An Engineering Perspective on Reverse Engineering the Brain

James Albus

Senior Fellow
Krasnow Institute for Advanced Studies
George Mason University

Founder & Chief (Retired) – Intelligent Systems Division
National Institute of Standards and Technology

james.albus@gmail.com

Outline

An Engineering Viewpoint

The Neuroscience

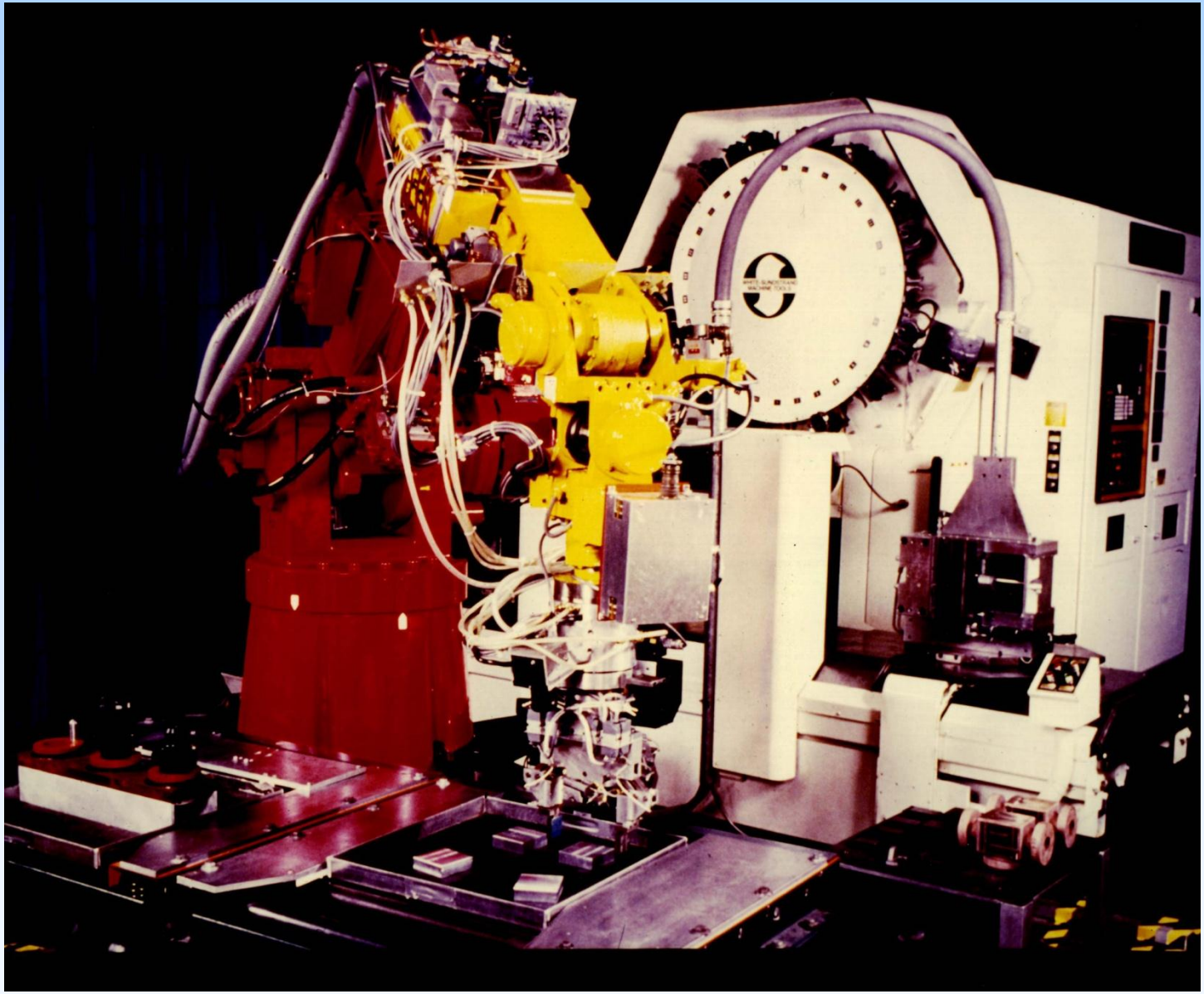
Reverse Engineering the Brain

Intelligent Systems Engineering

Intelligent Control Projects ~ $100M total over 43 years

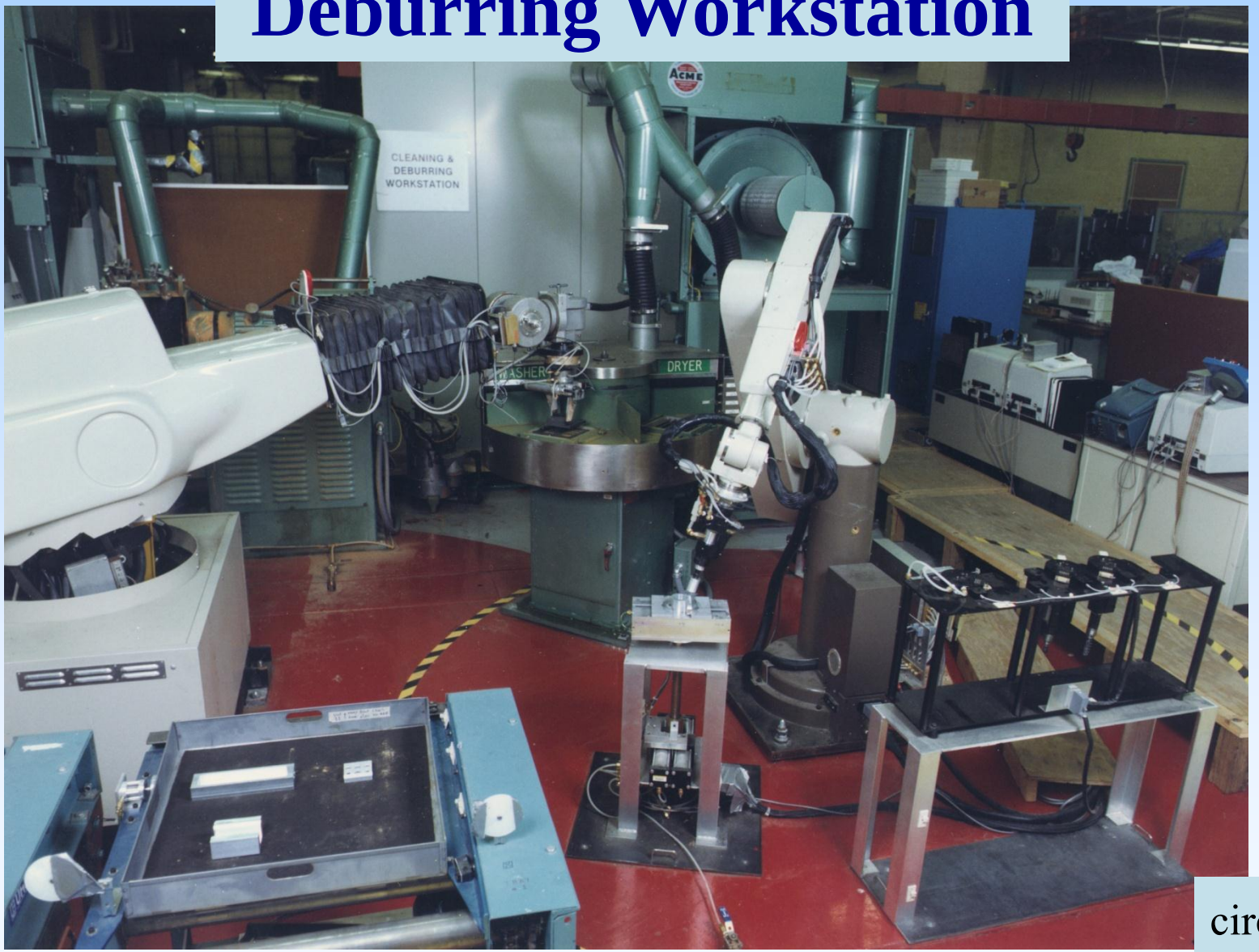
65-75 NASA-NBS -- Cerebellum model for learning control (CMAC neural net)
73-85 Navy/NBS – Robot control, Automated Manufacturing Research Facility
86-87 DARPA -- Multiple Unmanned Undersea Vehicles (MAUV)
88-89 DARPA -- Submarine Operational Automation System (SOAS)
90-92 GD Electric Boat -- Next generation nuclear submarine control
86-88 NASA -- Space Station Flight Telerobotic Servicer (NASREM)
87-89 Bureau of Mines -- Coal mine automation
87-91 U.S. Postal Service -- Stamp distribution center, General mail facility
86-08 Army -- TEAM, TMAP, MDARS, Picatinny Arsenal UGV, Demo I and III ARL Collaborative Technology Alliance, JAUGS, VTA, FCS-ANS
96-97 Navy – Double Hull Robot, Multiple UAV SWARM
94-95 DARPA / General Motors – Enhanced CNC & CMM Control
99-01 Boeing – Cell Control, Riveting, Hi Speed machine tool
92-01 Commercial CNC - plasma & water jet cutting
96-98 DARPA – MARS, PerceptOR
02-04 Boeing/SAIC – FCS Autonomous Navigation System, Integrated Combat Demo
02-07 AirForce – RoboCrane Paint Stripping Robot for Large Aircraft
08-09 DOT – Intelligent vehicles, Foveal-Peripheral Vision for Driving
06-07 DARPA – Learning Applied to Ground Robotics (LAGR)
08-10 DARPA – EATR Foraging Robot

