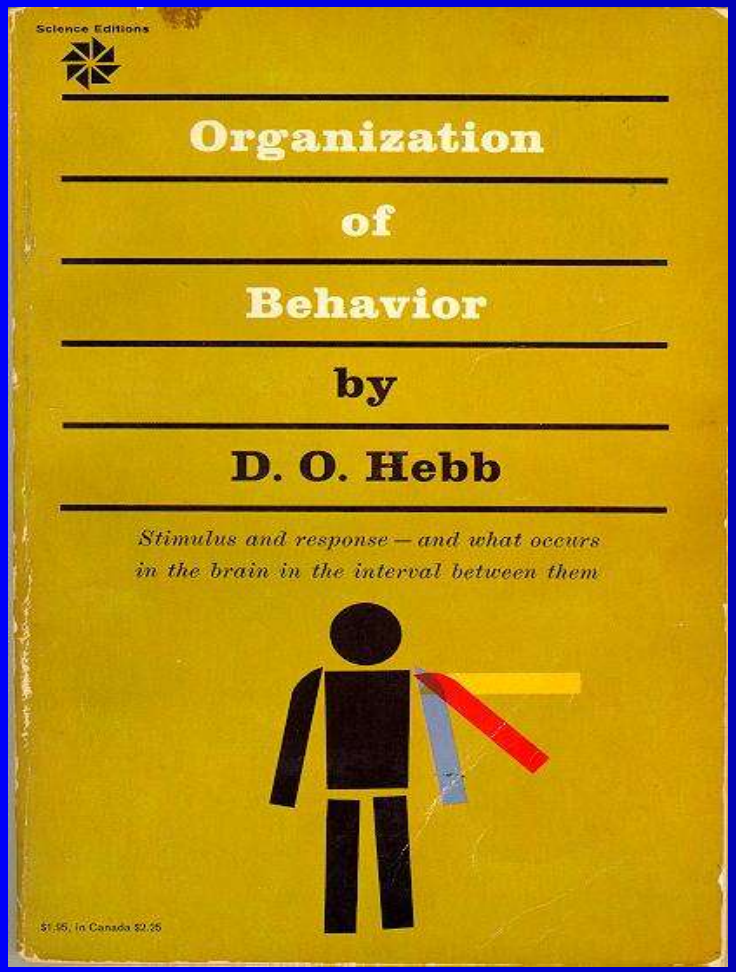
From AI & BI to NN & CI

- Roots of a new discipline, 1960s to 1987
  - From Biological Intelligence(Hebb): universal learning
  - From AI: reinforcement learning & MLP pattern recog
- Early NSF 1988-1992 (as DARPA tries SYNAPS)
  - NN = parallel hardware? Optics to Mead. ATT Zip.
  - Neurocontrol: a new paradigm; The Terminator movies
- Main history 1993-2006
  - China, Learning & Intelligent Systems, “3 brains in 1”
  - Massive new apps – hit-to-kill, MD11, Ford, CC, coal
- But where are we now?
  - Deep learning in AI, grid, no full PhD, Winter Soldier

Hebb 1949: Intelligence As An Emergent Phenomenon or Learning



"The general idea is an old one, that any two cells or systems of cells that are especially active at the same time will tend to become 'associated,' so that activity in one facilitates activity in the other" -- p.70 (Wiley 1961 printing)

The search for the General Neuron Model (of Learning)

"Solves all problems"

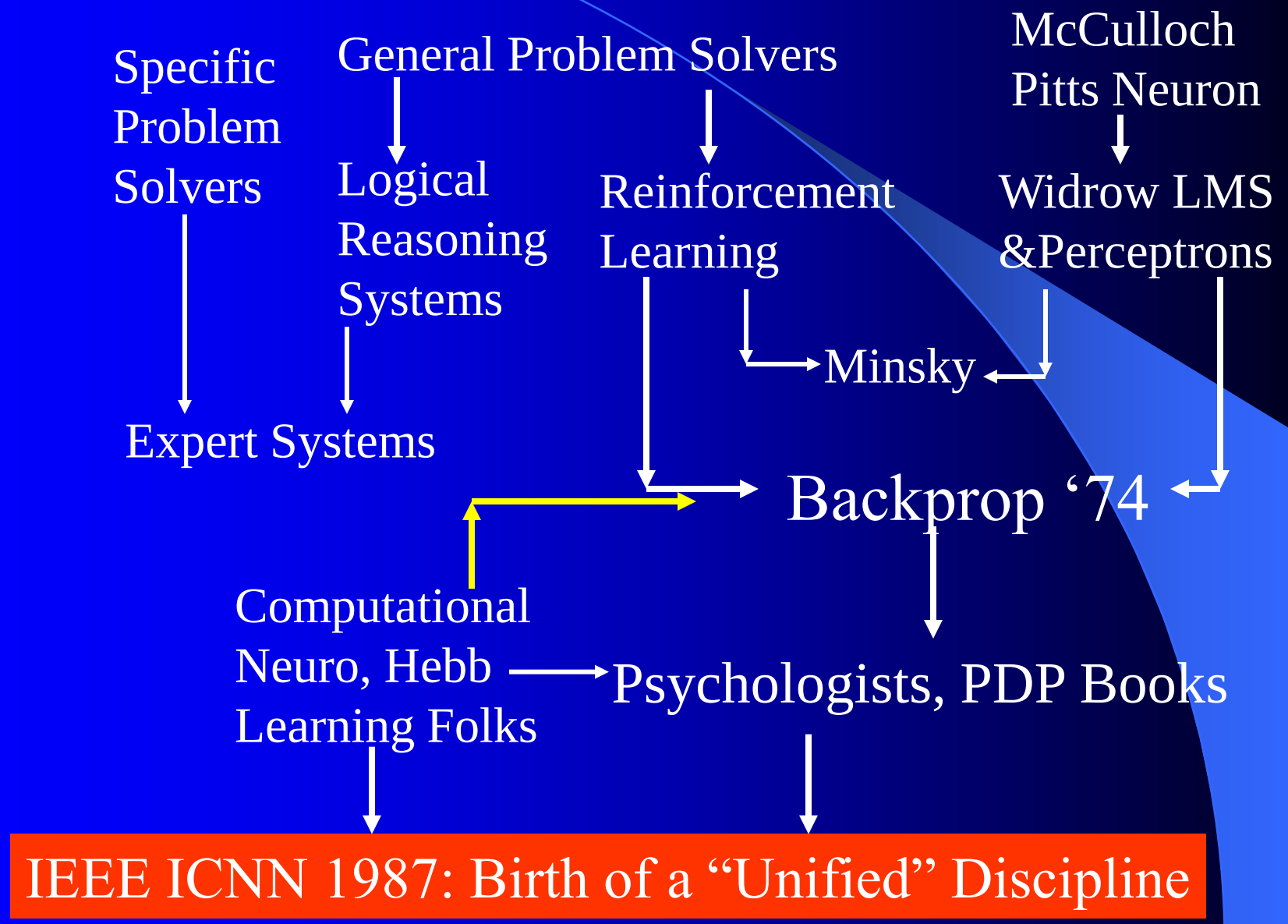
Claim (1964) : Hebb's Approach Doesn't Quite Work As Stated

- Hebbian Learning Rules Are All Based on Correlation Coefficients
- Good Associative Memory: one component of the larger brain (Kohonen, ART, Hassoun)
- Linear decorrelators and predictors
- Hopfield f(u) minimizers never scaled, but:
  - Gursel Serpen and SRN minimizers
  - Brain-Like Stochastic Search (Needs R&D)

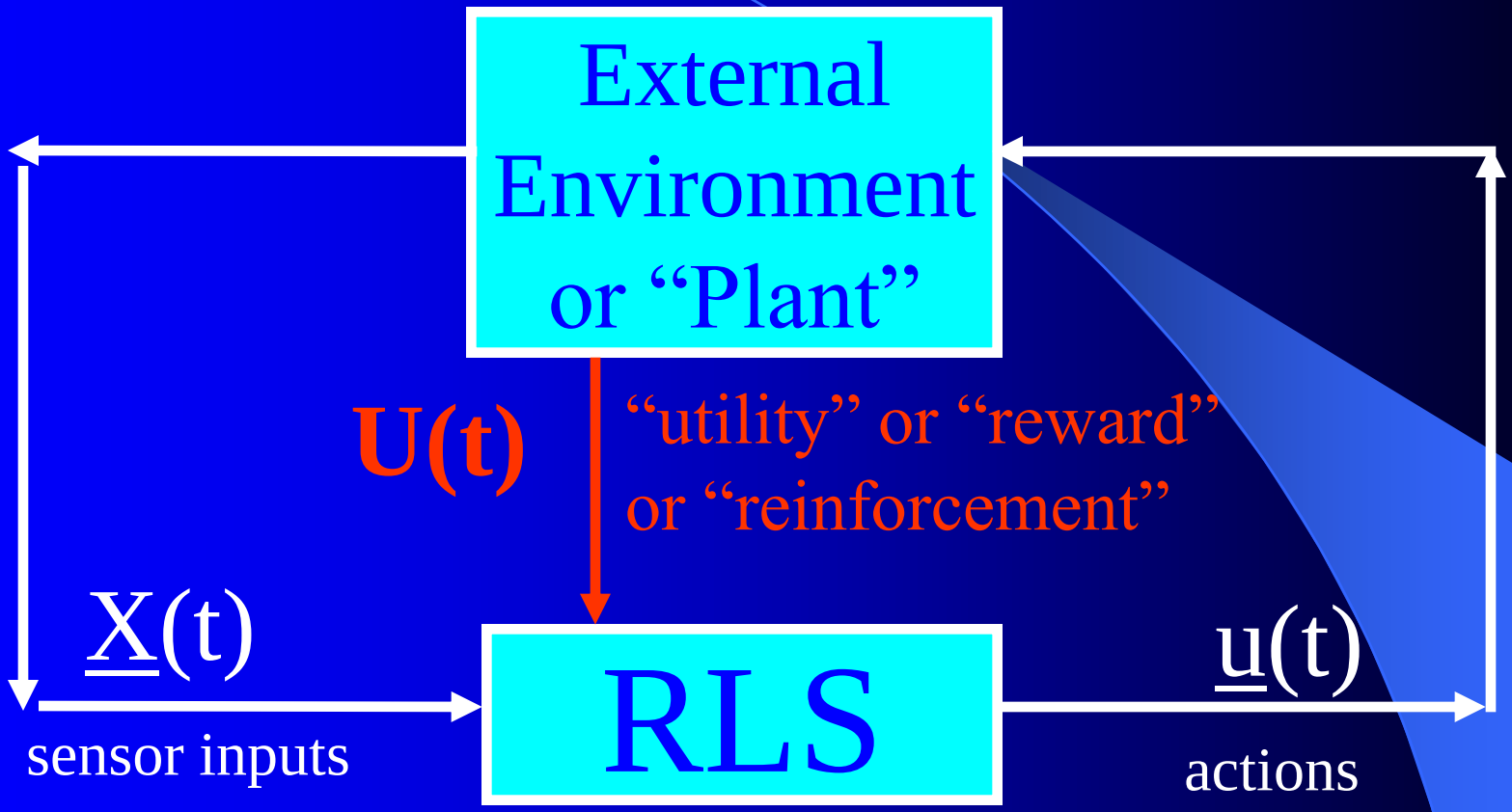
Understanding Brain Requires Models Tested/Developed Using Multiple Sources of Info

- Engineering: Will it work? Mathematics understandable, generic?
- Psychology: Connectionist cognitive science, animal learning, folk psychology
- Neuroscience: computational neuroscience
- AI: agents, games (backgammon, go), etc.
- LIS and CRI

Where Did ANNs Come From?

