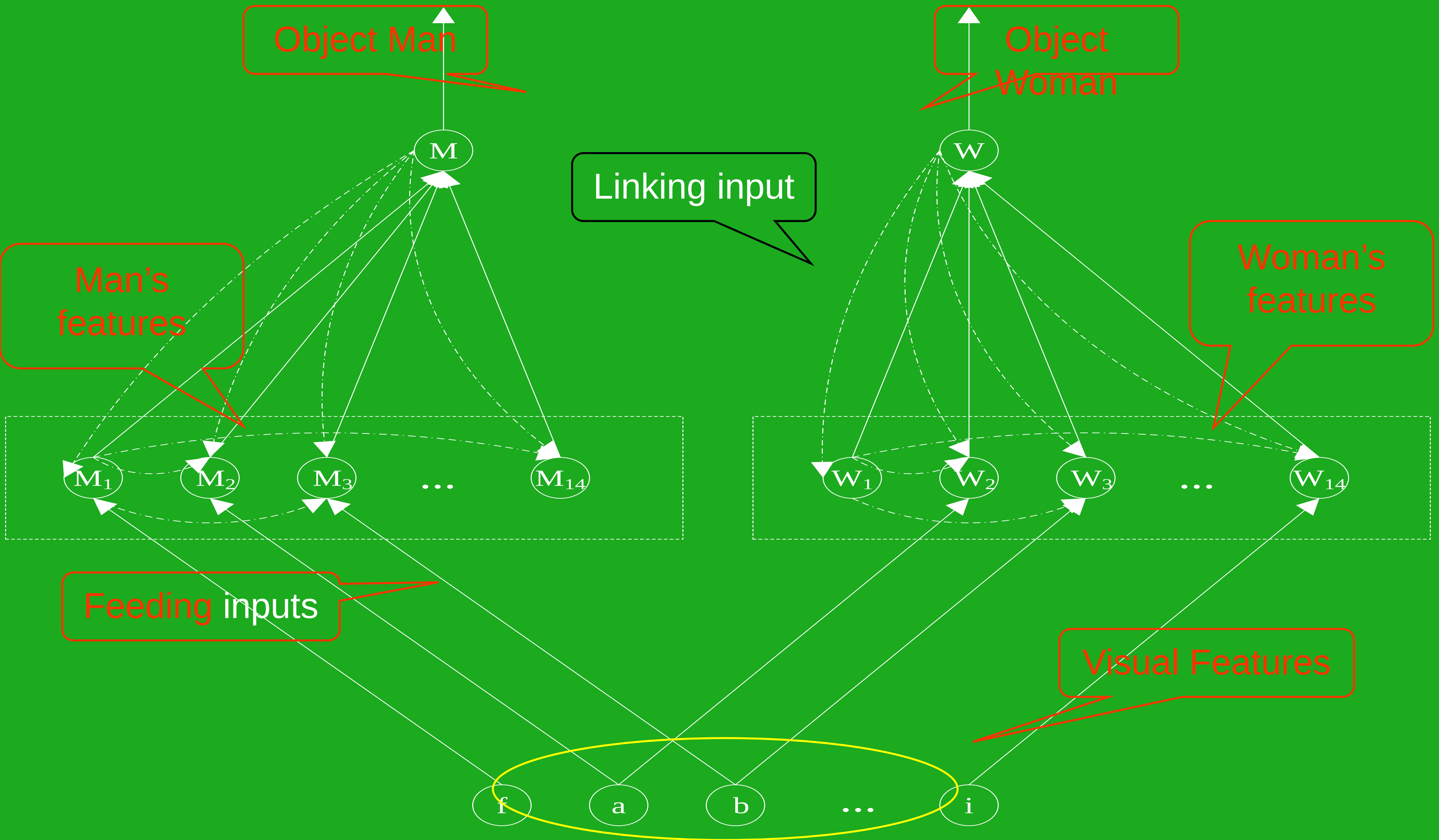
$$P(X \mid l_1, l_2, \ldots) = P(X) \cdot \prod_j w_j^{'} P(l_j)$$