Intelligent Machining Workstation



circa 1981

Intelligent Cleaning and Deburring Workstation



circa 1982

Intelligent Coal Mining Machine

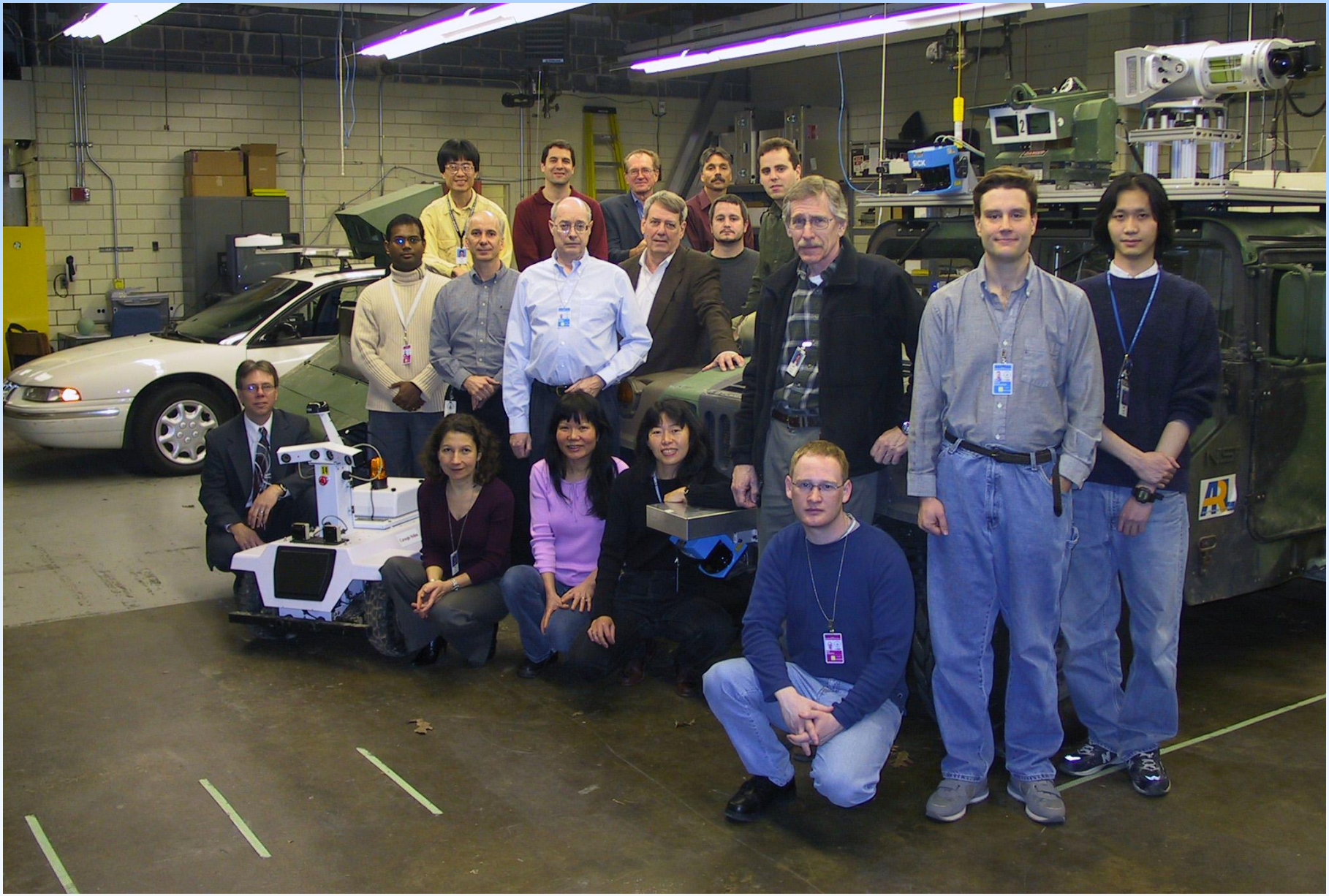
circa 1988

Multiple Autonomous Undersea Vehicles

Intelligent Vehicle Control

circa 1993

NIST Autonomous Mobility Team



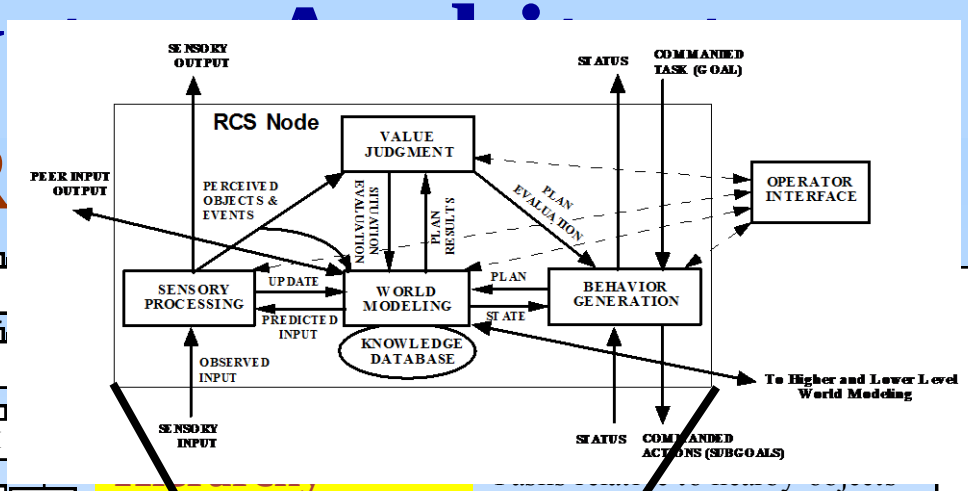
4D/RCS Reference Model Architecture for Unmanned Vehicle Systems

Adopted by GDRS for FCS Autonomous Navigation System
Adopted by TARDEC for Vetronics Technology Integration

- **Hierarchical structure of goals and commands**
- **Representation of the world at many levels**
- **Planning, replanning, and reacting at many levels**
- **Integration of many sensors stereo CCD & FLIR, LADAR, radar, inertial, acoustic, GPS, internal**

Intelligent Sy

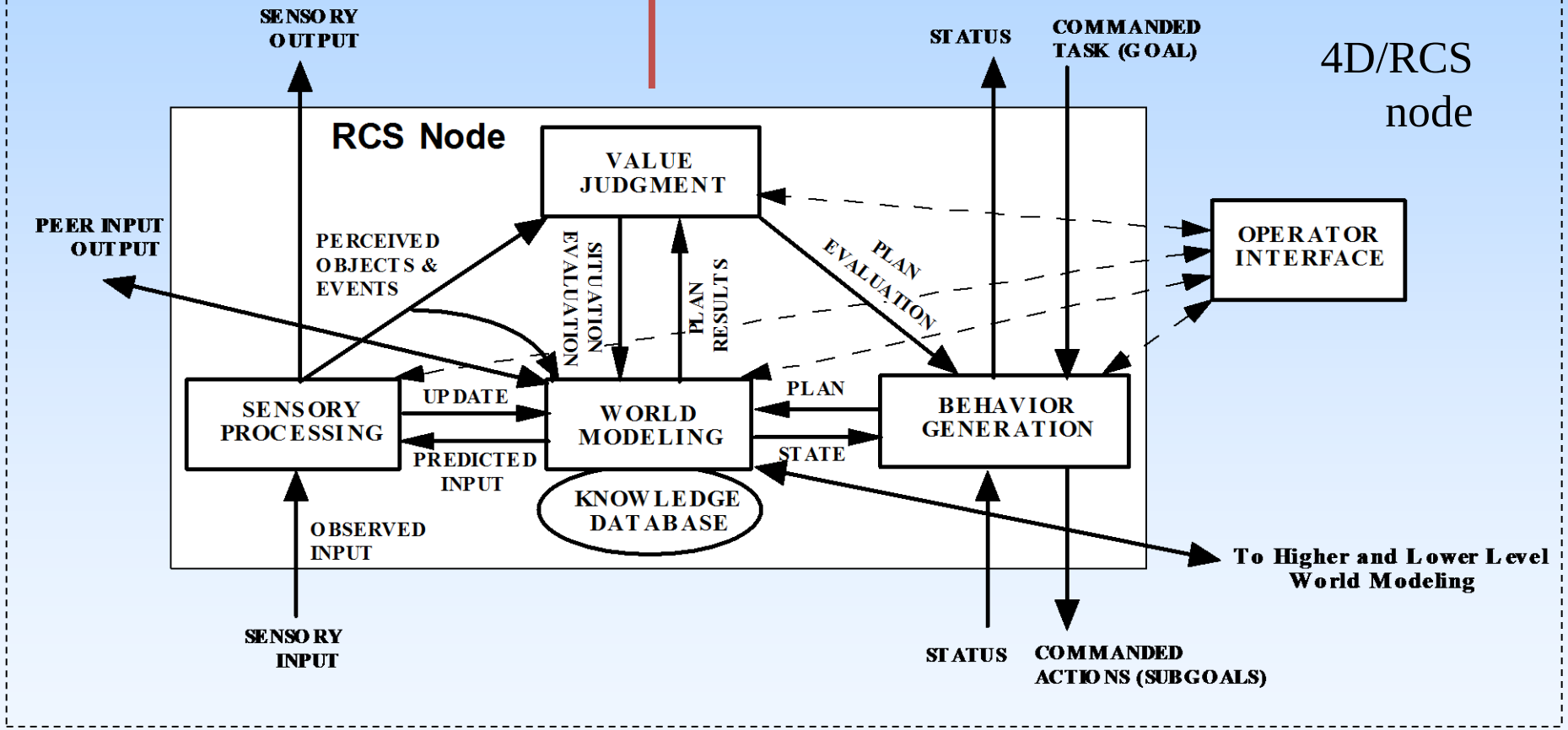
4D/RCS R



A 4D/RCS Computational Node

Mapping to the Brain

Posterior Cortex | Frontal Cortex



What is the Goal?

The Engineering Goal

To build machines that DO what the brain does

A Scientific Goal

To understand HOW the brain does what it does.

A second Scientific Goal

To understand how the brain LEARNS to do what it does

Overall Structure of Brain

Front to back:

Behavior generation in front

Sensory processing in back

Side to side:

Representation of right egosphere on left side

Representation of left egosphere on right side

Top to bottom:

Conscious self at top

Sensors and muscles at bottom

At the center:

Emotions, Appetites, & Internal state

What is the brain for?

The brain is first and foremost a control system

Early evolution => control of locomotion

Swimming motion & gait generation – coordination of actuators

Path planning – how to get from A to B

Decision making – where to go, when, why, how

Tactical behaviors – hunting for food, evading predators, . . .

Strategic behaviors – migrating, establishing territory, mating, . . .

Fine manipulation, language, and reasoning are recent developments

What are the Inputs?

Gravity sensors establish the horizontal plane for an internal egosphere representation

Body kinematics measured by proprioception

Body dynamics measured by vestibular sensors

Tactile input <= Arrays of sensors in the skin

Visual input <= Arrays of sensors in the retina

Audio input <= Arrays of sensors in the ears

Smell and taste input <= Sensors in nose and mouth

What are the Outputs?

Behavior – consistent with goals that are generated in the frontal cortex by processes that use:

- a rich internal model of the external world
- an internal model of body kinematics and dynamics
- an internal representation of needs and desires

Behavior – consisting of:

- control signals to muscles
- forces and velocities in the limbs and torso
- goal-driven tasks and subtasks on objects in the world

Behavior – that has many levels of resolution in:

- planning and coordination
- feedback error correction
- feed-forward control

Hierarchical Architecture

Brain is organized hierarchically

Unitary SELF at top
Millions of sensors and actuators at bottom

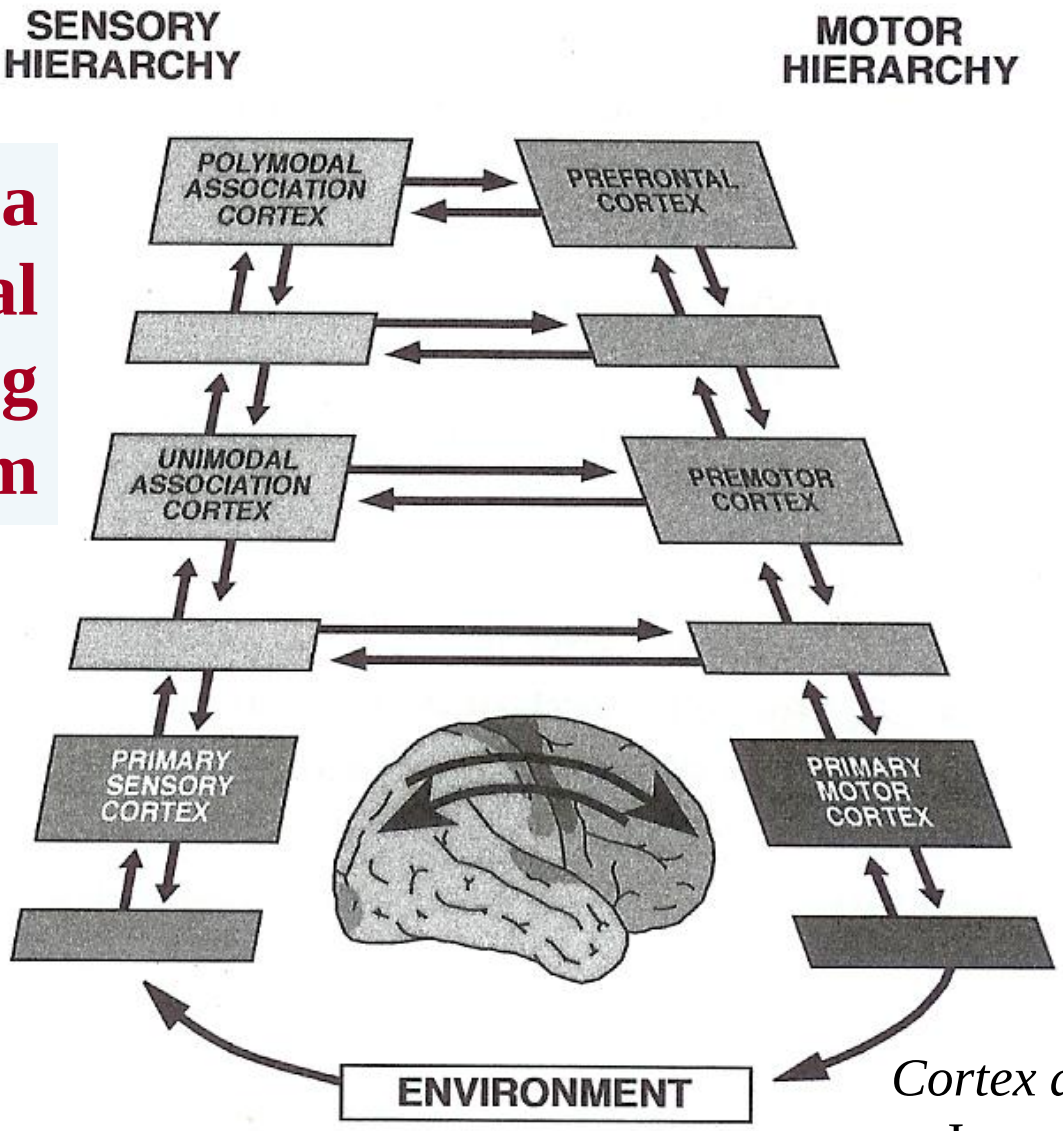
Complex strategies at top
Simple actions at bottom

Frontal hierarchy: **decision making, goal selection, priority setting, planning and execution of behavior**

Posterior hierarchy: **attention, segmentation, grouping, computing attributes, classification, establishing relationships**