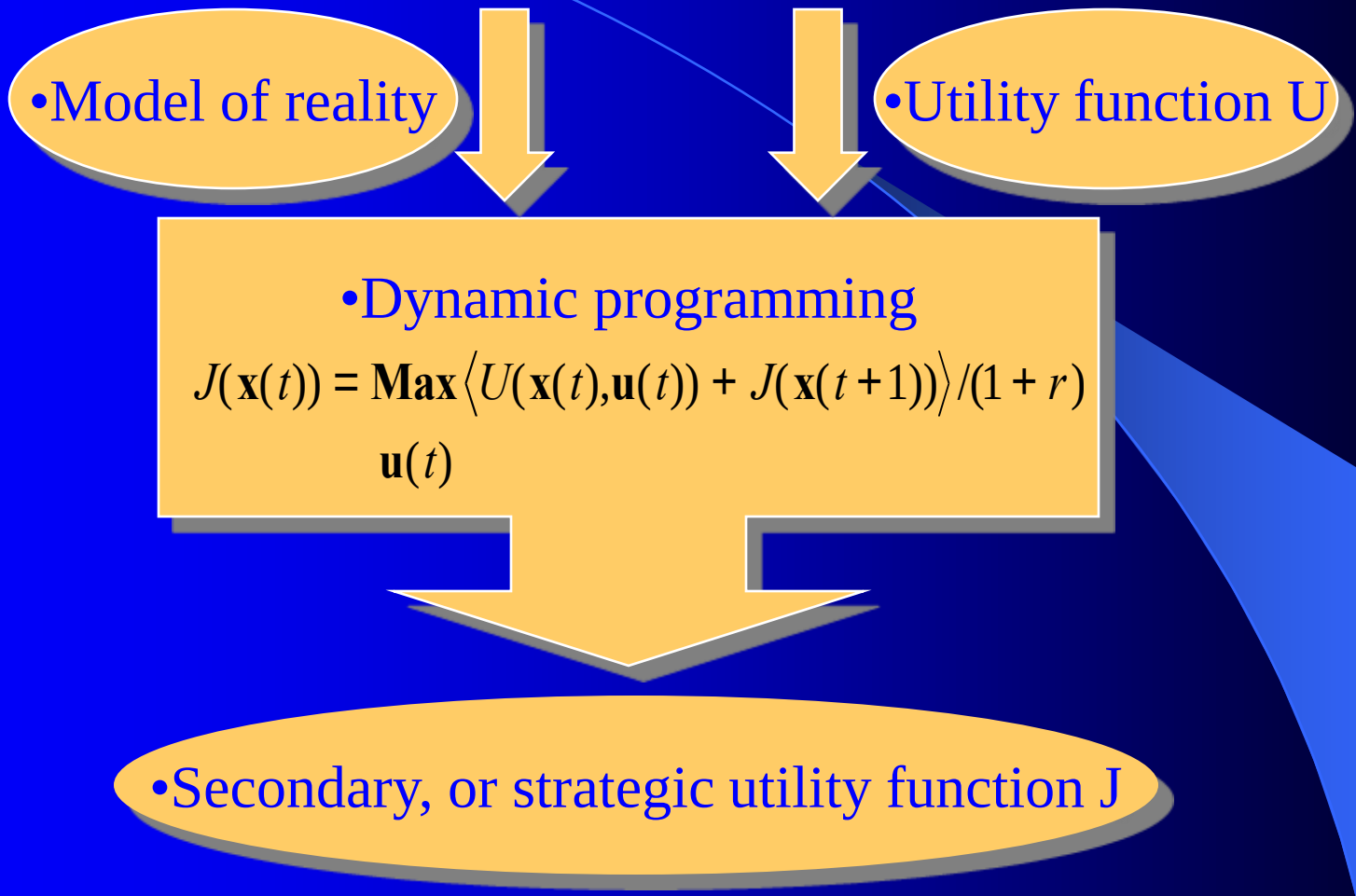


Reinforcement Learning Systems (RLS)



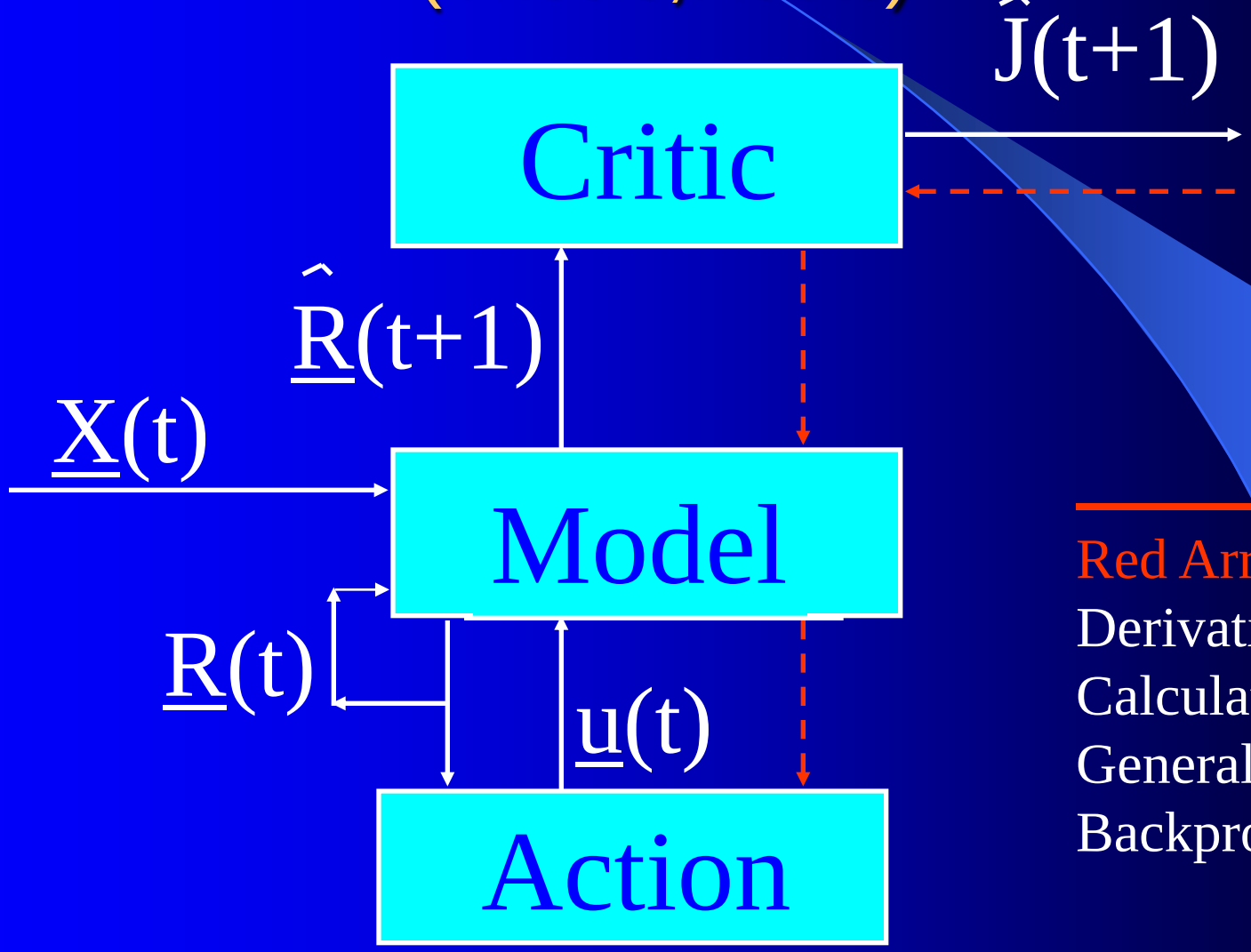
RLS may have internal dynamics and "memory" of earlier times t-1, etc.

Minsky failed but ADP Succeeded



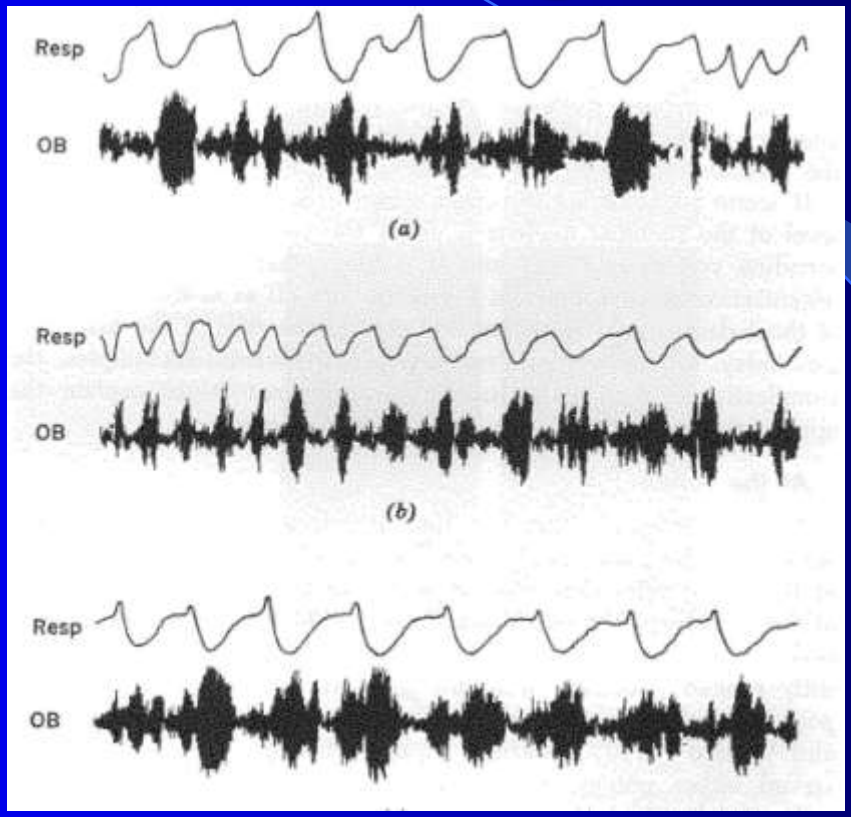
- See Lewis and Liu, 2012,
- and CLION video lectures

1971-2: Emergent Intelligence Is Possible If We Allow Three Types of Neuron (Thesis,Roots)



Red Arrows: Derivatives Calculated By Generalized Backpropagation

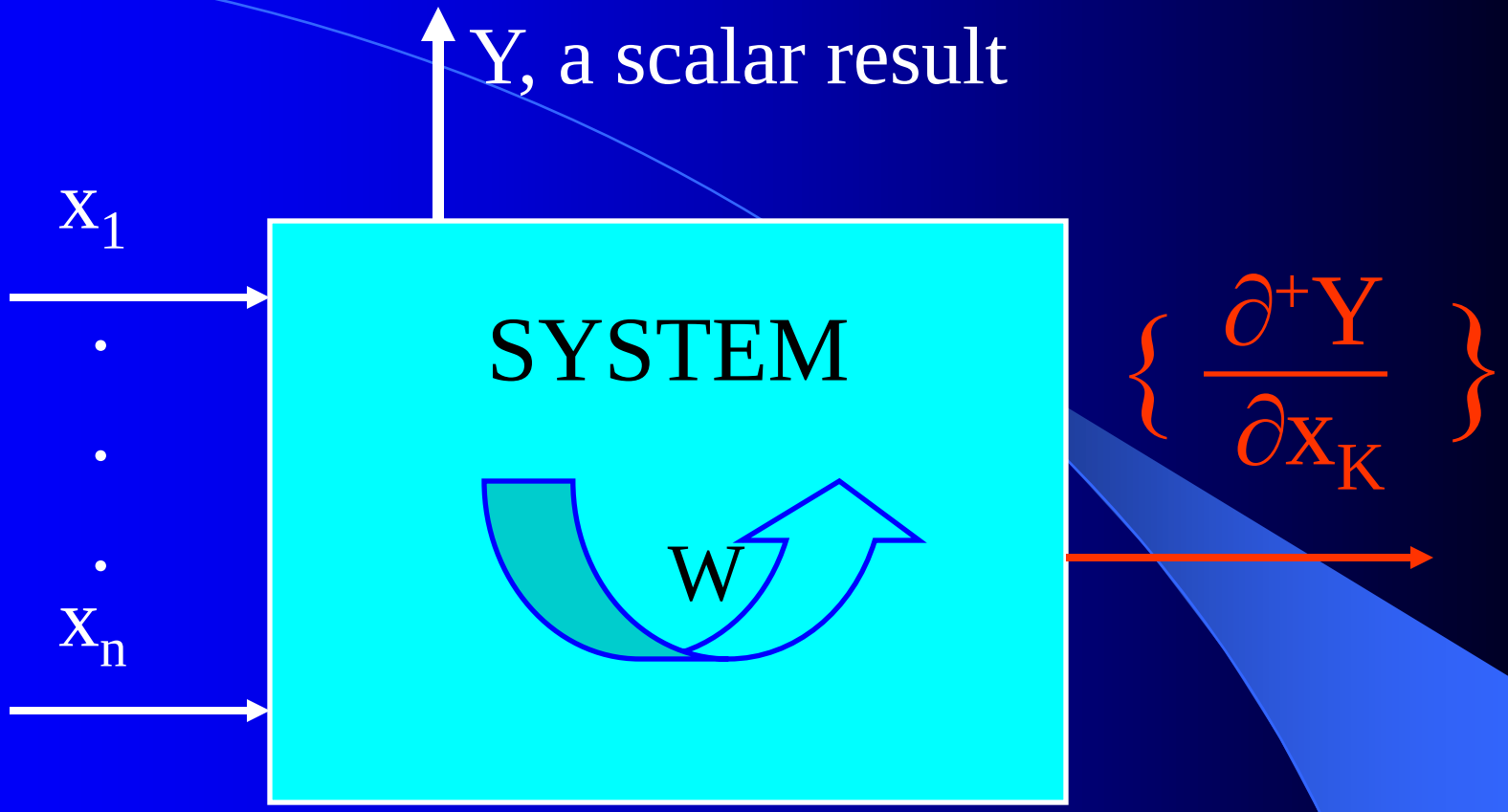
Offer to Minsky to Coauthor BP/TLU (see Talking Nets)