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

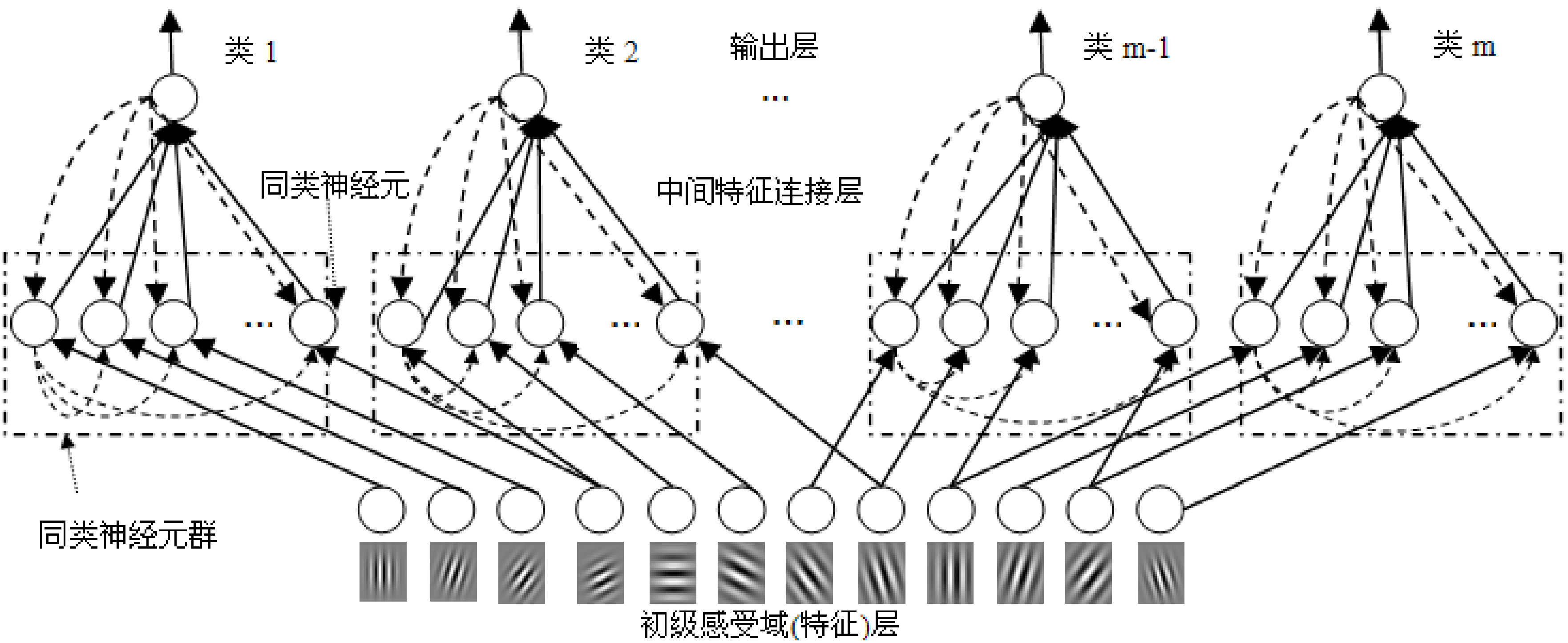
The architecture of Bayesian Linking Field Model