Cortical Architecture

The brain is a hierarchical signal processing & control system

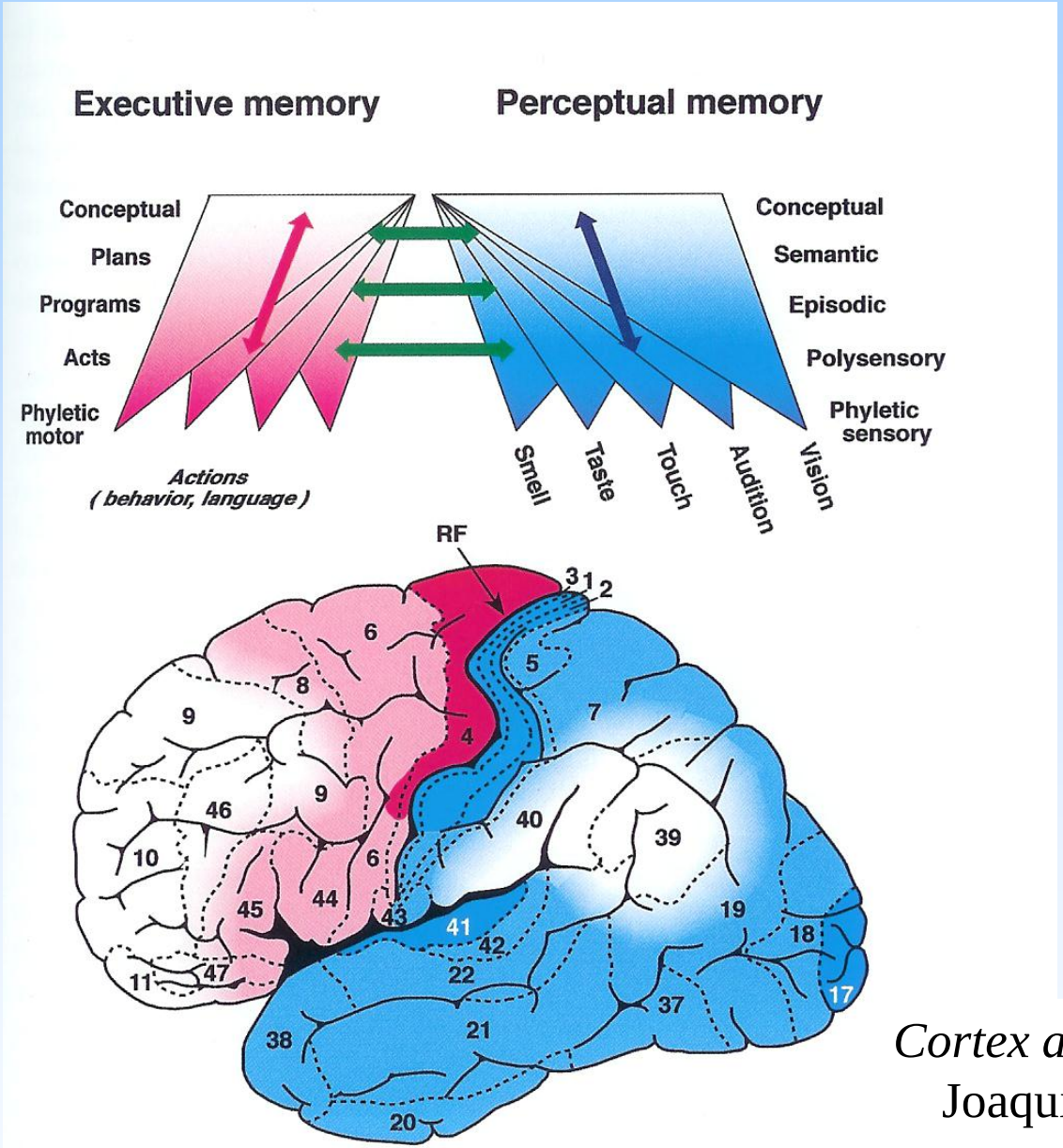


Cortex and Mind
Joaquin Fuster

Hierarchical Architecture

Brain hierarchy is not a pyramid

More neurons at the top



Cortex and Mind
Joaquin Fuster

Computational Mechanisms

Synapse is an electronic gate

-- complex biochemistry, site of long-term memory

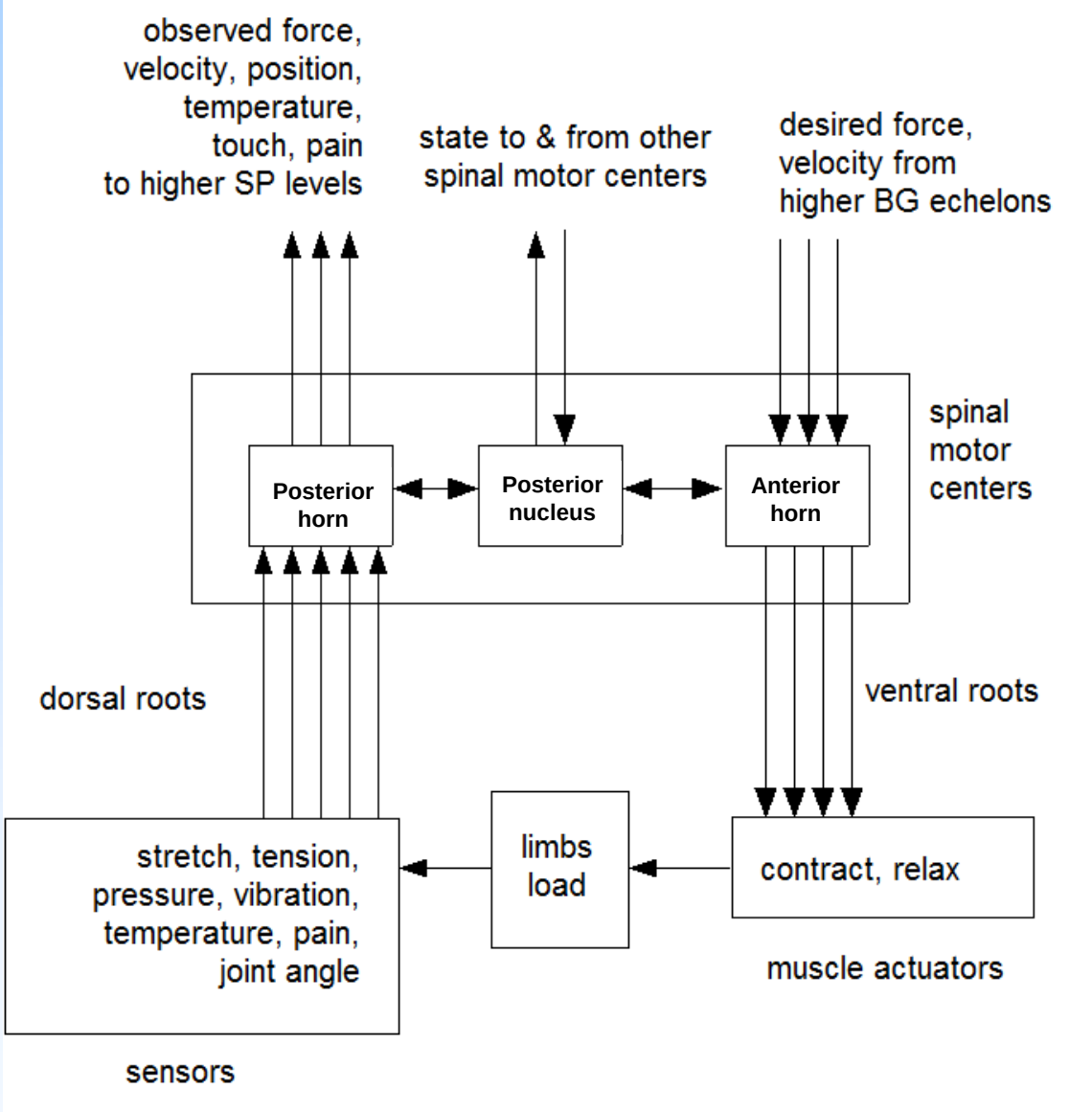
Neuron is a computational element

-- non-linear processes on many inputs, & decide

Neural Cluster is a functional unit

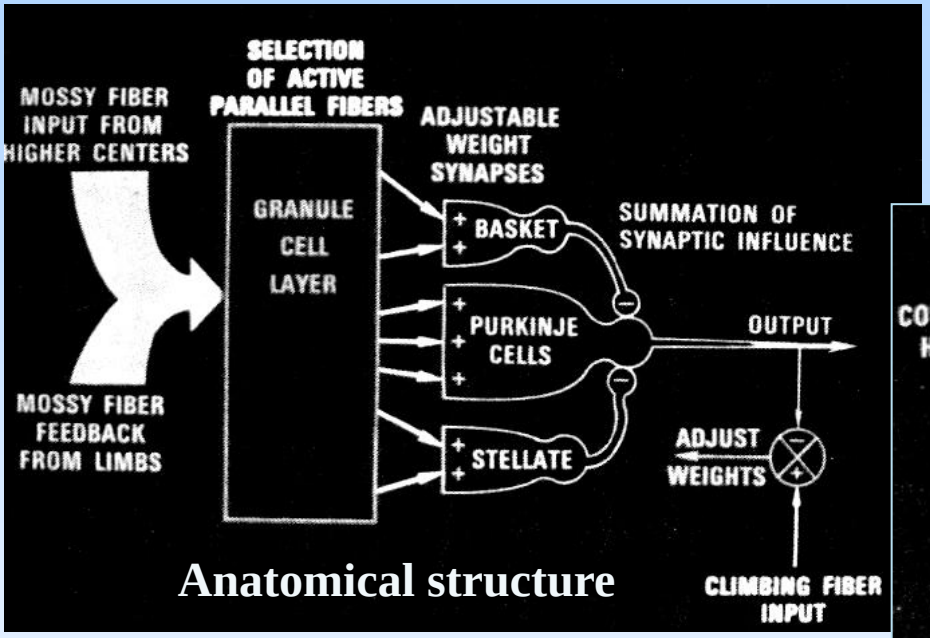
-- arithmetic or logical operations, correlation, convolution
-- coordinate transformation
-- finite-state automata
-- rules, grammar, direct and indirect addressing

Neural Clusters in Spinal Cord

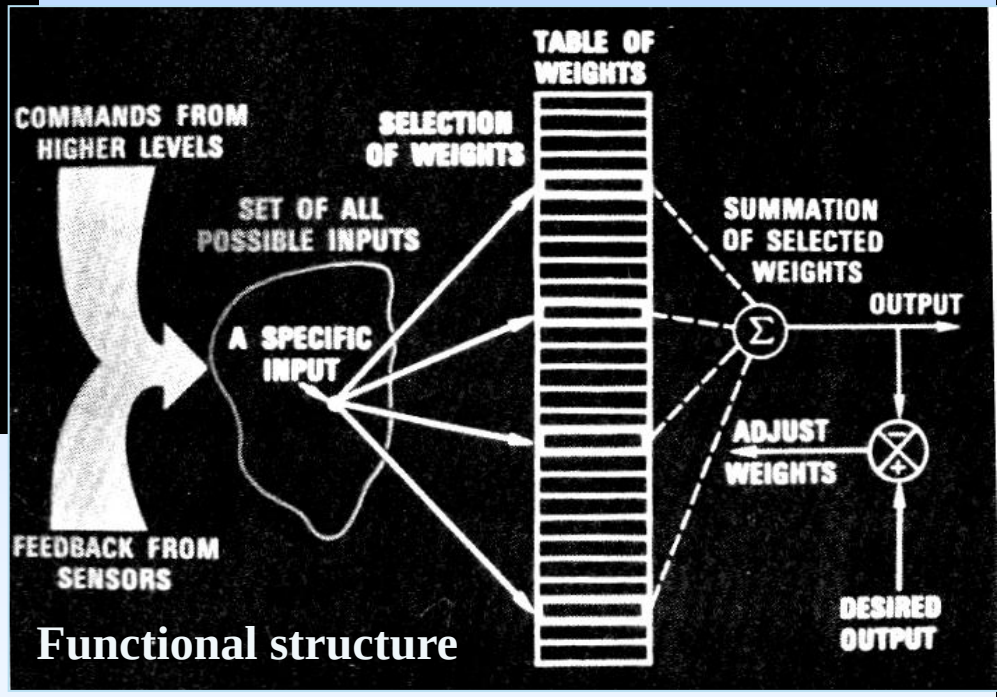


Neural Clusters in Midbrain
(e.g. Cerebellum)