- Real neurons are not 1/0 asynchronous binary digits! Every 100 ms or so, a “volley” of continuous intensity. Clocks, Richmond, LLinas

Harvard Committee Response

- We don't believe in neural networks – see Minsky (Anderson&Rosenfeld, Talking Nets)
- Prove that your backwards differentiation works. (That is enough for a PhD thesis.) The critic/DP stuff published in '77,'79,'81,'87..
- Applied to affordable vector ARMA statistical estimation, general TSP package, and robust political forecasting

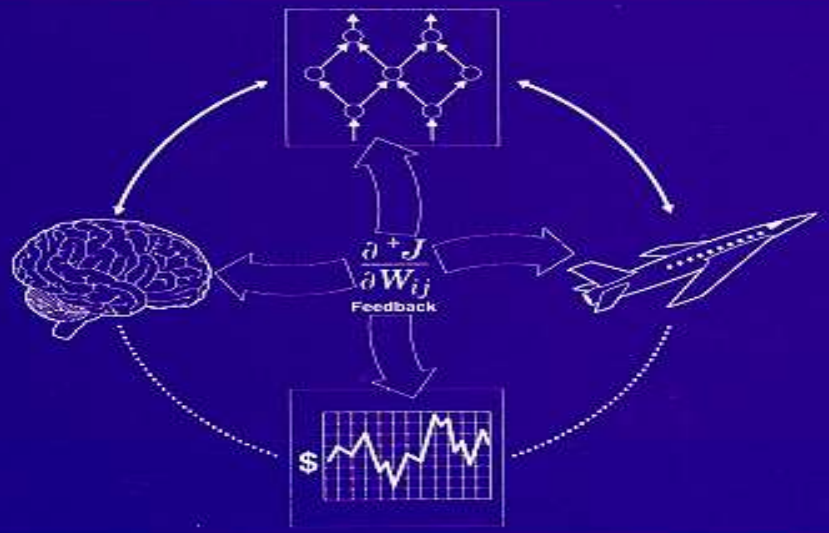


(Inputs x_k may actually come from many times)

Backwards Differentiation: But what kinds of SYSTEM can we handle? See details in AD2004 Proceedings, Springer, in press.

THE ROOTS OF BACKPROPAGATION

From Ordered Derivatives to Neural Networks and Political Forecasting



PAUL JOHN WERBOS

A Volume in the Wiley Series on ADAPTIVE AND LEARNING SYSTEMS FOR SIGNAL PROCESSING, COMMUNICATIONS, AND CONTROL
SIMON HAYKIN, SERIES EDITOR

National Science Foundation

Engineering Directorate

Computer & Info. Science Directorate

ECS

DMI

IIS

EPDT: Chips, Optics, Etc.

IHCS: Integration

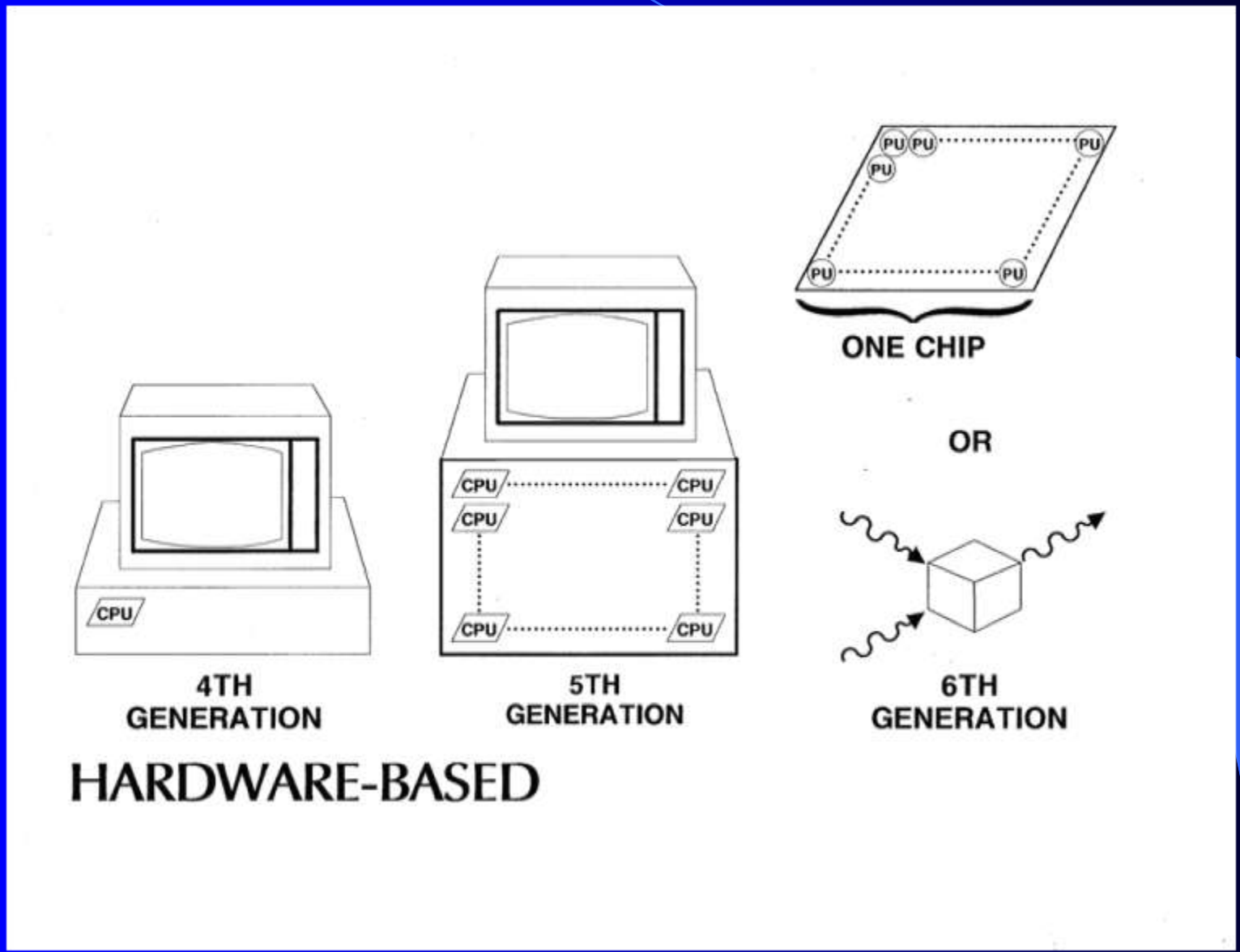
Power, Control, & Adaptive Networks

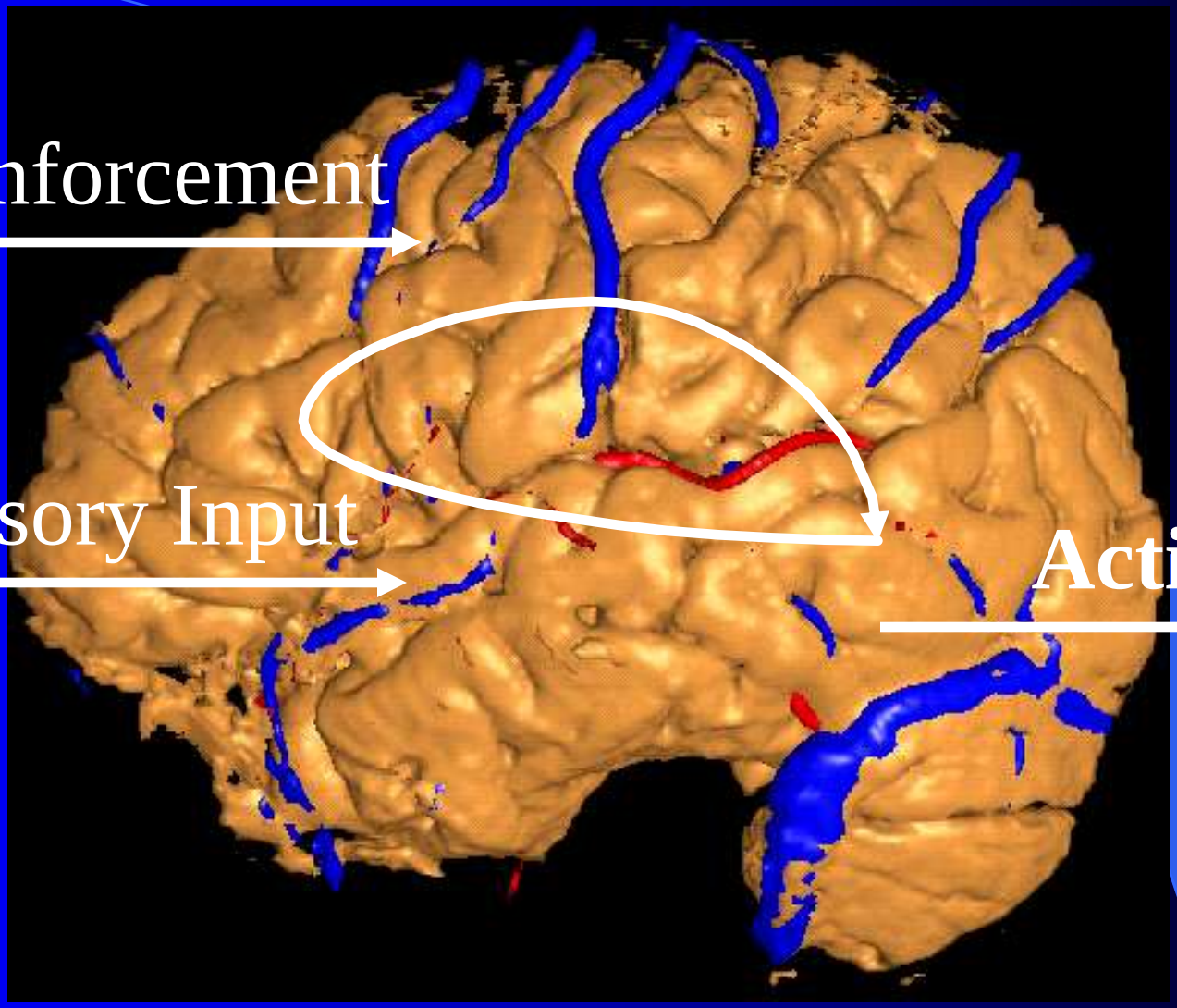
Robotics

AI

New Thrusts on Next Generation Cyberinfrastructure

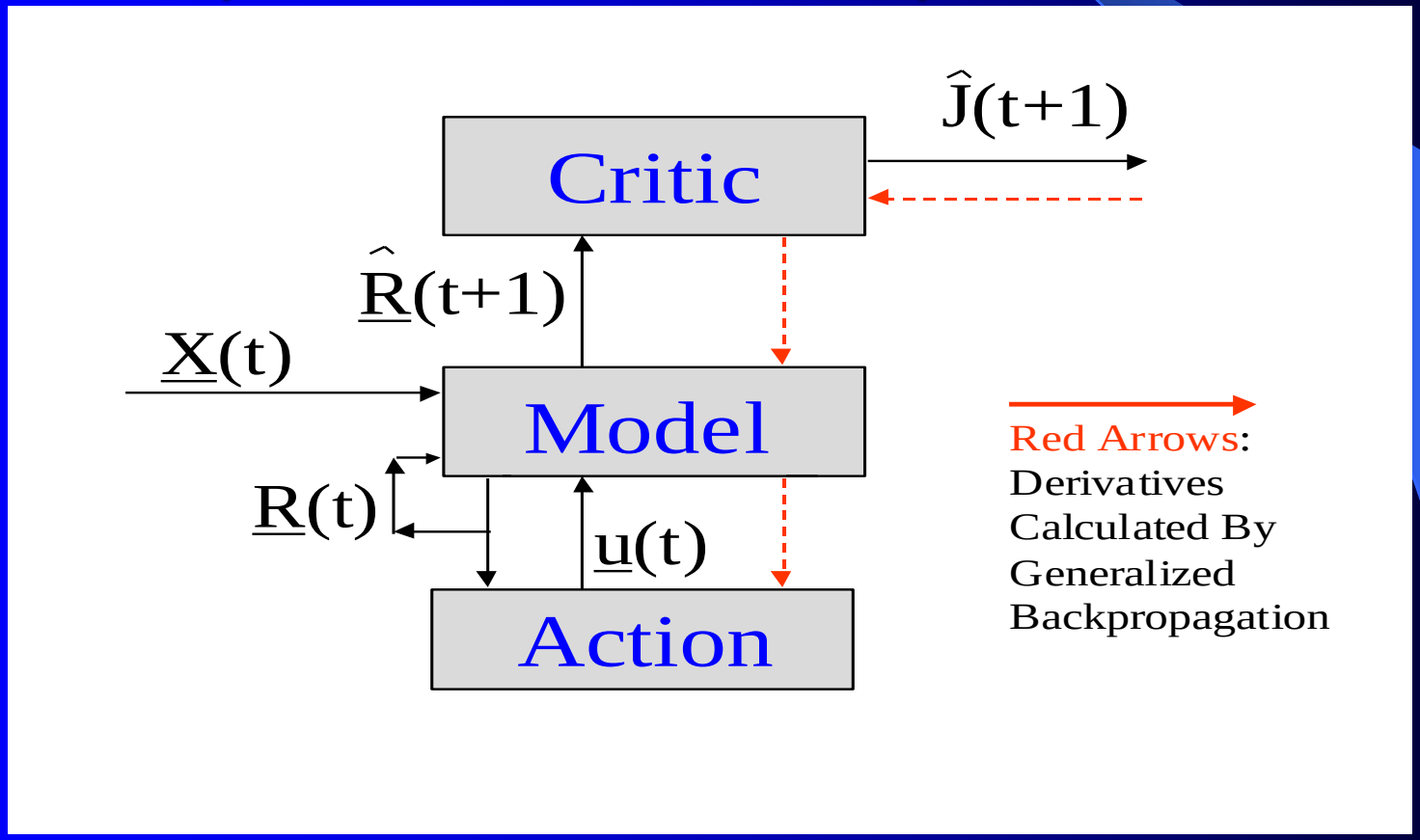
Neural Nets Defined as Brain-like Learning OR as Massively Parallel General Learning