中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY

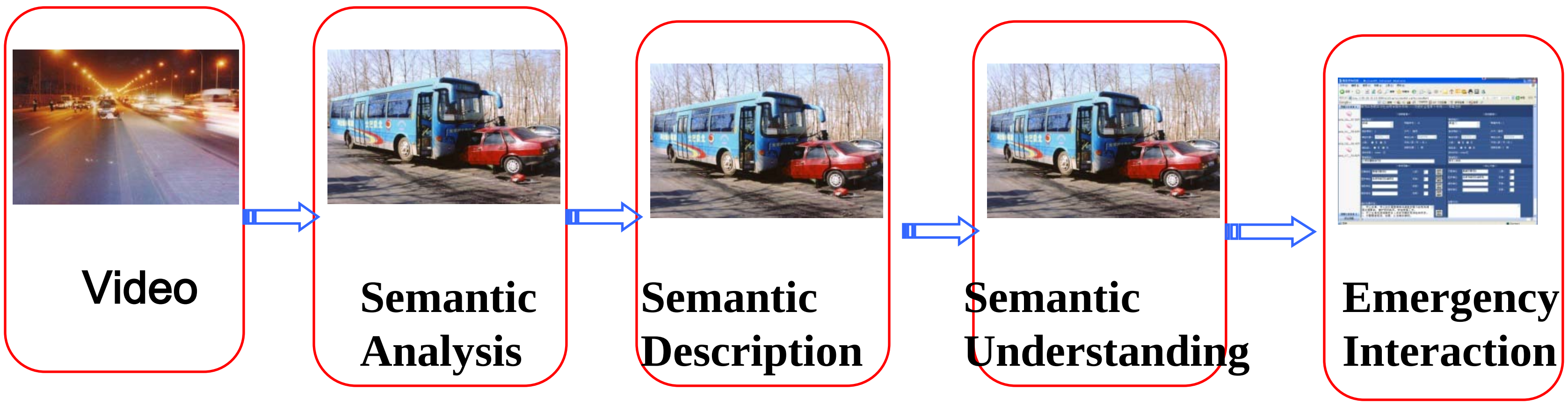


Computational Model for Feature Binding

Bayesian Link field Networks Model



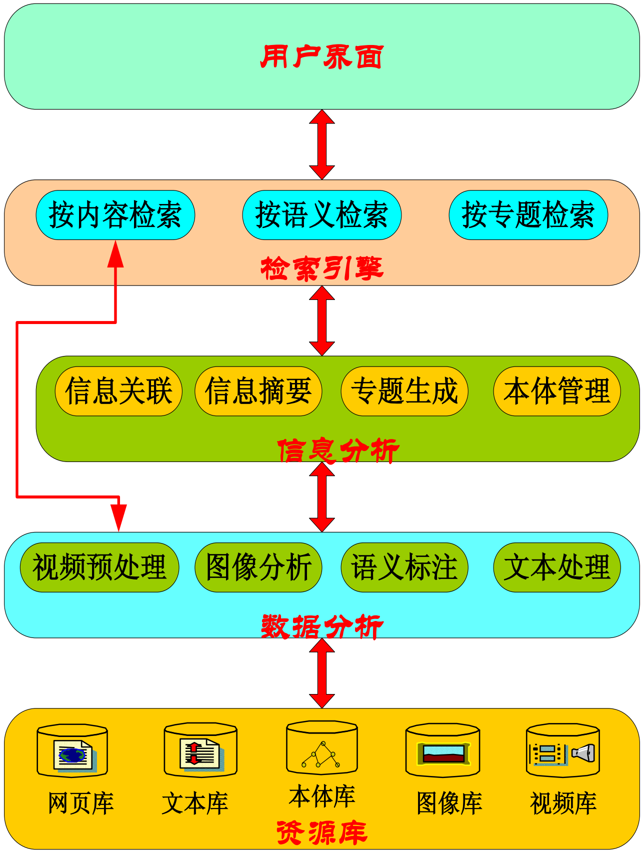
交通监测Traffic Surveillance



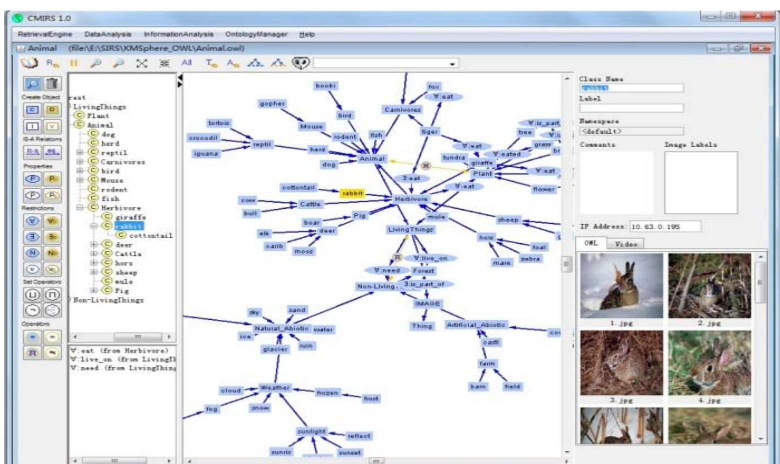
跨媒体智能检索系统CMIRS

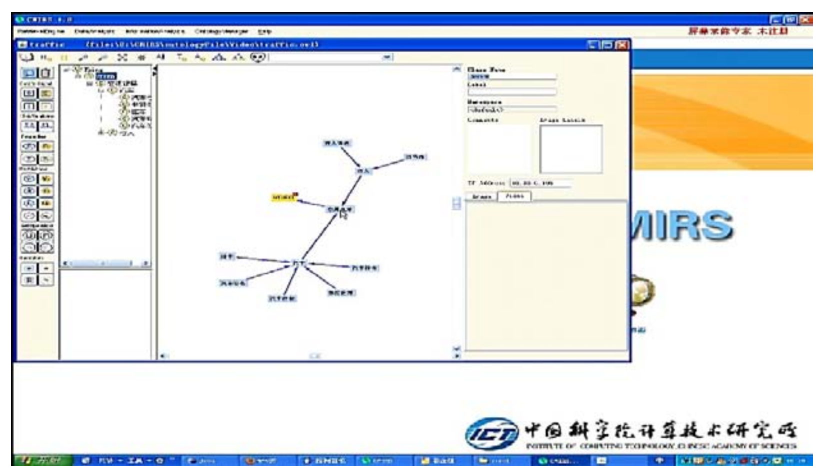
跨媒体智能检索系统CMIRS的逻辑结构分为五层，即资源库、数据分析、信息分析、检索引擎和用户界面。资源曾包括视频库、图像库、本体库、文本库、网页库等。数据分析包括视频预处理、图像分析、自动语义标注，以及文本和网页的处理。信息分析包括本体管理、信息摘要、专题生成、信息关联。本体管理采用本体知识管理系统KMSphere。在此基础上，检索引擎提供按语义检索、按内容检索、按专题检索三种方式，为用户方便使用。用户界面提供用户检索信息、反馈信息的输入，将检索结果以多媒体形式输出给用户。