Random access table-look-up computation with generalization



Anatomical structure

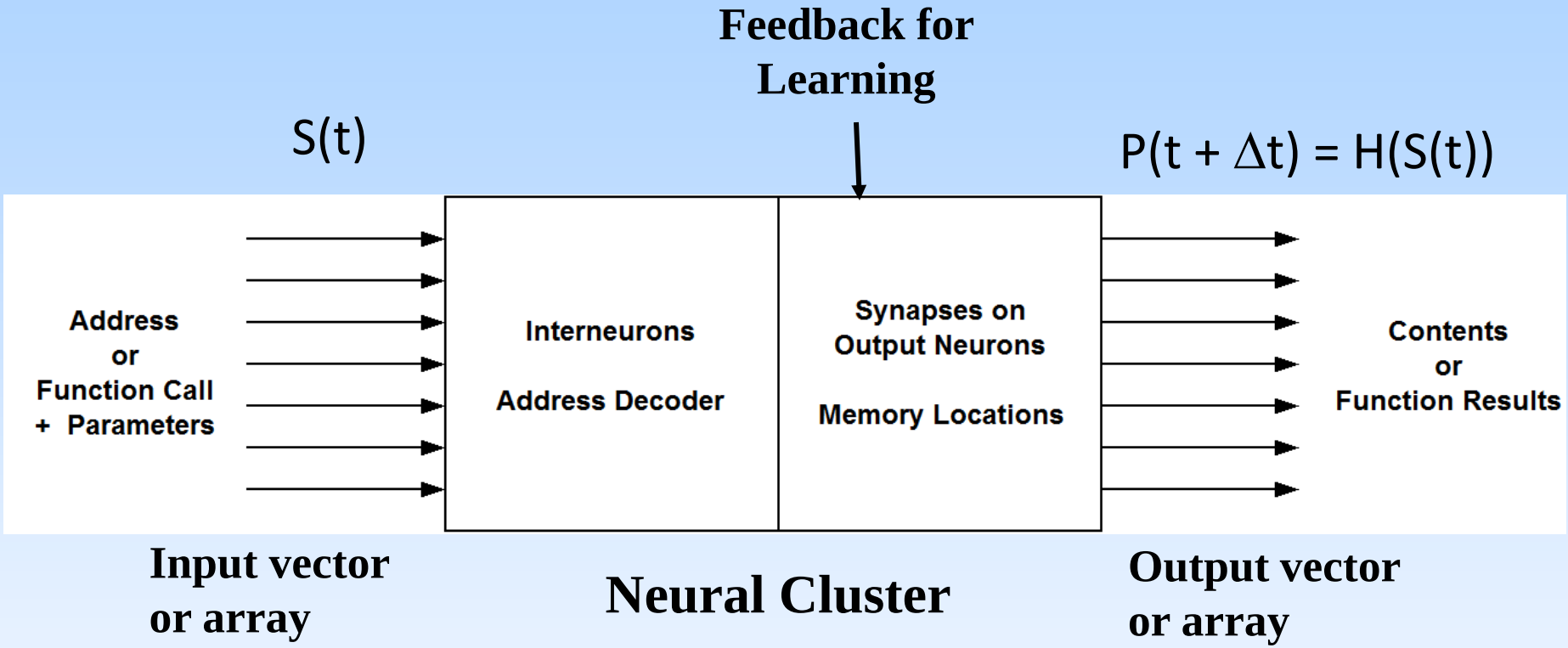


Functional structure

Input	Output
Command & feedback	Action
Address	Contents
Address	Pointer
If (Situation)	Then (Consequent)

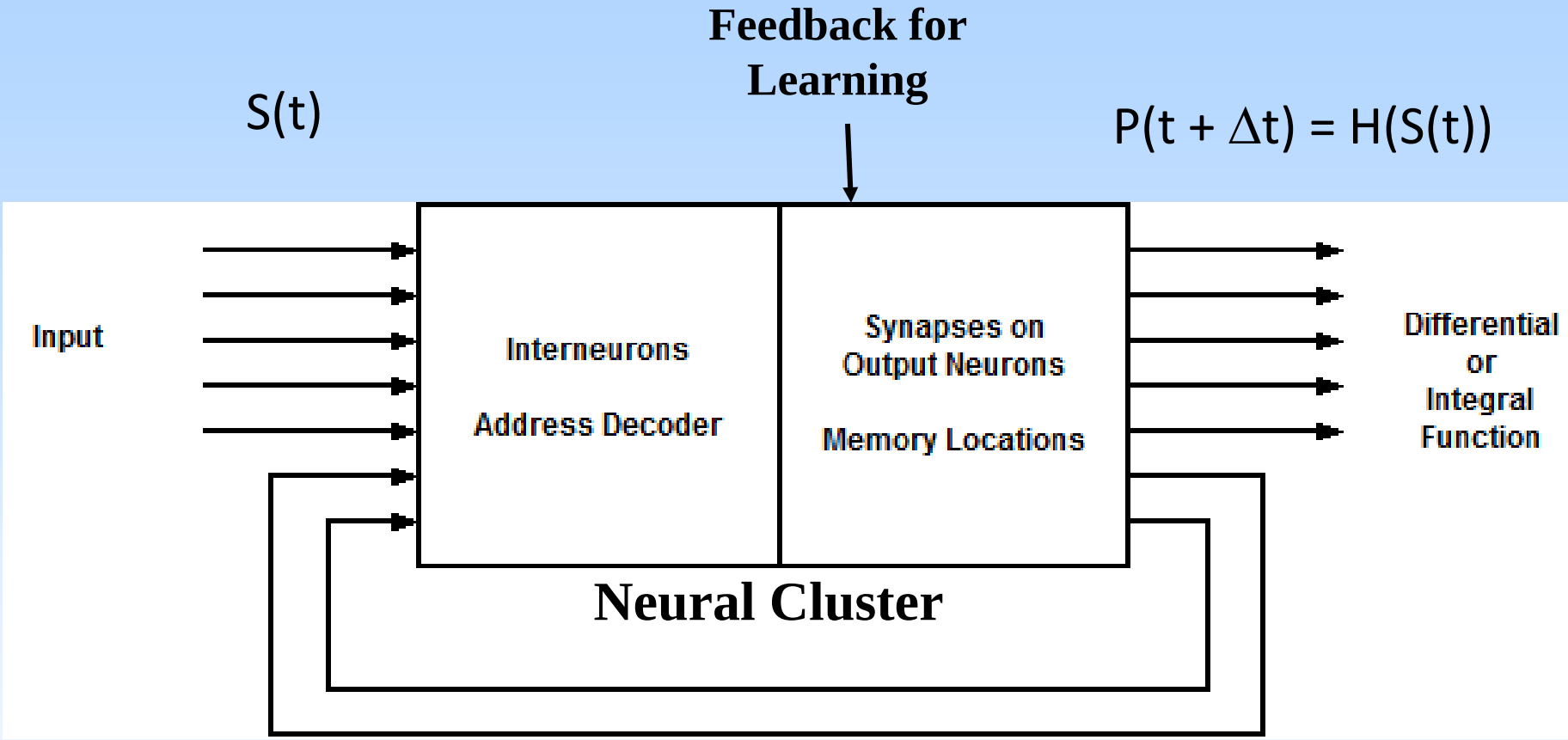
Marr 1969, Albus 1971

General Functional Model



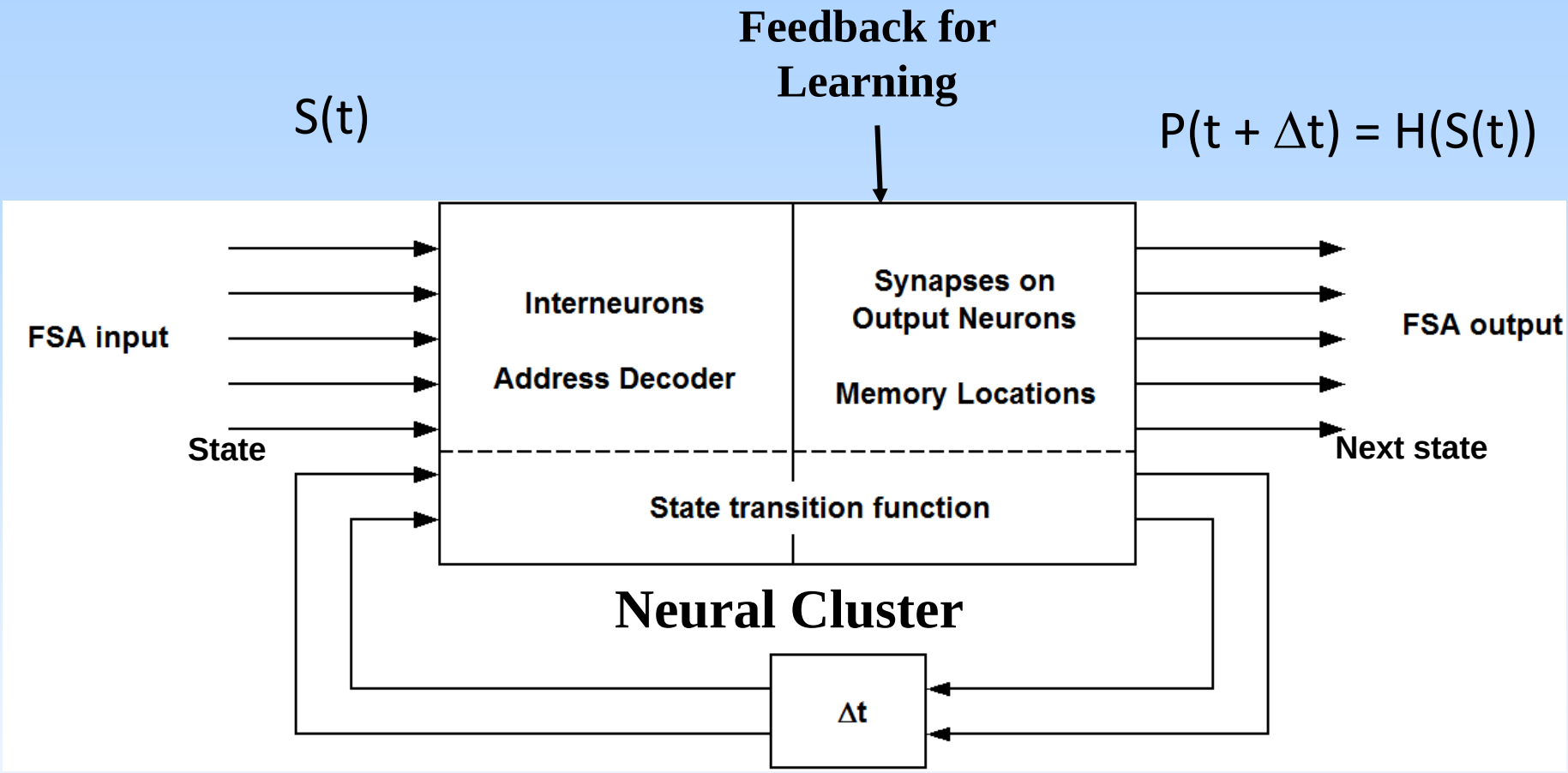
memory storage & recall, arithmetic or logical functions,
IF/THEN rules, goal-seeking reactive control,
forward & inverse kinematics, direct & indirect addressing

Functional Model + Feedback



differential and integral functions, dynamic models, phase-lock loops, time and frequency analysis, recursive estimation, Kalman filtering

A Neural Finite State Automaton



Markov processes, scripts, plans, behaviors, grammars, Bayesian networks, semantic nets, narratives

Cortical Structure

Cortex is a 2D sheet – 2000 cm^2 area x 3 mm thick

Cortical sheet is partitioned into functional regions

Regions are arranged in hierarchical layers

Each region is segmented into arrays of columns

Each column has capabilities of a fsa + memory

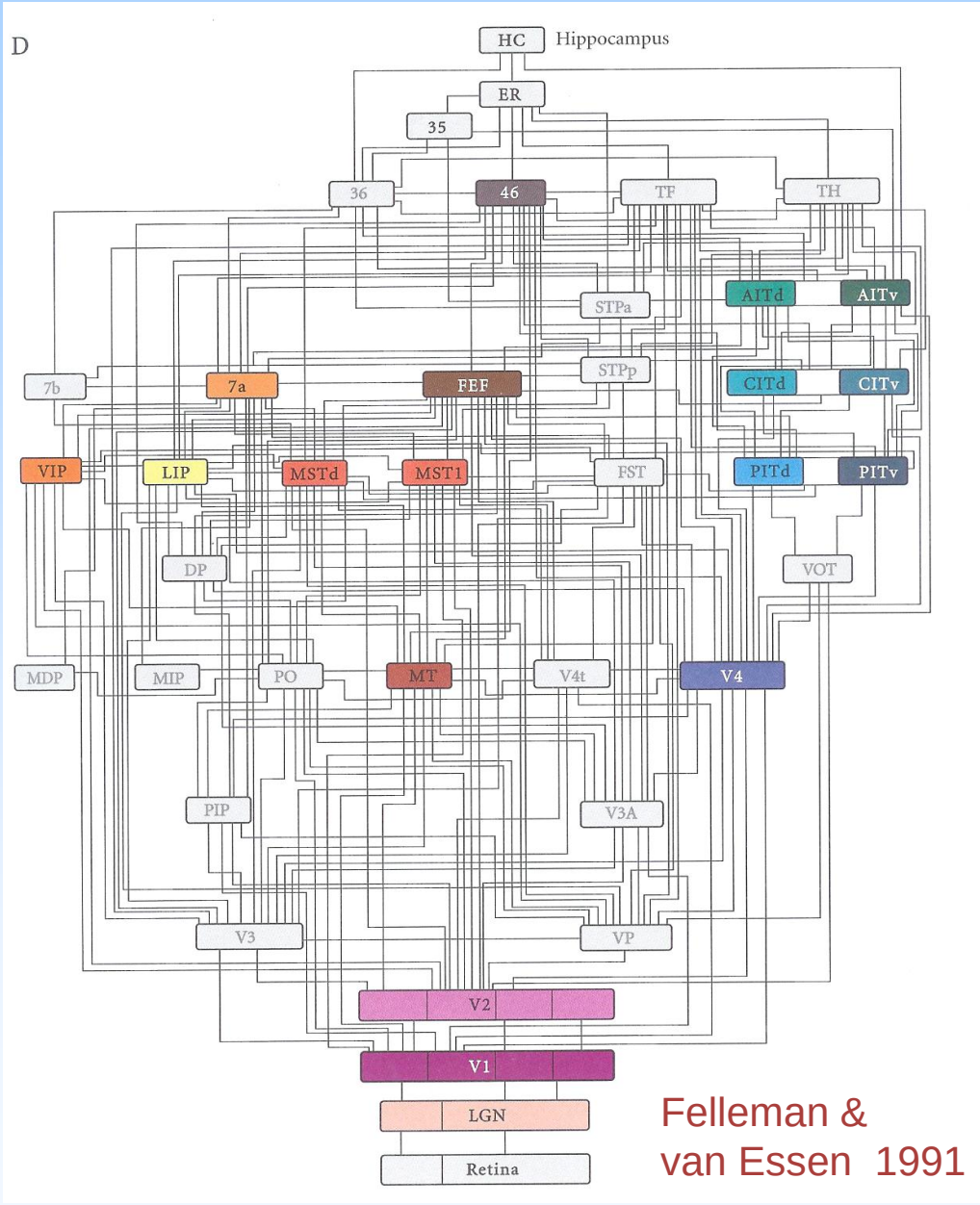
Circuit diagram of visual system in brain