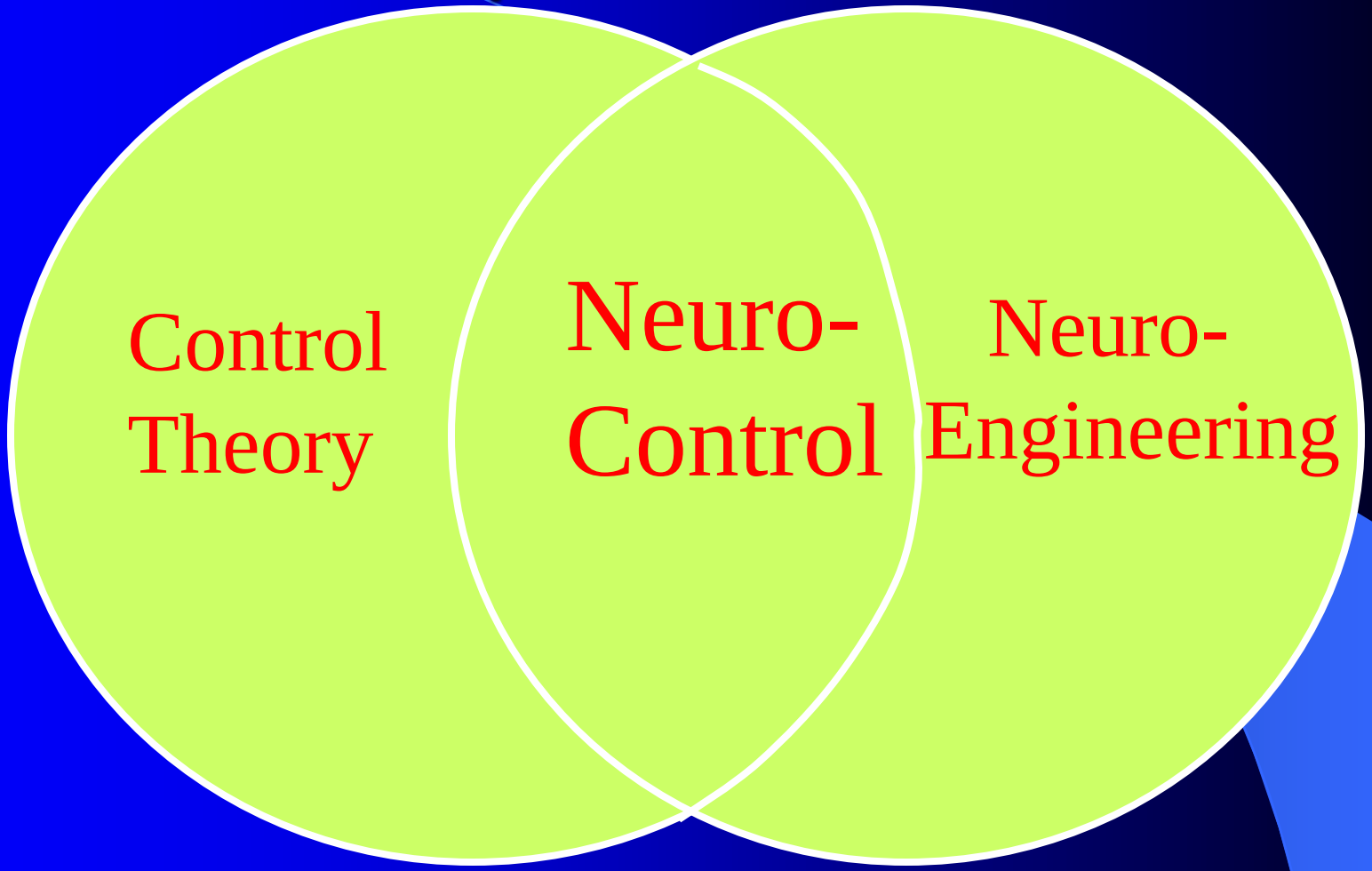


The Brain As a Whole System
Is an Intelligent Controller

To Fill IN the Boxes:
(1) NEUROCONTROL, to Fill in Critic or Action;
(2) System Identification or Prediction (Neuroidentification) to Fill In Model



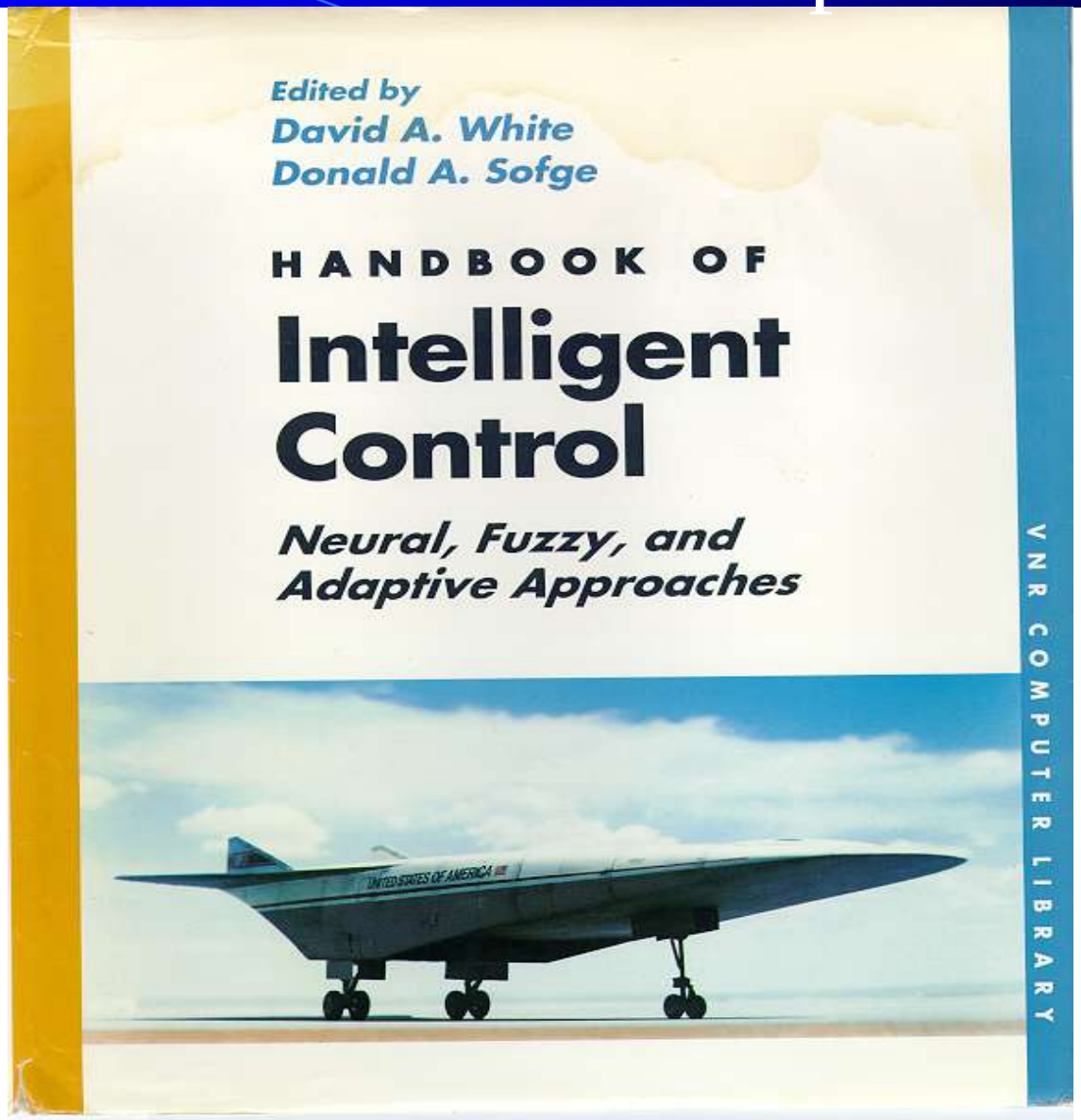
NSF Workshop Neurocontrol 1988



Miller, Sutton, Werbos, MIT Press, 1990

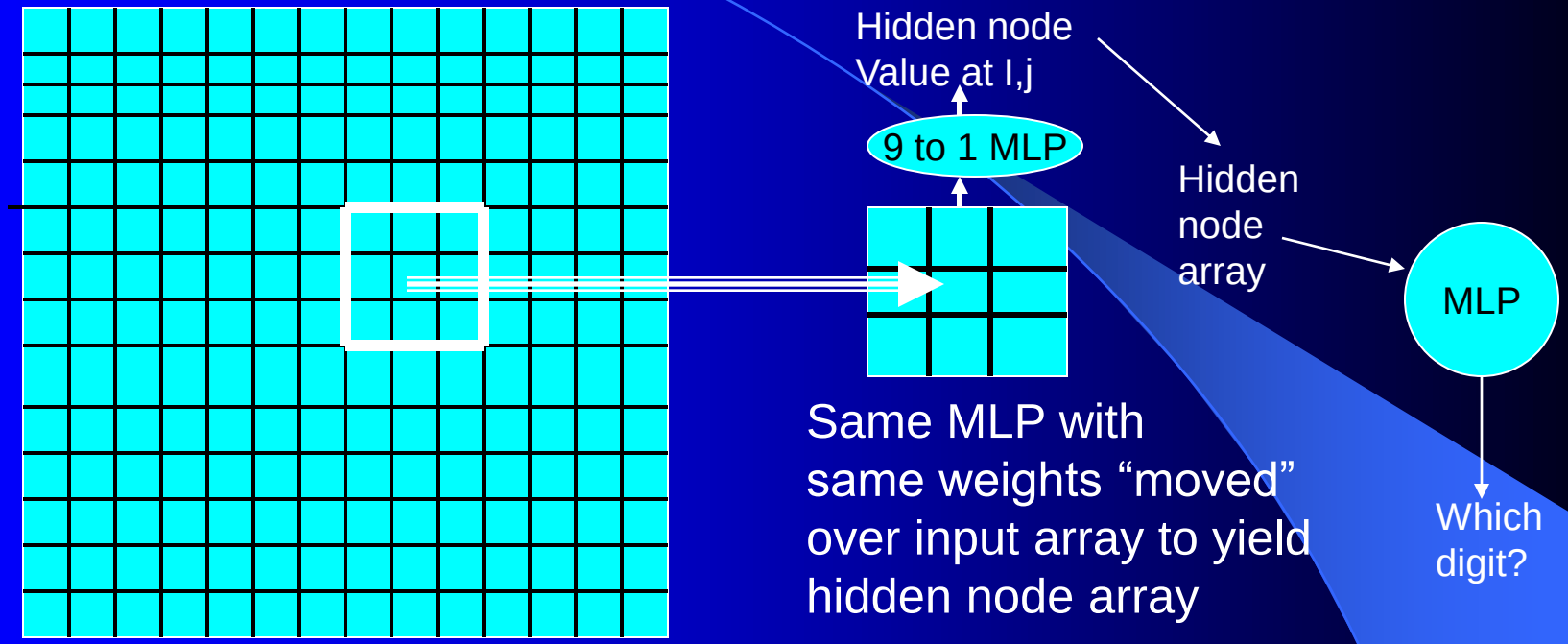
Neurocontrol is NOT JUST Control Theory!