跨媒体智能检索系统CMIRS旨在支持图像的批处理输入、预处理、特征提取、自动标注和聚类索引等管理功能，而且能在生成式建模和自动标注算法的基础上实现图像的语义检索功能。为便于管理和提高效率，特征提取和自动标注的相关功能都在后台执行，而图像的语义检索则在前台的检索引擎子系统执行。



跨媒体智能检索系统CMIRS





展 望

中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY

☛ 生命科学、纳米技术、认知科学、信息技术的结合将催生新型智能机器

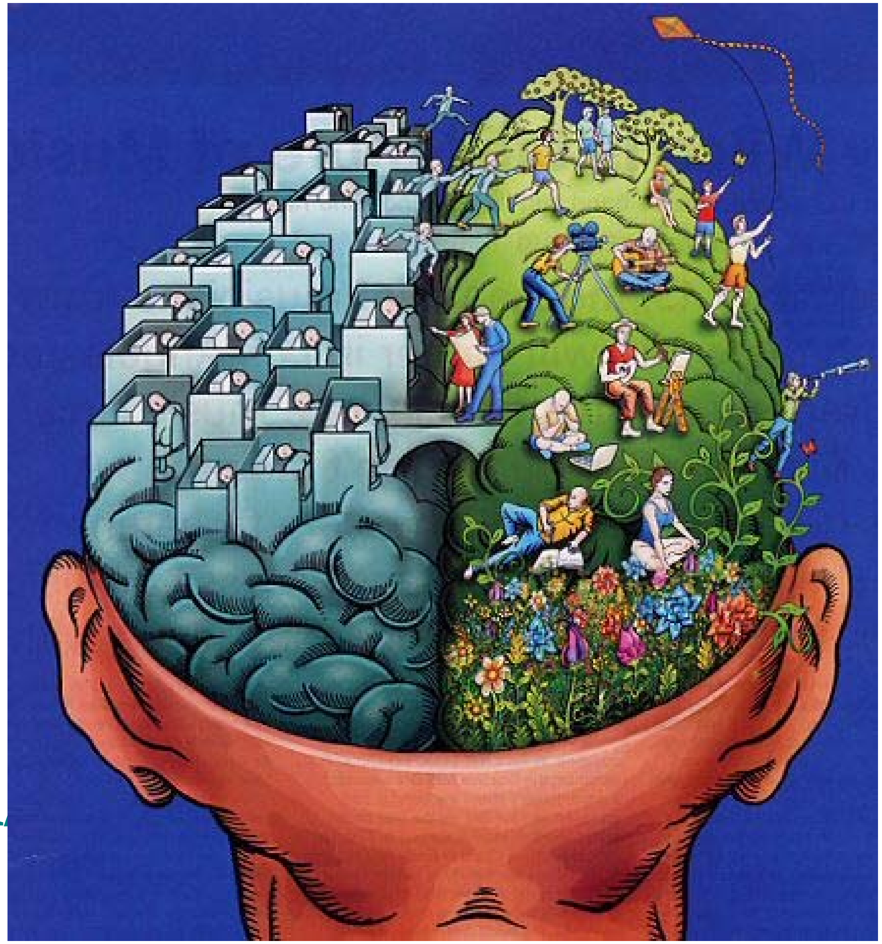
☛ 信息技术面临重大变化， 智能科学与技术将引领发展。

☛ 心智模型CAM 将为新型智能机器提供理论基础。

谢　　谢

中科院计算所
INSTITUTE OF COMPUTING TECHNOLOGY

Question!



Intelligence Science

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