12 layers

32 areas

Each area is an array of Cortical Columns

Each area represents the Visual Field of Regard



Felleman & van Essen 1991

Cortical Column Structure

Microcolumns

100 – 250 neurons

30 – 50 μ diameter, 3000 μ long

Posterior: detect patterns, compute attributes

Frontal: evaluate alternatives, recommend actions

Hypercolumns (a.k.a. columns)

100+ microcolumns in a bundle

500 μ in diameter, 3000 μ long

Posterior: segmentation, grouping, classification, relationships

Frontal: set goals, make plans, control action

There are about 10^6 hypercolumns in human cortex

Communication in the Brain

Axon is an active fiber connecting one neuron to others

(transmits a scalar variable on a publish-subscribe network with bandwidth ~ 500 Hz)

Two kinds of axons:

- **Drivers – Preserve topology and local sign**

  Data vectors or arrays of attributes and state-variables

  i.e., images, objects, events, attributes and state -- e.g., color, shape, size, position, orientation, motion

- **Modulators – Don't preserve topology or local sign**

  Context & broadcast variables, addresses, and pointers

  e.g., select & modify algorithms, set parameters, define relationships

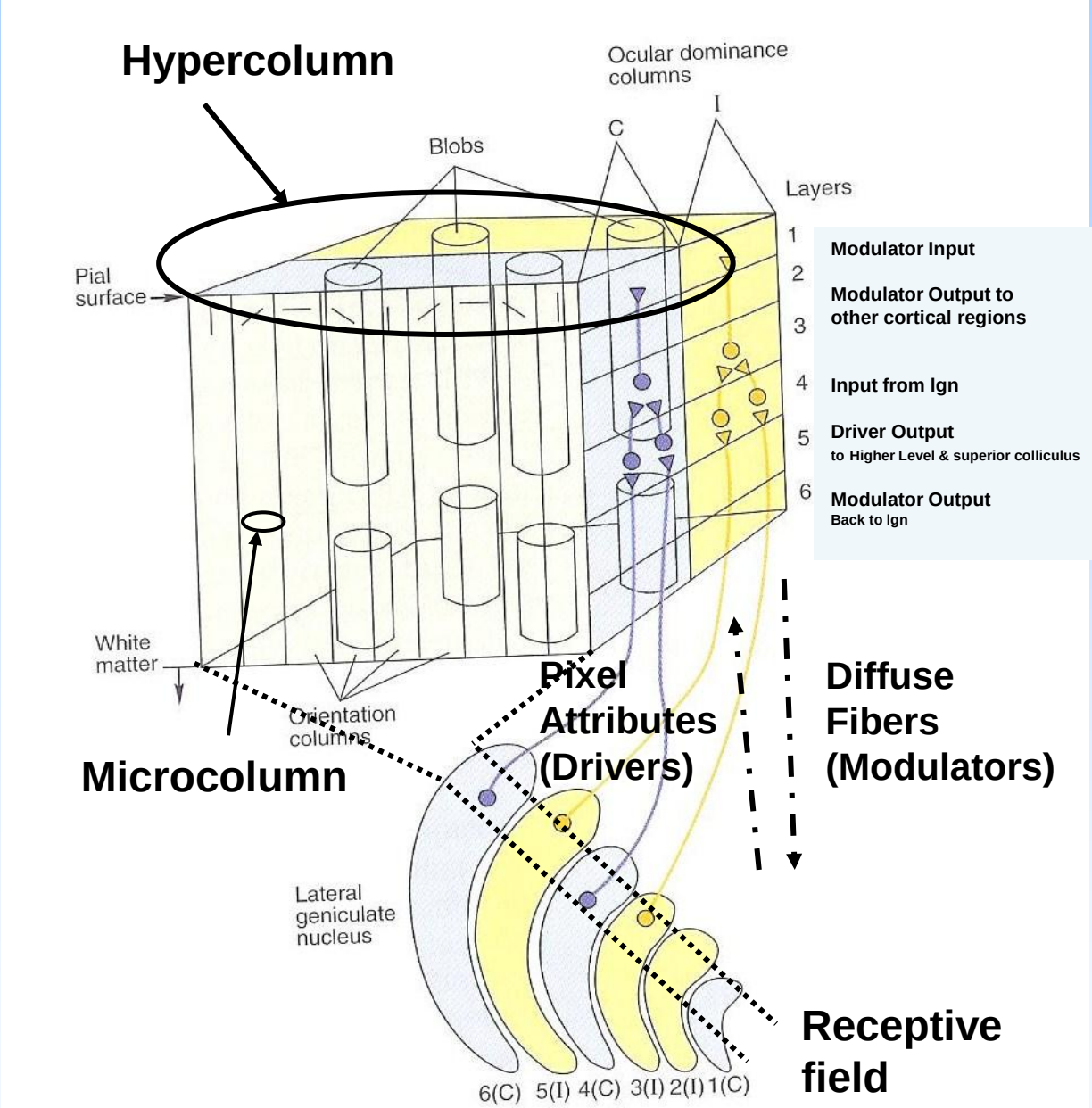
Exploring the Thalamus
Sherman & Guillery 2006

Early Cortical Vision Processes

Cortical Columns in V1

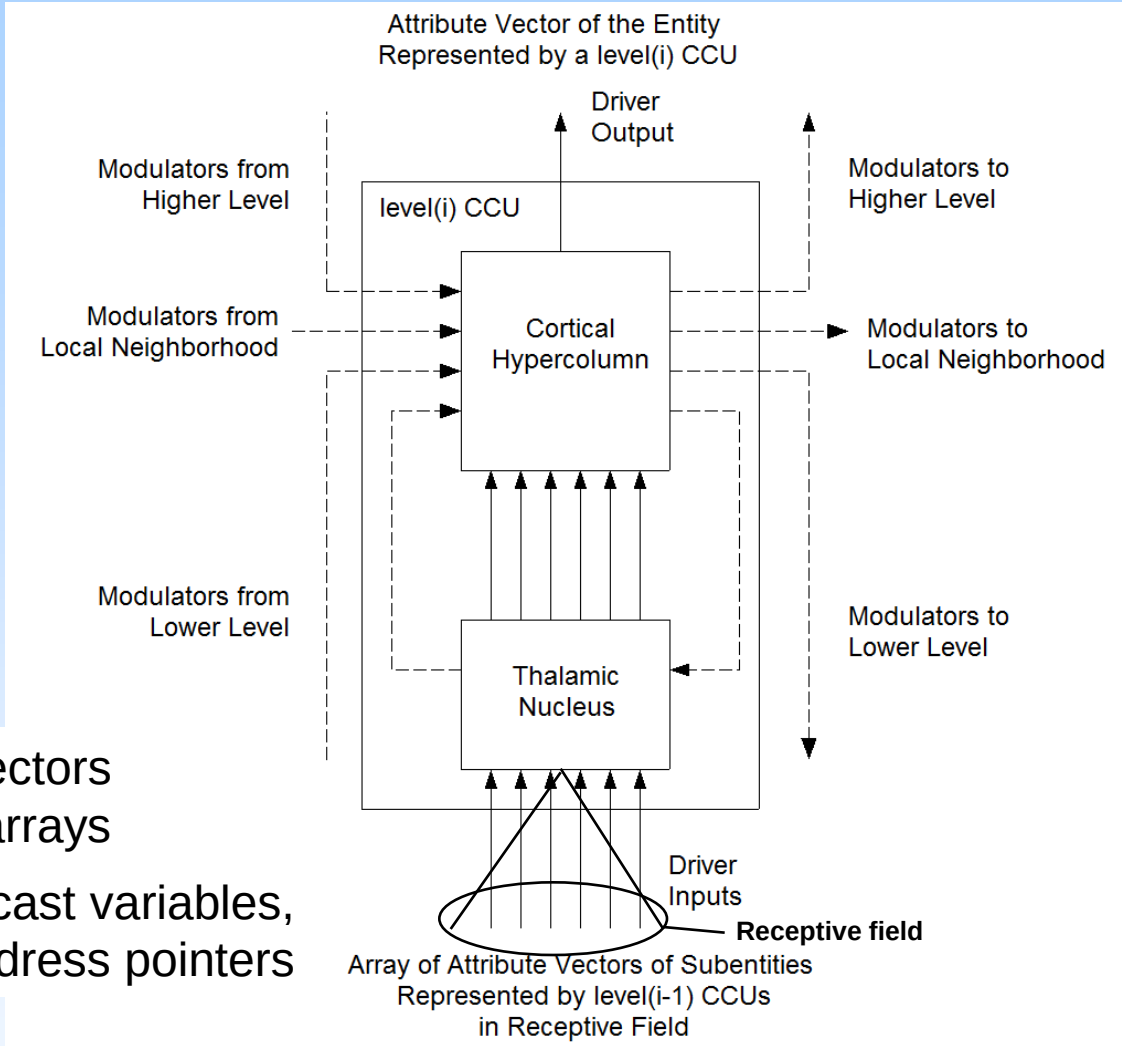
+

Lateral Geniculate in Thalamus



Cortical Hypercolumn + Thalamic loop

Cortical Computational Unit (CCU)

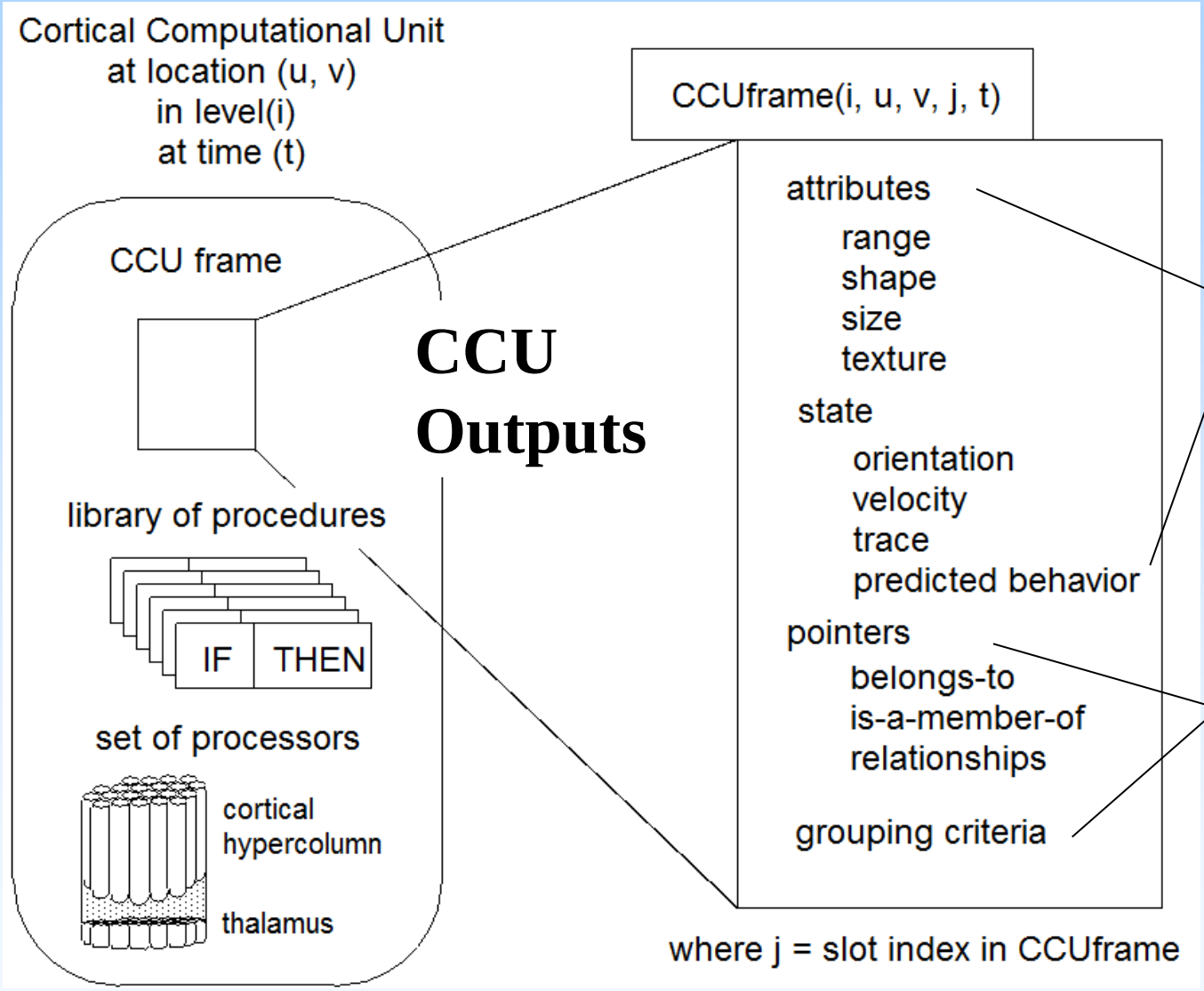


drivers = attribute vectors & arrays

modulators = broadcast variables, & address pointers

windowing, segmentation, grouping, computing group attributes & state, filtering, classification, setting and breaking relationships