NSF/McAir Workshop 1990



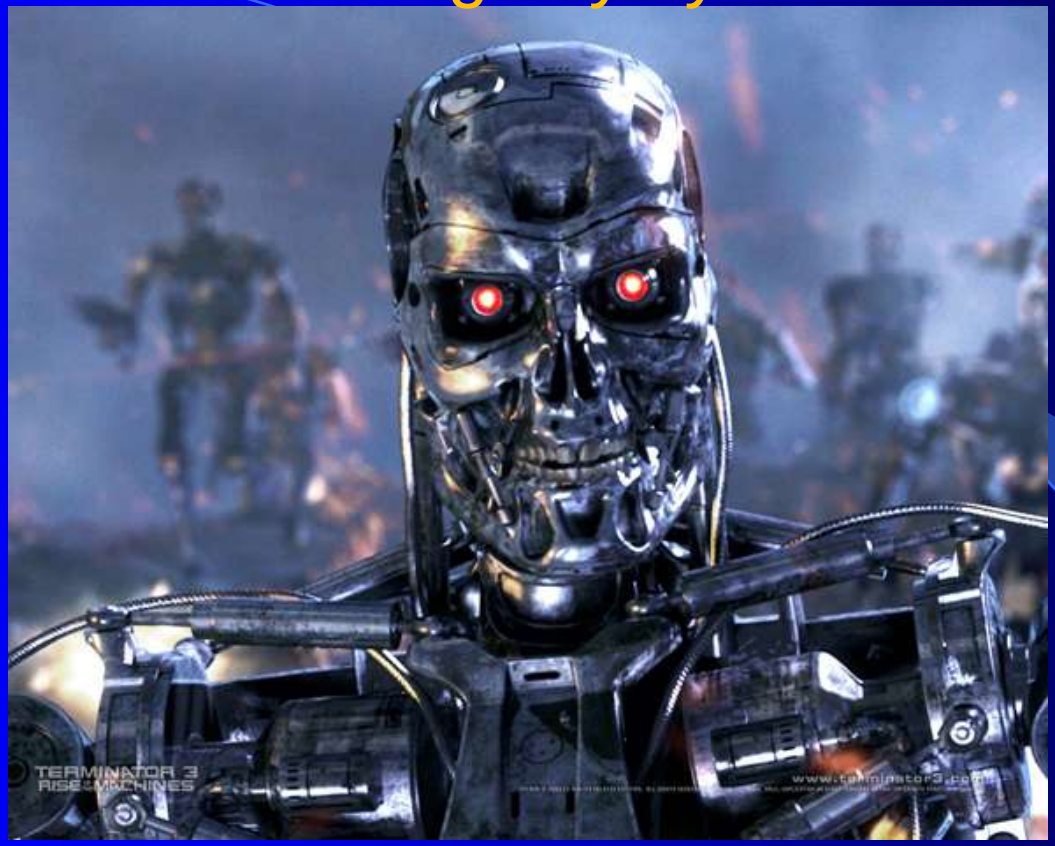
White and Sofge eds, Van Nostrand, 1992

Moving Window Net: Clue Re Complexity LeCun Calls this "Convolutional Neural Net"



- Best ZIP Code Digit Recognizer Used "Moving Window" or "conformal" MLP! (Guyon, LeCun, AT&T story, earlier…)
- Exploiting symmetry of Euclidean translation crucial to reducing number of weights, making large input array learnable, outcomes.
- NEW IN 2010: WORLD'S BEST OBJECT RECOGNIZER!

IJCNN91 Seattle: Boeing says you MUST see Terminator

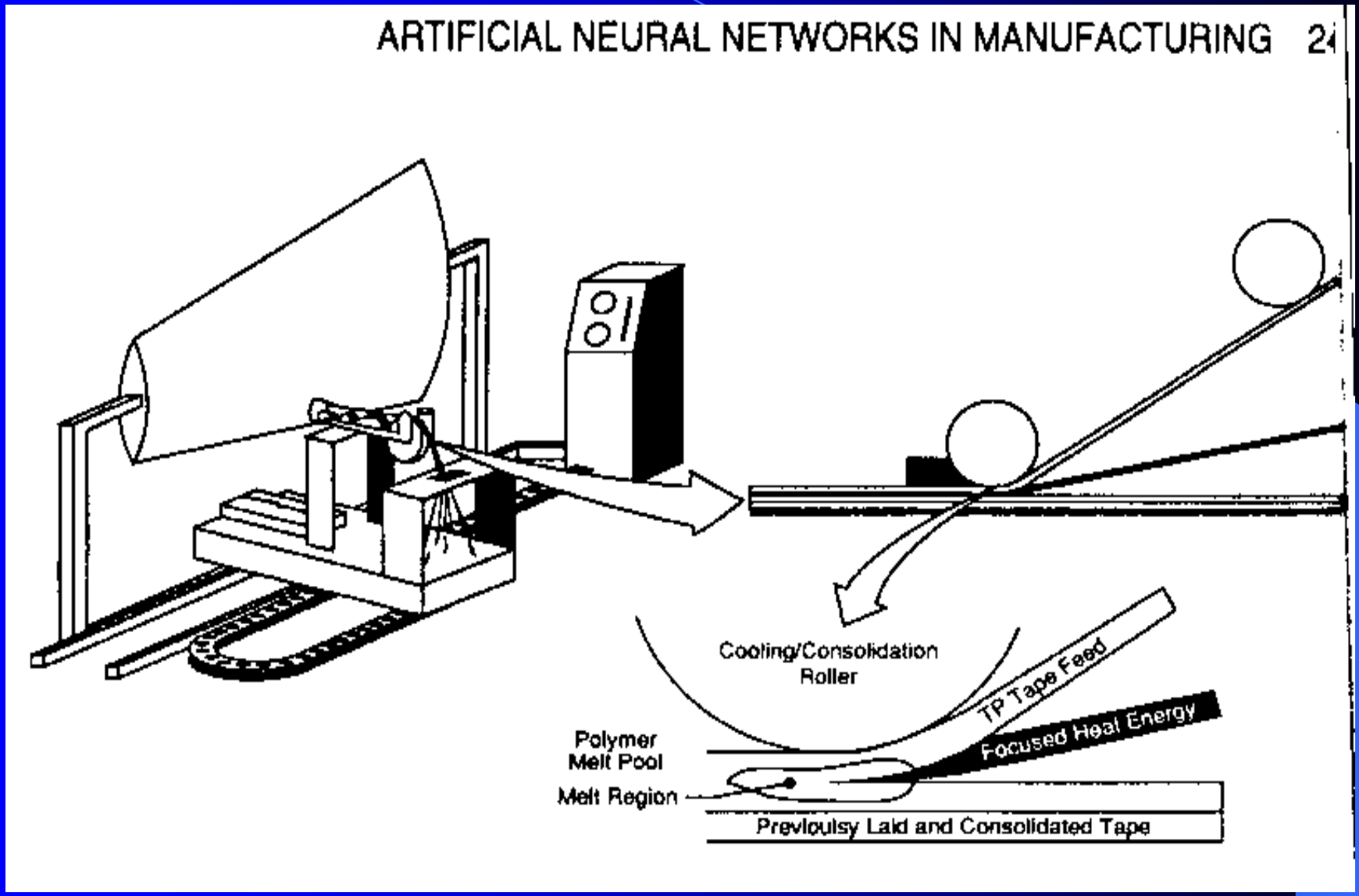


Actual company Cyberdyne/Neurodyne funded by me that week
Bad nano guy a morph between me and Neurodyne
Starts with NN Theater Missile Interception (as in Seattle!)
2 key items – robot arm (award that week) but what of chip?
Schwarzennegger voice – briefing on the ship that week
Movie explains information can be sent backwards through time

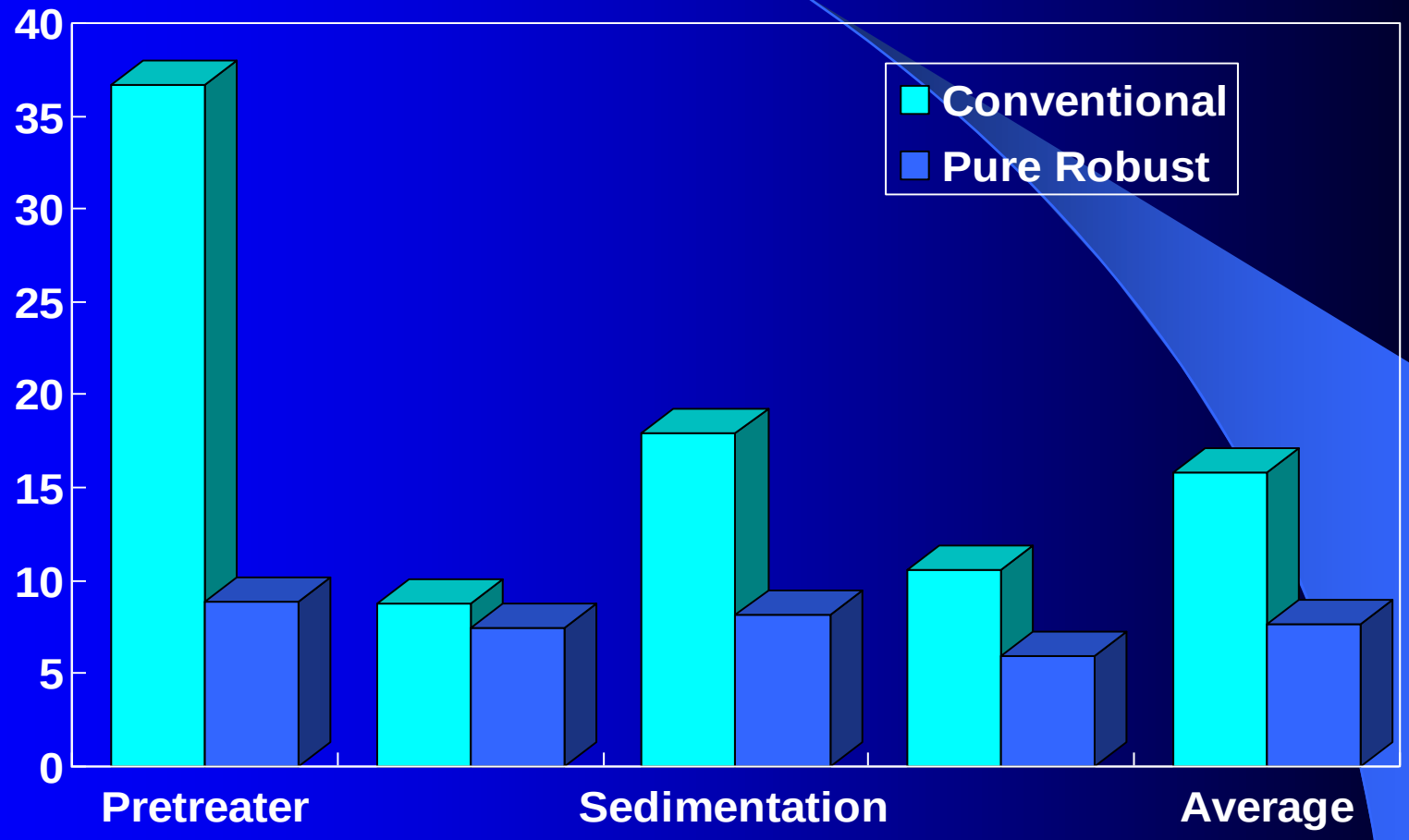
S.N. Balakrishnan: Using DHP, Reduced Error in Hit to Kill Missile Interception more than order of magnitude vs. all previous methods