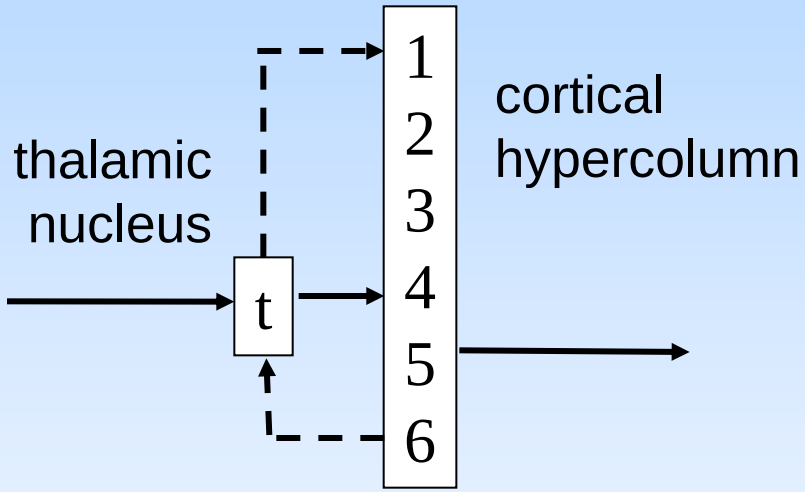
CCU Data Structure Hypothesis



Cortico-Thalamic Loop

drivers = attribute vector array

modulators = address pointers

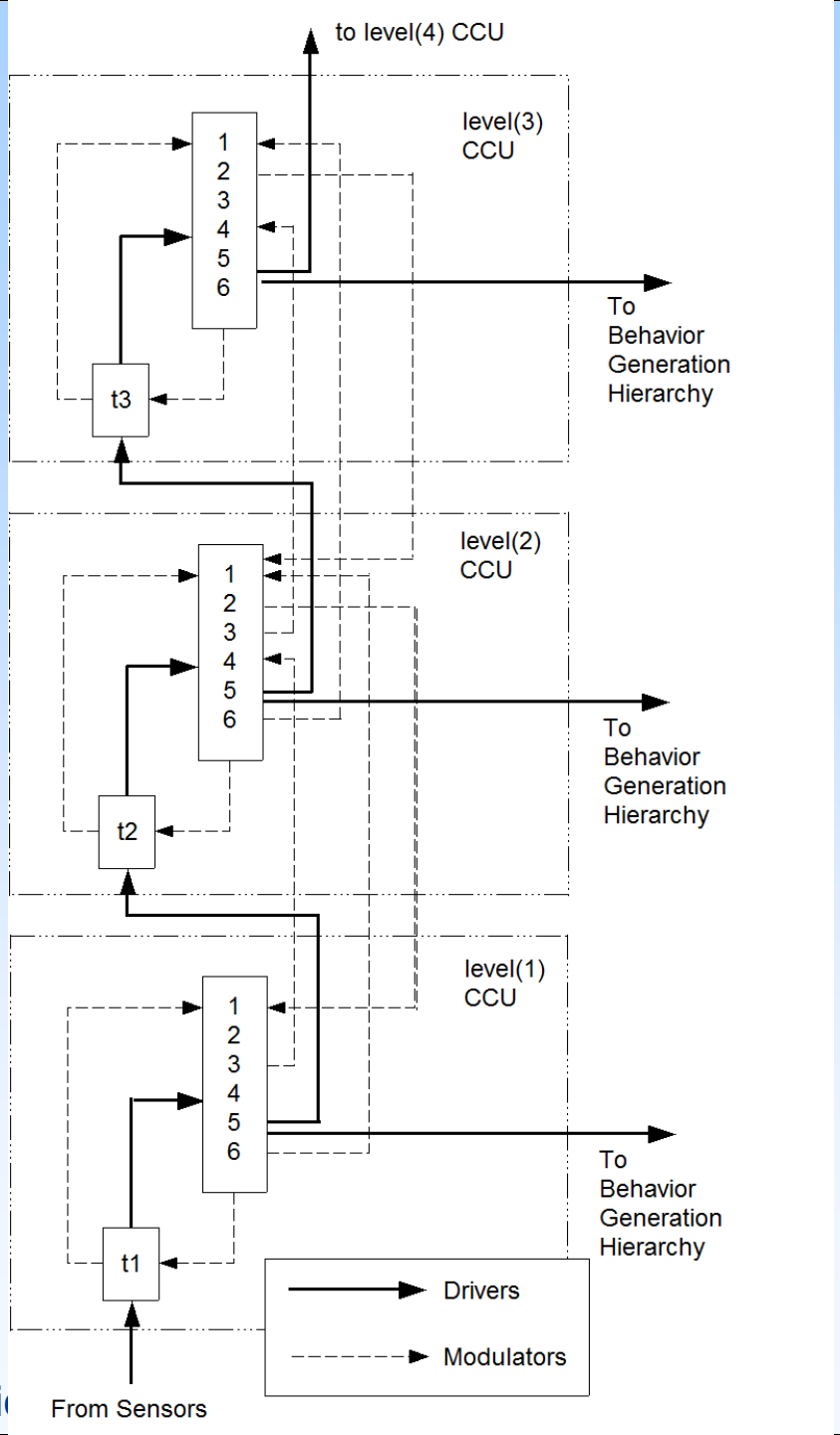


A Cortical Computational Unit (CCU)

windowing
segmentation & grouping
compute group attributes
recursive filtering
classification

Posterior Cortico-Thalamic Loop Hierarchy

windowing
segmentation & grouping
compute group attributes
recursive filtering
classification
at each level



Two types of hierarchies

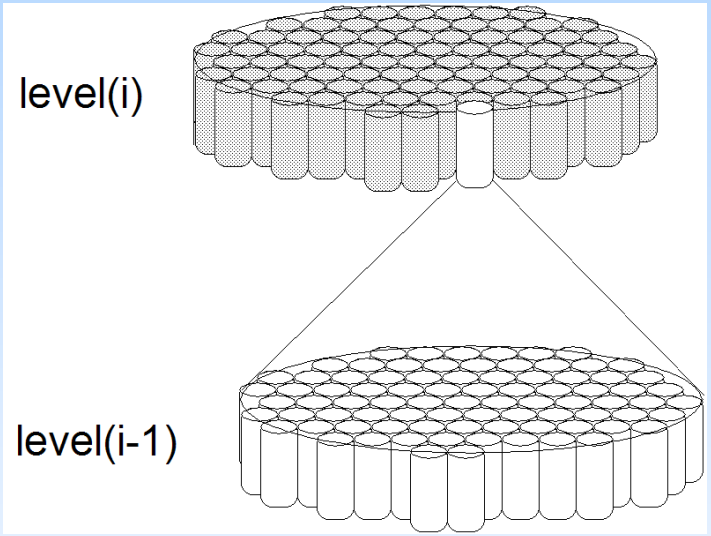
1. **Receptive field hierarchies**

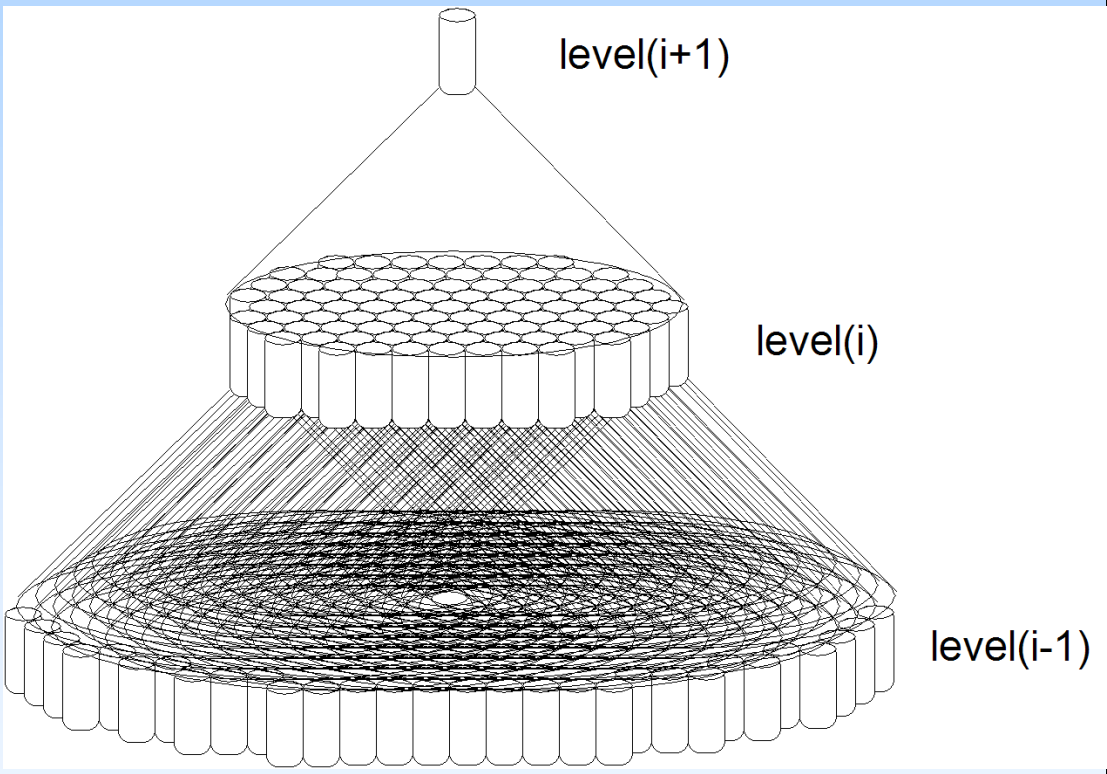
   Receptive field hierarchies are defined by driver anatomical connectivity and are relatively fixed

2. **Entity and Event hierarchies**

   Entity and event hierarchies are defined by modulator activity that can establish or break *belongs-to* and *has-part* pointers in ~ 10 ms