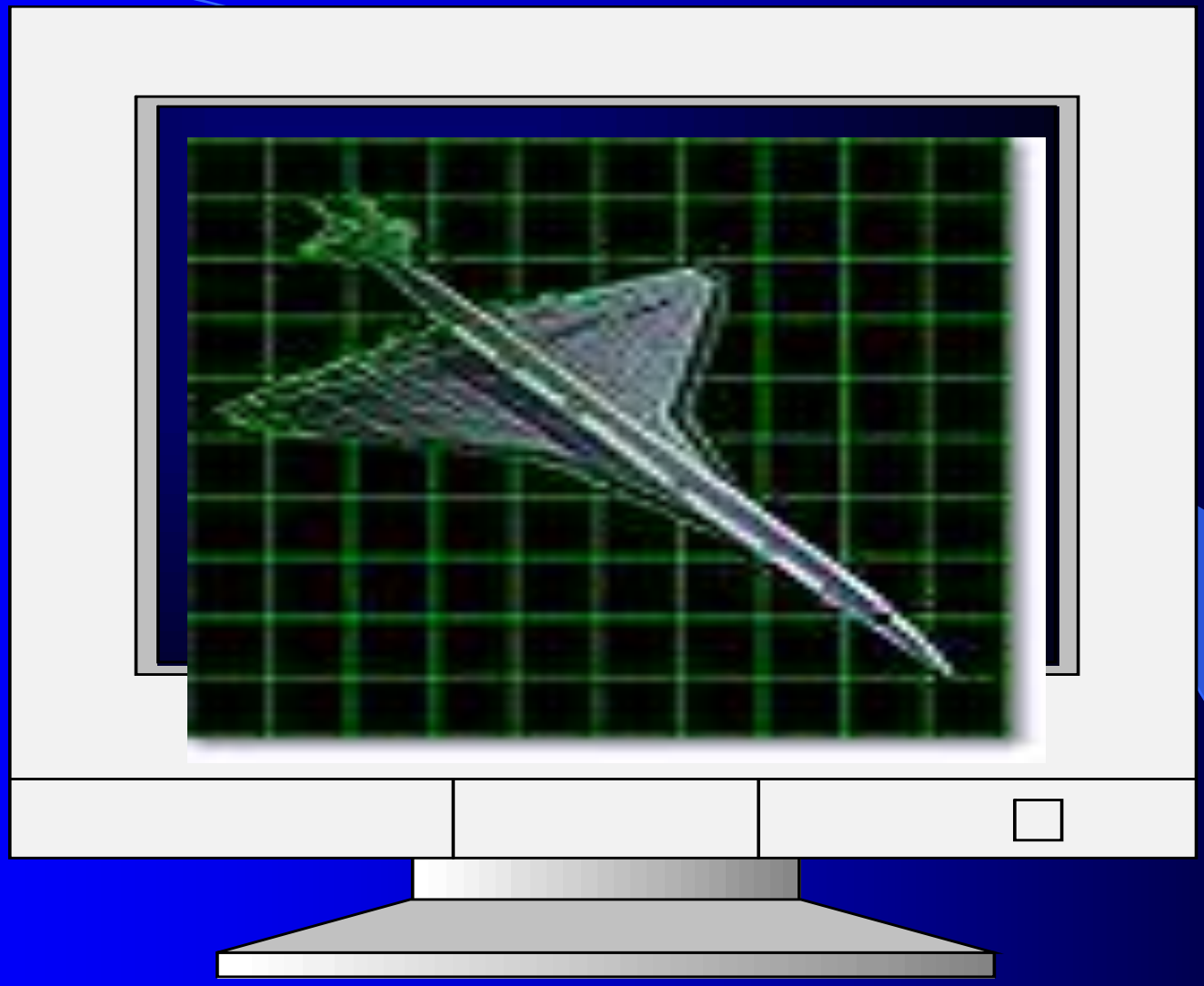
* First proven in comparative study by Cottrell for BMDO across hundreds of methods, including his own
- Lockheed presentation at Marshall Institute update on ballistic missile defense
- Widely published in many papers with variations in AIAA journals
- Invited plenary talk in China including Harbin
- Paper in Si et al, Handbook of Learning and Adaptive Dynamic Programming, IEEE/Wiley 1994. (Also: SNAC. Broom balancer)

McAir Process for Thermoplastic C-C Parts



Prediction Errors (HIC p.319)





1st Neural Flight Control (90's): "Cloning" best human pilot in slowed-down "game" led to Robust controller for National Aerospace Plane model much faster than previous best controller. The neural contractor AAC became for a time Lead company in US hypersonics effort.

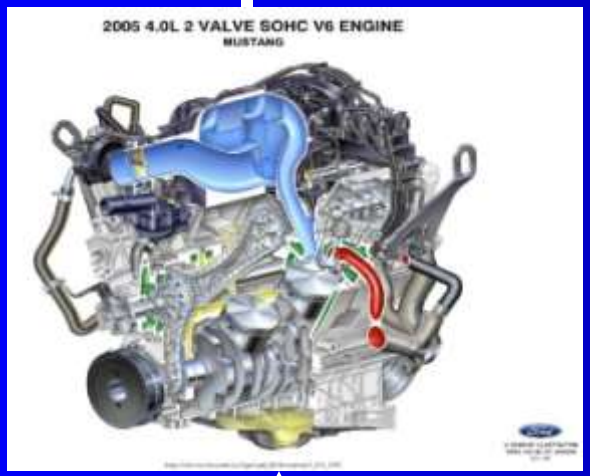
Venayagamoorthy/Wunsch/Harley ADP Turbogenerator Control

- Stabilized voltage & reactance under intense disturbance where neuroadaptive & usual methods failed
- Being implemented in full-scale experimental grid in South Africa
- Best paper award IJCNN99

Neural Network in Commercial Power Grid Hardware

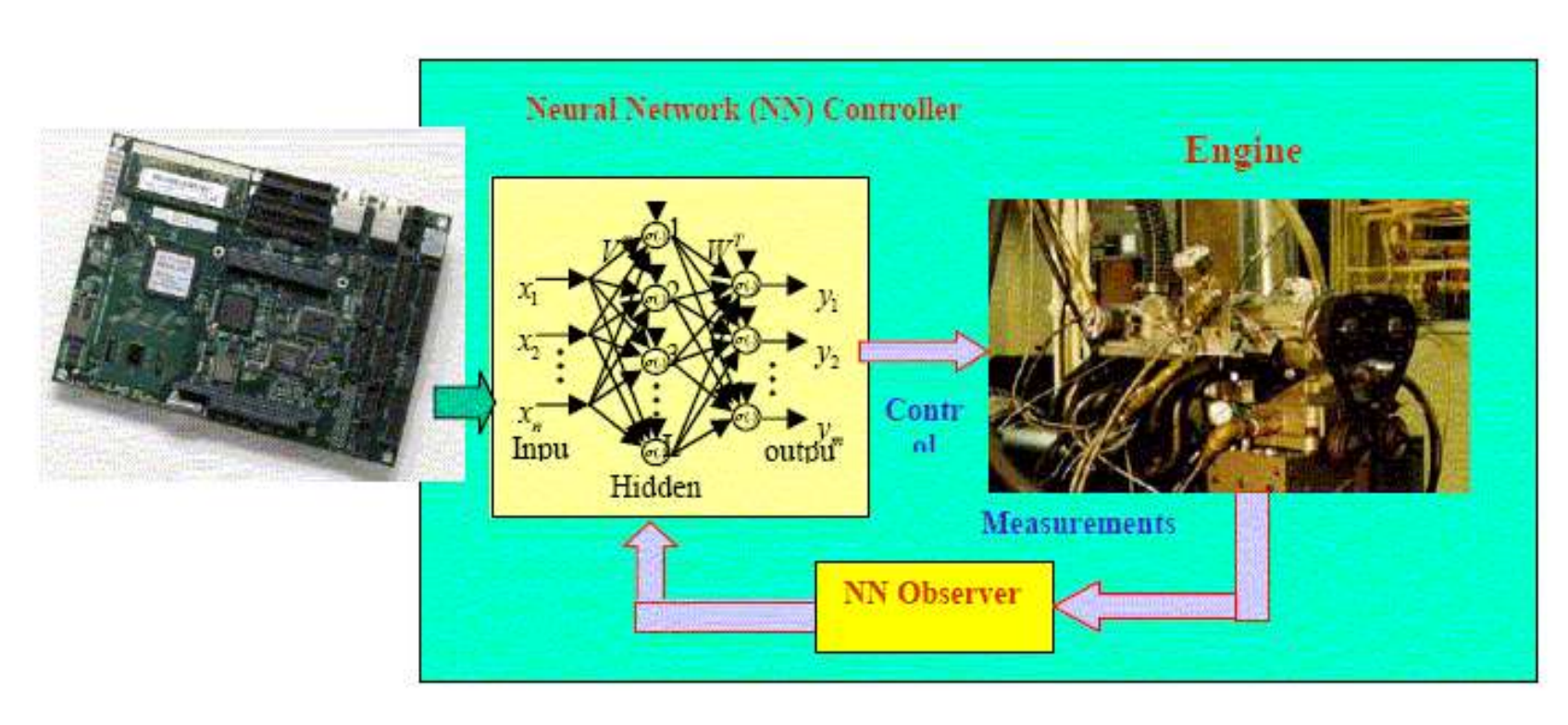
- First deployment of deployment of recurrent neural network in the field in a commercial electric power grid. (Improved prediction to allow unprecedented monitoring and control of harmonics.) Harley, Georgia Tech.

Winner of IJCNN07 Forecasting Contest: Ford 1998: “All Ford Cars Will Have TLRNs by 2001, for on-board Diagnostics”



- How can one neural network predict and diagnose all Ford engines, without retraining, even as they change over time? TLRN: adaptive prediction even without learning! ICNN05: “A neural network which can predict anything.”
- IJCNN07, Prokhorov: TLRN prediction and control can improve mpg of Prius hybrid by 15% “at zero cost”!
- Technology now at Siemens, Toyota, other big companies

ADP Controller Cuts NOx emissions from Diesel Engines by 98%



J. Sarangapani MUST NSF grant

IJCNN07: JS shows mpg up 7% in ordinary car engines with ADP

Prokhorov shows mpg up 15% in Prius hybrid with Neural MPC

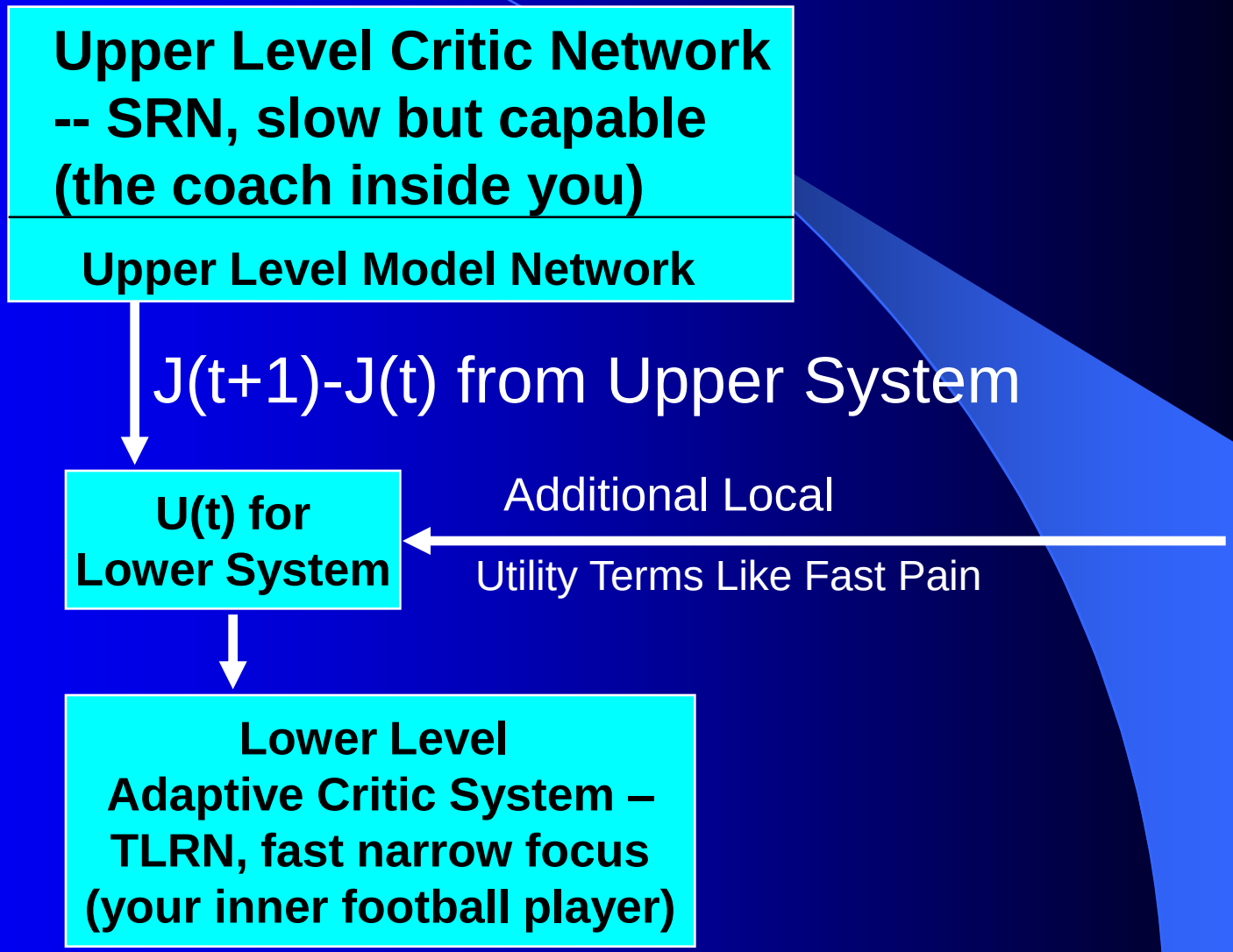
NASA Ames (Dryden) autolands Full MD11 with flaps locked up, safety breakthrough

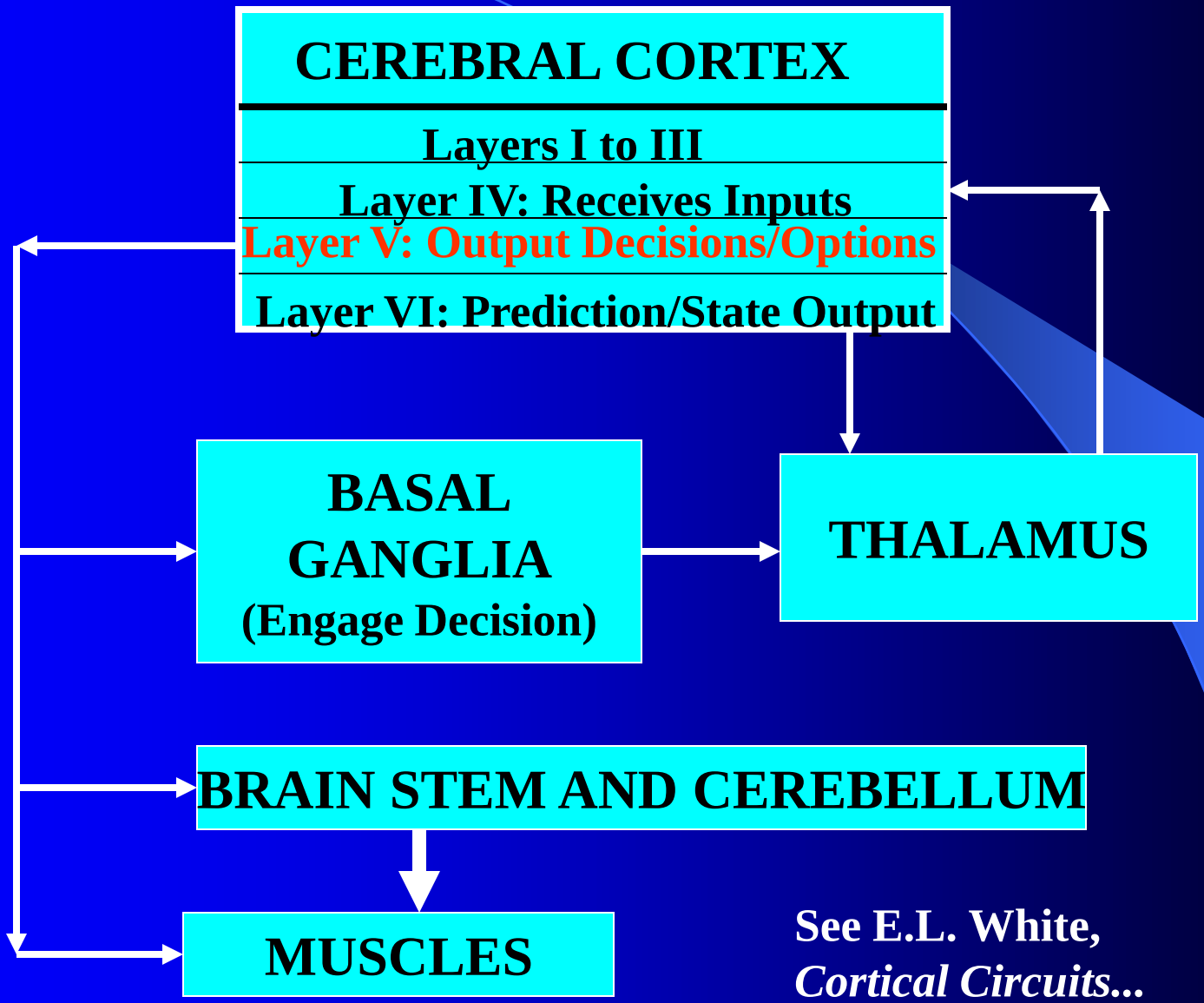
- Never deployed – like space re-entry technology
- Developer Jorgensen recently out of NASA

NSF Learning and Intelligent Systems (LIS) Calls for Work To Combine Filters (&China Using Multiple Sources of Info, Machine Learning

- Engineering: Will it work? Mathematics understandable, generic?
- Psychology: Connectionist cognitive science, animal learning, folk psychology
- Neuroscience: computational neuroscience
- AI: agents, games (backgammon, go), etc

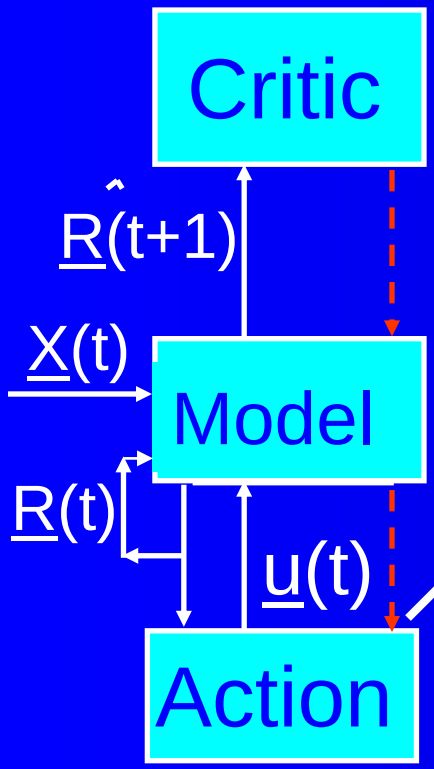
SMC 87: Model of Brain as Two Brains (Two Entire FLS Systems) in One





See E.L. White, *Cortical Circuits...*

RLADP From Vector to Mammal: see http:arxiv.org 2014 MLCI



0. Vector Intelligence – HDP, DHP, GDHP, etc.

Add new spatial complexity logic (ObjectNets +…, Suitable for CNNs)

1. First ever system which learned master class chess Fogel, Proc IEEE 2004

Add ability to make Decisions, plays (Modified Bellman eqs for Multiscale t.)

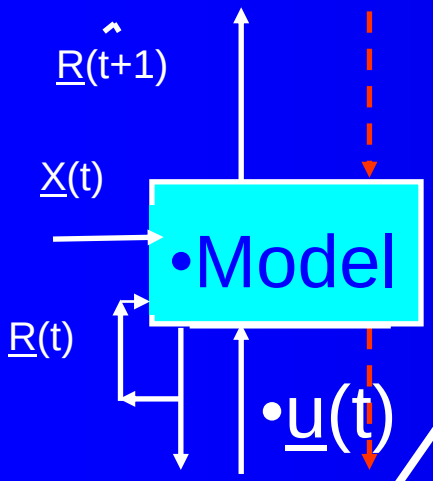
2. reptile

Add Creativity System (Cognitive map of space of possible decisions)

3. Mouse

Roadmap for Cognitive Prediction

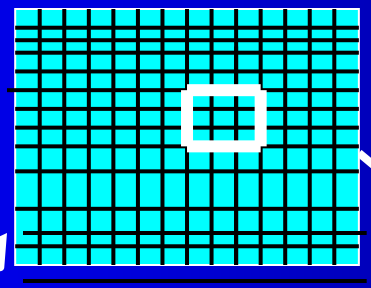
Reward direct simplicity



0. Vector Prediction (robustified SRN/TLRN)
HIC Chapter 10 on web.

Reward symmetry

1. AT&T winning ZIP code recognizer and new COPN work



Networks for inputs with more spatial complexity using symmetry – CSRN, ObjectNets, ….

Predicts What Will Happen Over Multiple Time Intervals Harmonized

2. reptile

Space-like cognitive map of the space of Possibilities, to support higher creativity

3. Mouse

To see how you could do better than even them, and break the world records again… or to see the research needs to fulfill gthis roadmap… see

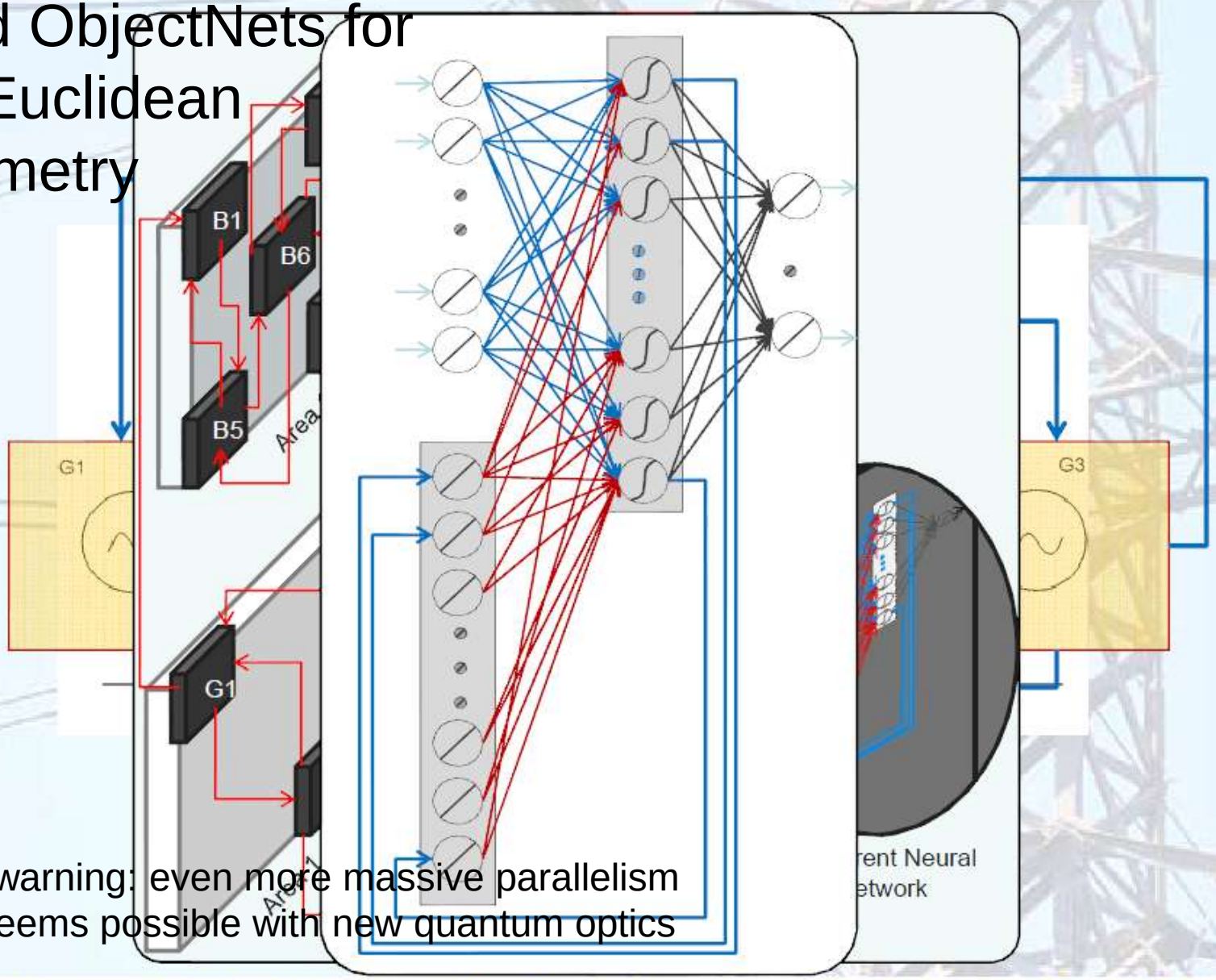
www.werbos.com/Erdos.pdf

David Fogel (Proc IEEE 2004): World's First System which LEARNED Master-Class Performance in Chess

- Evolutionary computing (EC) to train a game-player worked for tic-tac-toe, but not checkers
- EC to train a multilayer perceptron (MLP) to serve as a CRITIC (an ADP value function) was enough to beat checkers but not chess
- EC to train a feedforward Object Net as a Critic was enough to beat chess
- Prediction: A full (recurrent) ObjectNet Critic can get to master class in Go. Will Wunsch get there first?