CCU Receptive Field Hierarchy





Defined by driver neurons flowing up the processing hierarchy

CCU Entity/Event Hierarchy

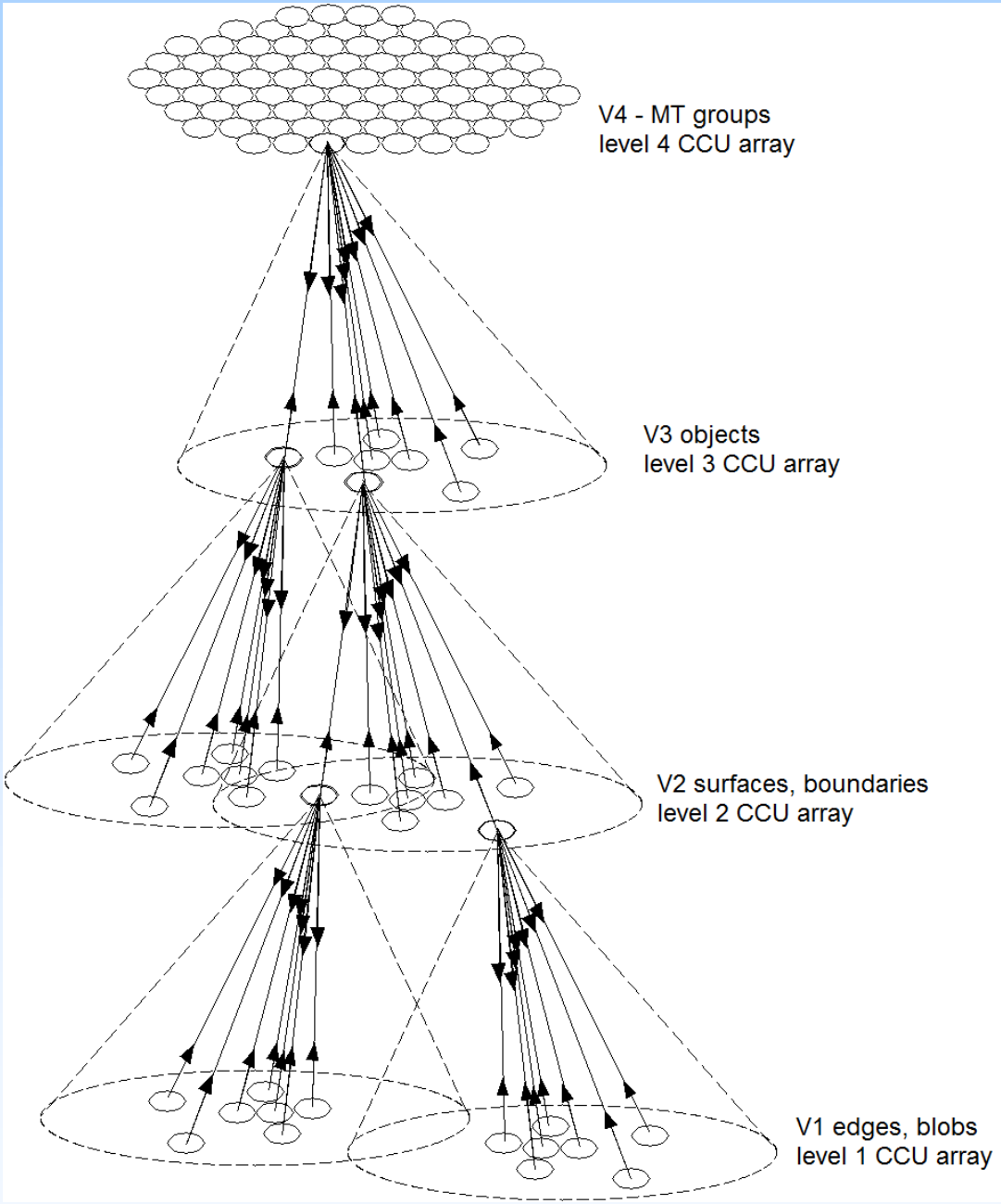
Defined by pointers

result of segmentation & grouping processes

Pointers link pixels to symbols & vice versa

Provides symbol grounding

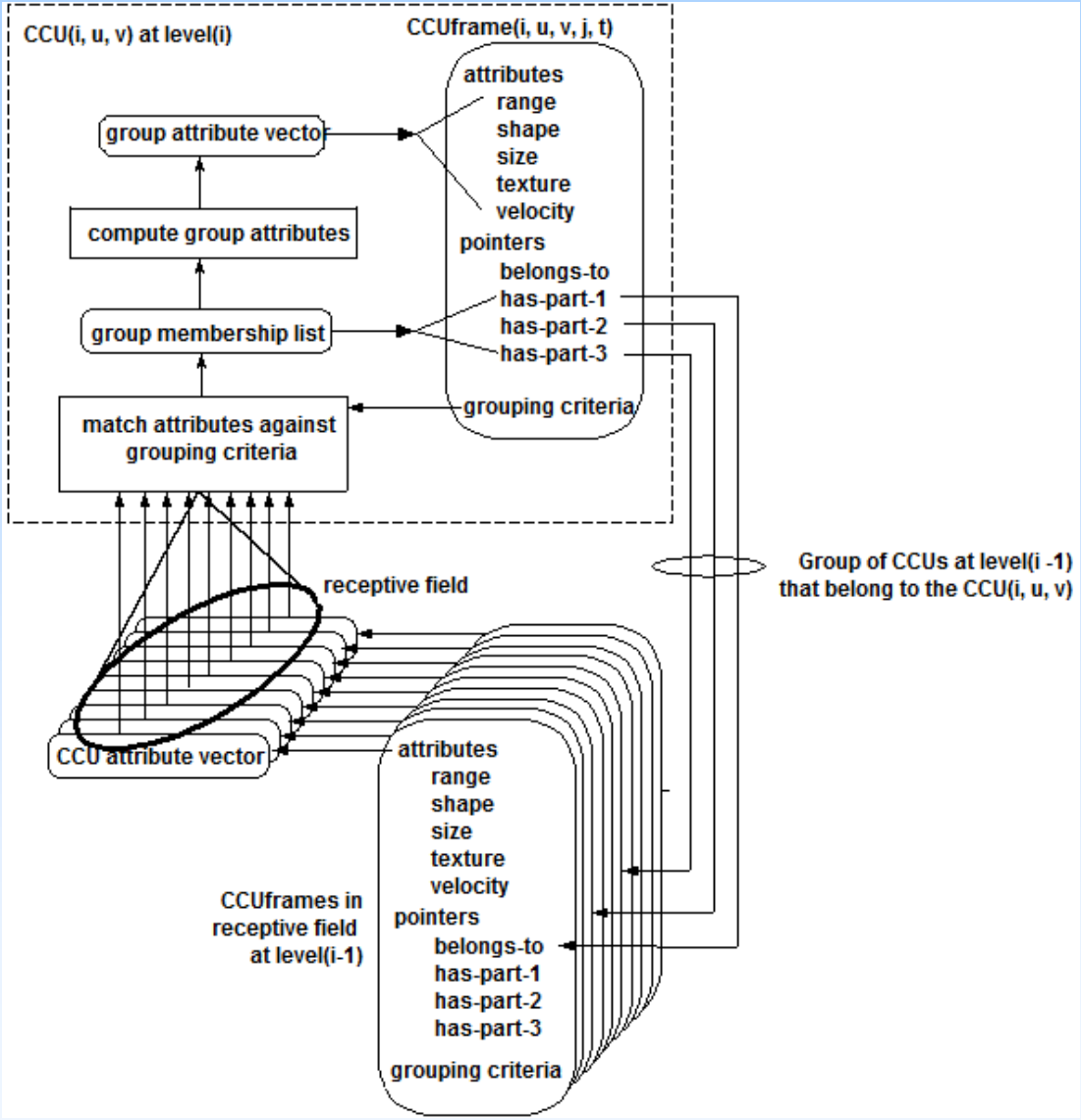
Pointers reset top to bottom within a saccade ~ 150 ms



Segmentation & Grouping Process

Each level detects patterns within its receptive field & sets or breaks pointers

This produces an Entity/Event Hierarchy

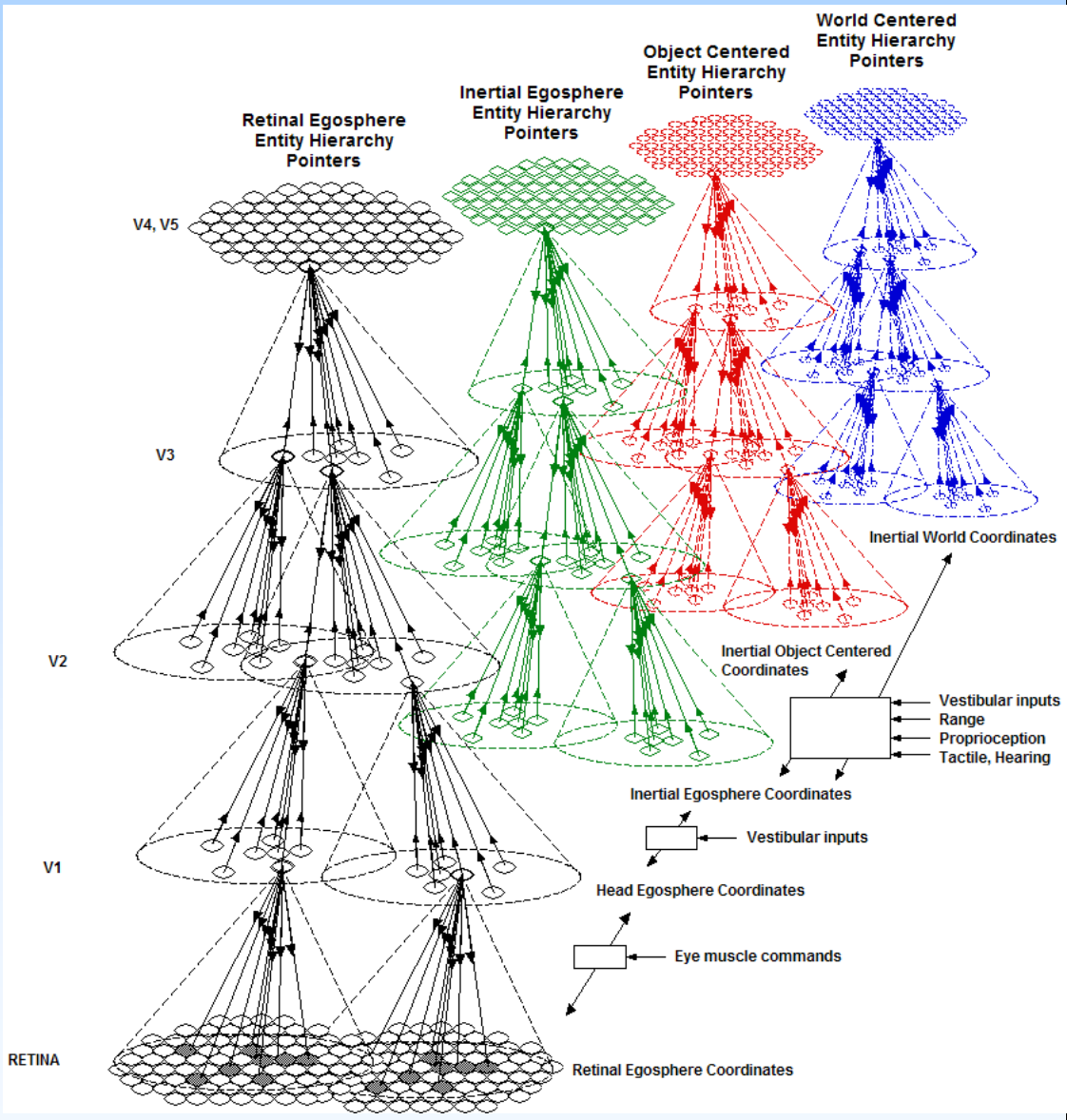


CCU Coordinate Frames

Each level has multiple coordinate frames

Oculomotor signals, vestibular inputs, & range estimates provide transform parameters