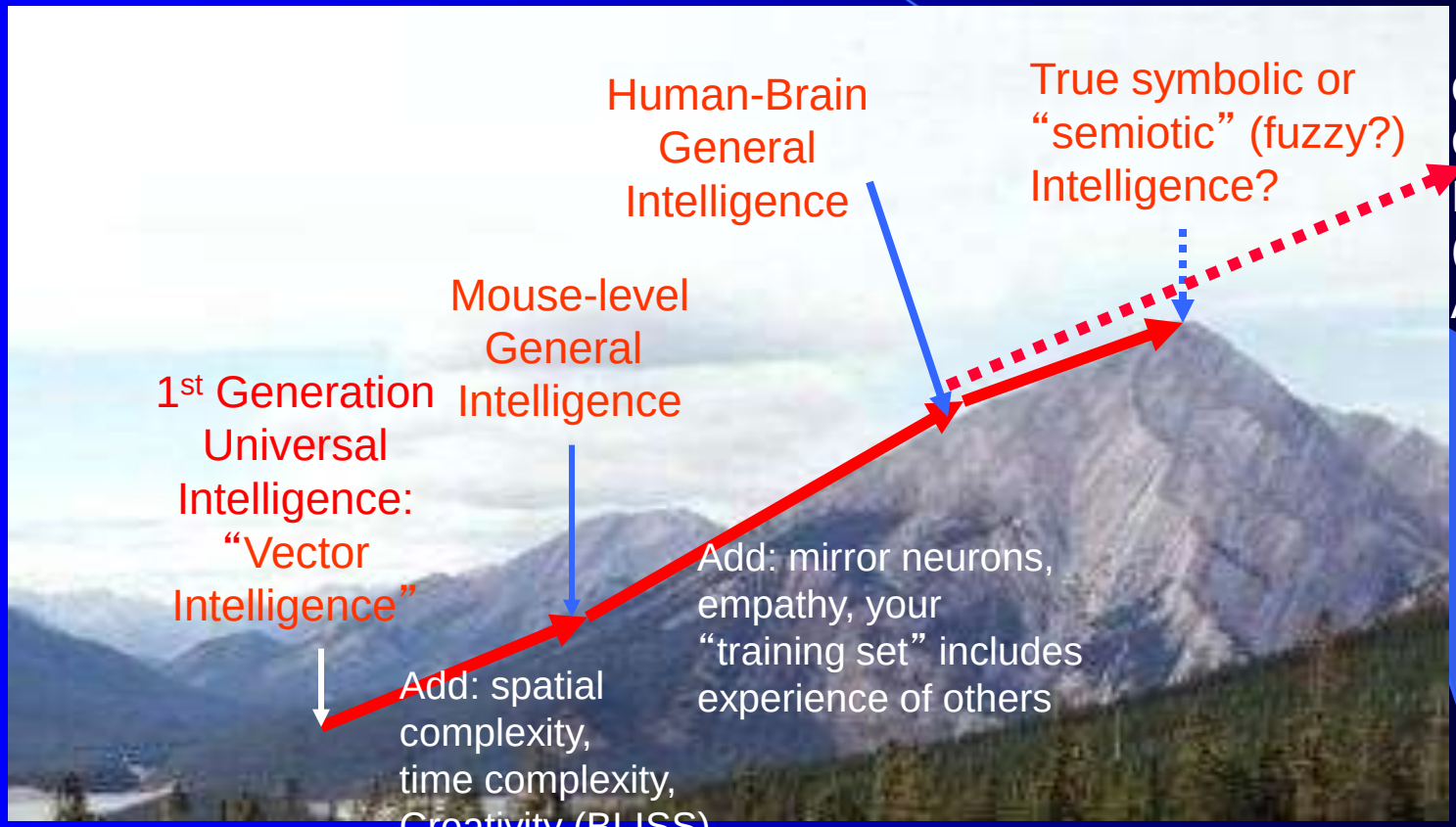
Cellular Neural Networks for WAMS

Need ObjectNets for NonEuclidean Symmetry



Early warning: even more massive parallelism now seems possible with new quantum optics

From Brain to Mind: What Can We Learn Of Use Beyond the Level of the Mouse Brain?



www.werbos.com/pi/Confucius_talk.pdf

And Neural Networks 2012 , arxiv 2015

"NSF is currently supporting research to develop a '4th generation intelligent grid' that would use intelligent system-wide optimization to allow up to 80% of electricity to come from renewable sources and 80% of cars to be pluggable electric vehicles (PEV) without compromising reliability , and at minimum cost to the Nation (Werbos 2011)." -- NSTC Smart Grid Policy June 2011

Werbos 2011: IEEE Computational Intelligence Magazine, August 2011

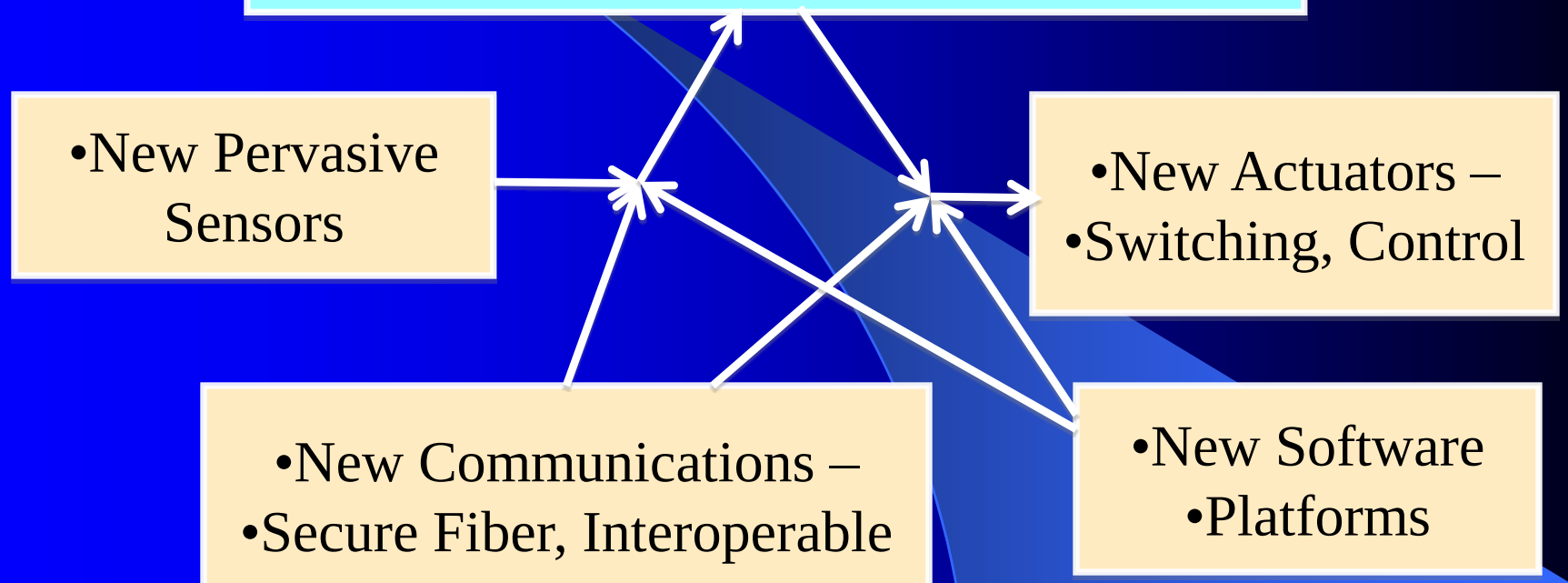
•New Adaptive Intelligent Algorithms For Anticipatory Optimization

•-- At User Sites

•-- At ISO/RTO

•-- Elsewhere

•-- Distributed but an Integrated System



2008: ENG-Wide Cognitive Optimization & Prediction under EFRI

Pr(A|B)=Pr(B|A)* Pr(A)/Pr(B)

COPN

Prediction

Memory

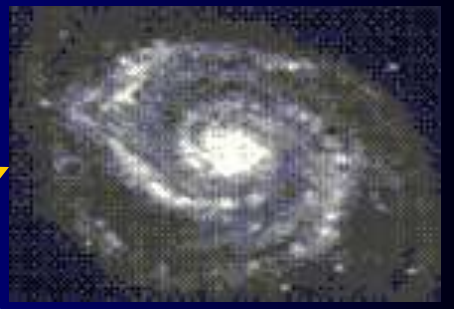
. . .

Clustering

Optimization

J(t)=Max<J(t+1)+U>

$$\frac{\partial^{+} z_n}{\partial z_i} = \frac{\partial z_n}{\partial z_i} + \sum_{j=i+1}^{n-1} \frac{\partial^{+} z_n}{\partial z_j} \frac{\partial z_j}{\partial z_i}$$



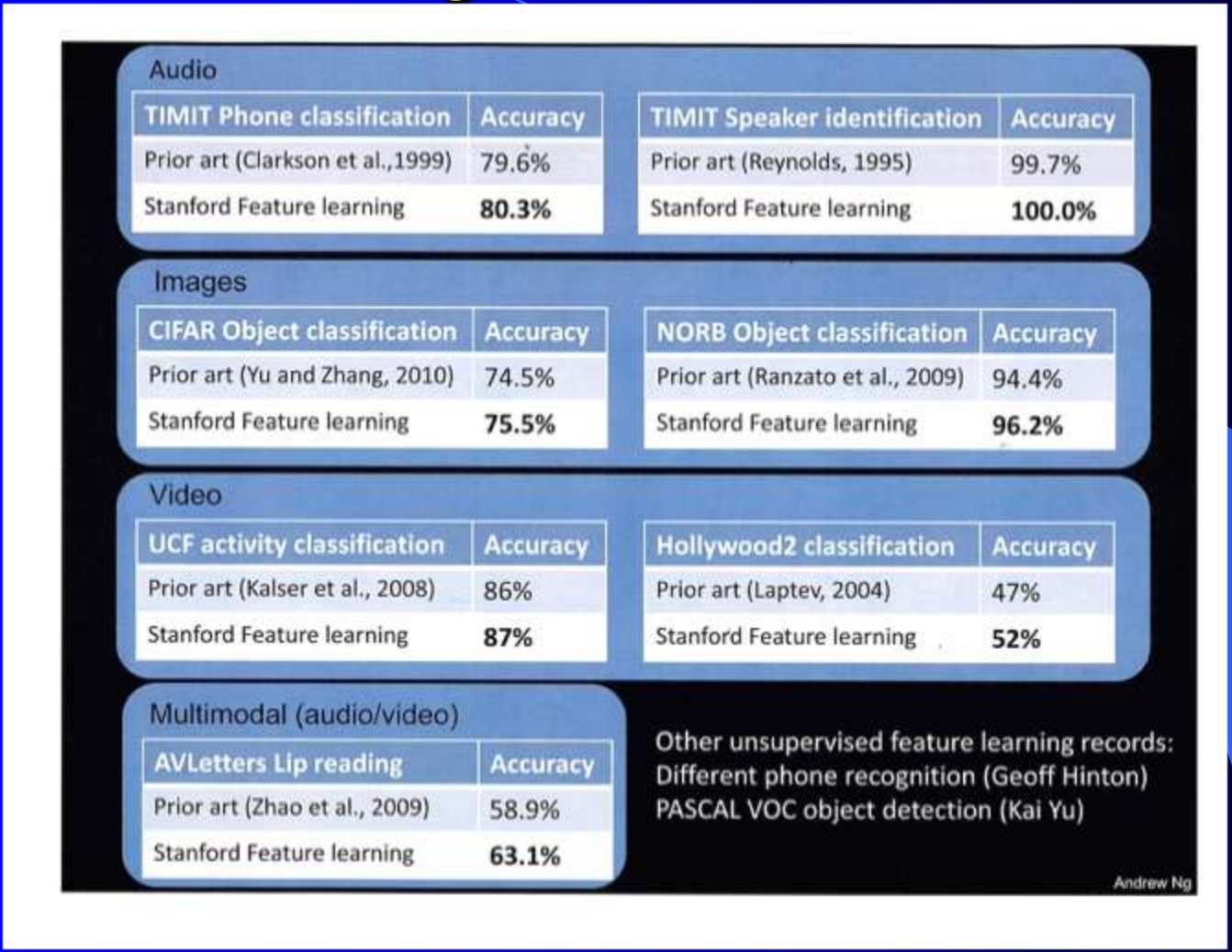
The Economic Crunch of 2008

- Finance spends a lot on prediction and optimal decision. But they had many failures in 2008. Today I have time for just one.
- The trigger of the collapse:
  - Big financial firms predicted low probability of big loss in packages of mortgages
  - Given M mortgages, i=1,…,M, estimate Pr(default-i) from "FICO scores"
  - Assume independent probabilities such that

    Pr(total default) = Pr(default-1) * Pr(default-2) *…*Pr(default-M)

But FICO does not give a probability! It shifted from neural nets to SVM, from scores based on a probability method to scores based on Vapnik thinking, when this became popular. Also, no cross-time analysis or external variable conditions reported.

New Performance Breakthroughs in Prediction/Recognition



New world records (under NSF COPN) using relatively simple neural networks with a symmetry addition…

Winter Soldier: Another Warning

Will IBM Watson Save us from misuse of real algorithms to serve an emerging cabal of a few? (Orson Scott Card, Empire)

Or is faith and wide use of artificial intelligence a worse threat than artifcial intelligence? Will we kill ourselves by stupidity?

Will control of brains by folks who do not understand them lead to really gross loss of freedom, as in this guy (or in "Clone Armies") even if nonsurgical stimulation?