Coordinate transforms computed in parallel at each level in ~ 10 ms



World Centered Hierarchy of Entity Pointers

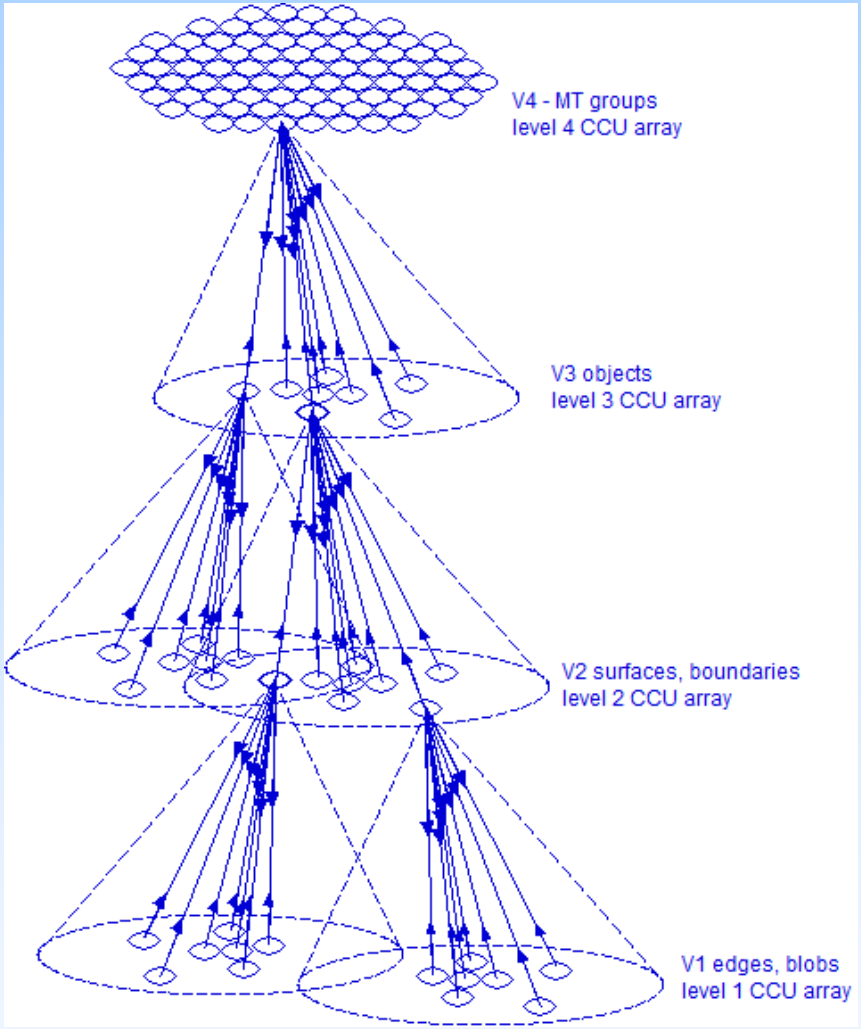
Coordinate frame is fixed in the world

Entities have continuity in space & time

Entities have state
– position, orientation, velocity

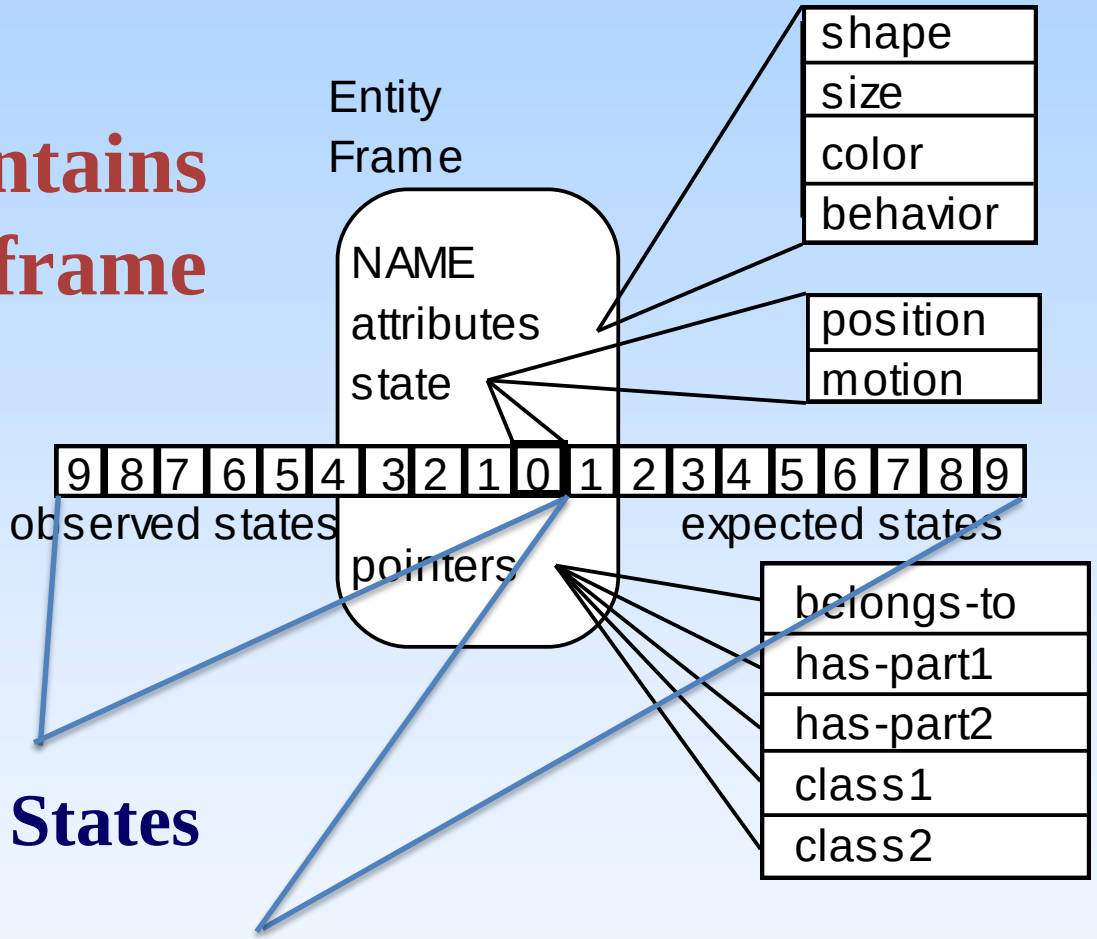
Entities have attributes
-- size, shape, color, texture, behavior

Entities have relationships
-- class, rank, spatial, temporal, causal



Temporal Continuity for an Entity

Each CCU contains an entity frame



Each entity frame contains:

a Trace of Observed States

&

a Prediction of Expected States

Reverse Engineering the Brain

What does that mean?

Building computational machines that are functionally equivalent to the brain

in their ability to perceive, think, decide, and act in a purposeful way to achieve goals in complex, uncertain, dynamic, and possibly hostile environments, despite unexpected events and unanticipated obstacles, while guided by internal values and rules of conduct.

Functional equivalence ::= producing the same input/output behavior

Reverse Engineering the Brain

Will require a deep understanding of how the brain works and what the brain does

- **How is information represented in the brain?**
- **How is computation performed?**
- **What are the functional operations?**
- **What are the knowledge data structures?**
- **How are messages encoded?**
- **How are images processed?**
- **How are relationships established and broken?**
- **How are signals transformed into into symbols?**

How does the brain generate the incredibly complex colorful, dynamic internal representation that we consciously perceive as external reality?

Reverse Engineering the Brain

Cited as a Grand Challenge by
U.S. National Academy of Engineering

Focus of DARPA SyNAPSE & NEOVISION2, IARPA ICArUS, & European Union FACETS & POETICON Programs

Decade of the Mind Initiative

A proposed 10 year $4B program
to understand the mechanisms of mind

Krasnow Lead
Steering Committee of Top Scientists
Recent workshops at:

Sandia National Labs Jan 13 –15, 2009
Berlin Sept 10 – 12, 2009
Singapore October 18-20, 2010

Why Now?

The science & technology is ready

Neurosciences – computation and representation in the brain

- biochemistry, synaptic transmission, brain imaging, neuron modeling
- neuroanatomy, neurophysiology, network & whole brain models

Cognitive Modeling – representation and use of knowledge

- mathematics, logic, language, learning, problem solving
- psychophysics, cognitive psychology, functional brain modeling

Intelligent Control – making machines behave appropriately

- control theory, cybernetics, AI, knowledge representation, planning
- manipulation, locomotion, manufacturing, vehicles, weapons

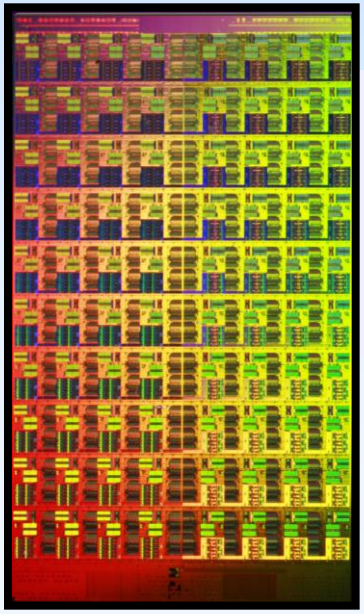
Computational Power – speed and memory that rival the brain

- supercomputer = 10^{15} ops today, laptop > 10^{15} ops in 20 years

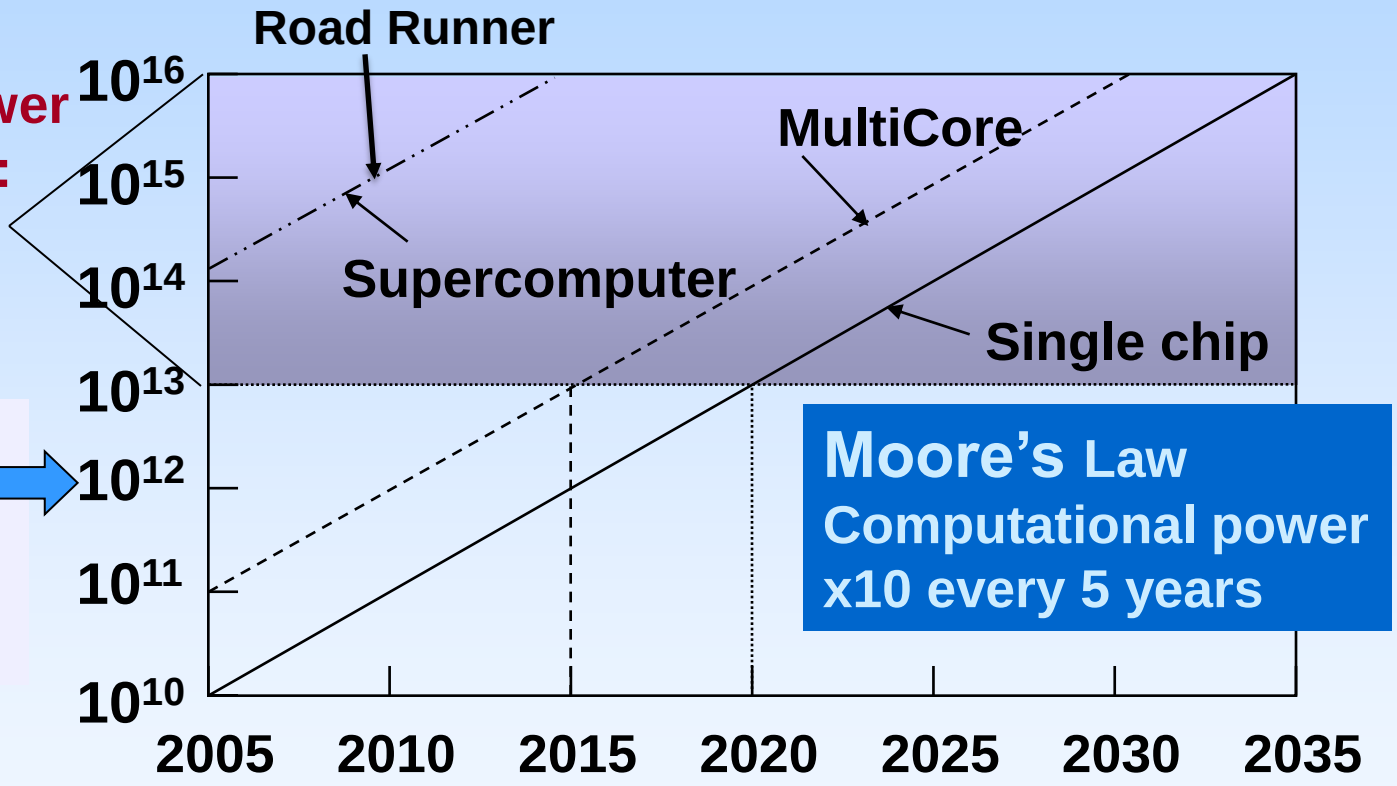
Progress is rapid, an enormous literature

Computational Power is Approaching a Critical Threshold

Estimated computing power of human brain: ~10^{13}-10^{16} ops



teraflop chip



Moore's Law
Computational power x10 every 5 years

Money Is Flowing

Military – FCS, UGV, UAV, UUV, USV, UGS
Commercial – manufacturing, transportation
Entertainment – video games, cell phones
Academic – neuroscience, computer science

Billions of $ will be invested over the next decade

Intelligent Machines Will Be Critical for Military Security and Economic Prosperity

What is the path to success for reverse engineering the brain?

Pick the right level of resolution

- **overall system level** (central nervous system)

AI and Cognitive Psychology

- **arrays of macro-computational units** (e.g., cortical regions)
- **macro-computational units** (e.g., cortical hypercolumns & loops)

CCUs

- **micro-computational units** (e.g., cortical microcolumns & loops)
- **neural clusters** (e.g., spinal and midbrain sensory-motor nuclei)
- **neurons** (elemental computational units) – input/output functions

Mainstream Neuroscience & Neural Nets

- **synapses** (electronic gates, memory elements) – synaptic phenomena
- **membrane mechanics** (ion channel activity) – molecular phenomena

Computational Requirements
for Engineering Human Brain at CCU level

10^6 CCUs running at 100 Hz requires
10^8 CCU modeling cycles per second

State of art supercomputer running at 10^{15} fops provides 10 million fops per CCU modeling cycle

Estimated communication load between CCUs
10^6 bytes per second for each CCU, or
10^{12} bps for full brain model

This appears to be within the state of the art for current supercomputers

Summary & Conclusions

Reverse Engineering the Human Brain appears feasible in the near term

- **The science and engineering can support it**
- **The benefits will justify the investment**
- **Near term success will require selecting the right level of resolution, e.g. CCU level**
- **Real-time modeling at CCU level of resolution appears achievable *now* with supercomputers and in ~ 20 years with laptop class computers**

The impact will be revolutionary

Thank you

Questions?

james.albus@gmail.com