The New Mathematics of Mind: A Path to Unify Modern Science and Traditional Spiritual Life

Paul J. Werbos, paul.werbos@gmail.com

Personal views, not representing any organization.

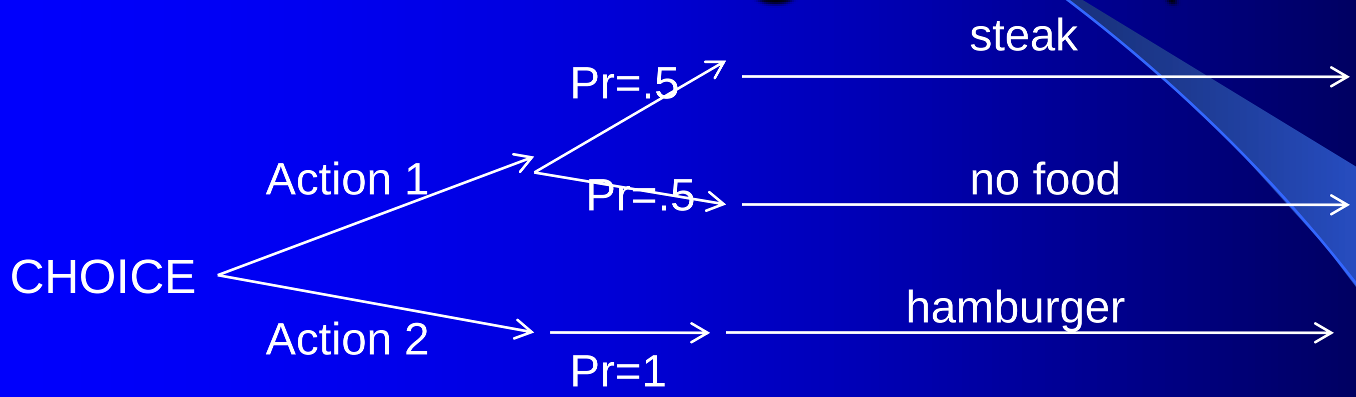
- Basic principles – questions, Von Neumann, Confucius
- 1st generation universal intelligence, vector intelligence
  - History and Basic Concepts
  - Success stories – manufacturing, vehicles, robots..
- Roadmap and new success from vector to mouse
- Beyond the individual brain: collective intelligence and quantum effects

www.werbos.com

Key Questions from Riken Brain Institute (Modified):

- How can we understand brain?
- How can we build brain?
- How can we cultivate or use our brains as much as possible?
- BUT: they really mean MIND or INTELLIGENCE. It requires a new mathematical roadmap or understanding to answer and unify these questions.

Von Neumann's Cardinal Utility: The Starting Point (1962)



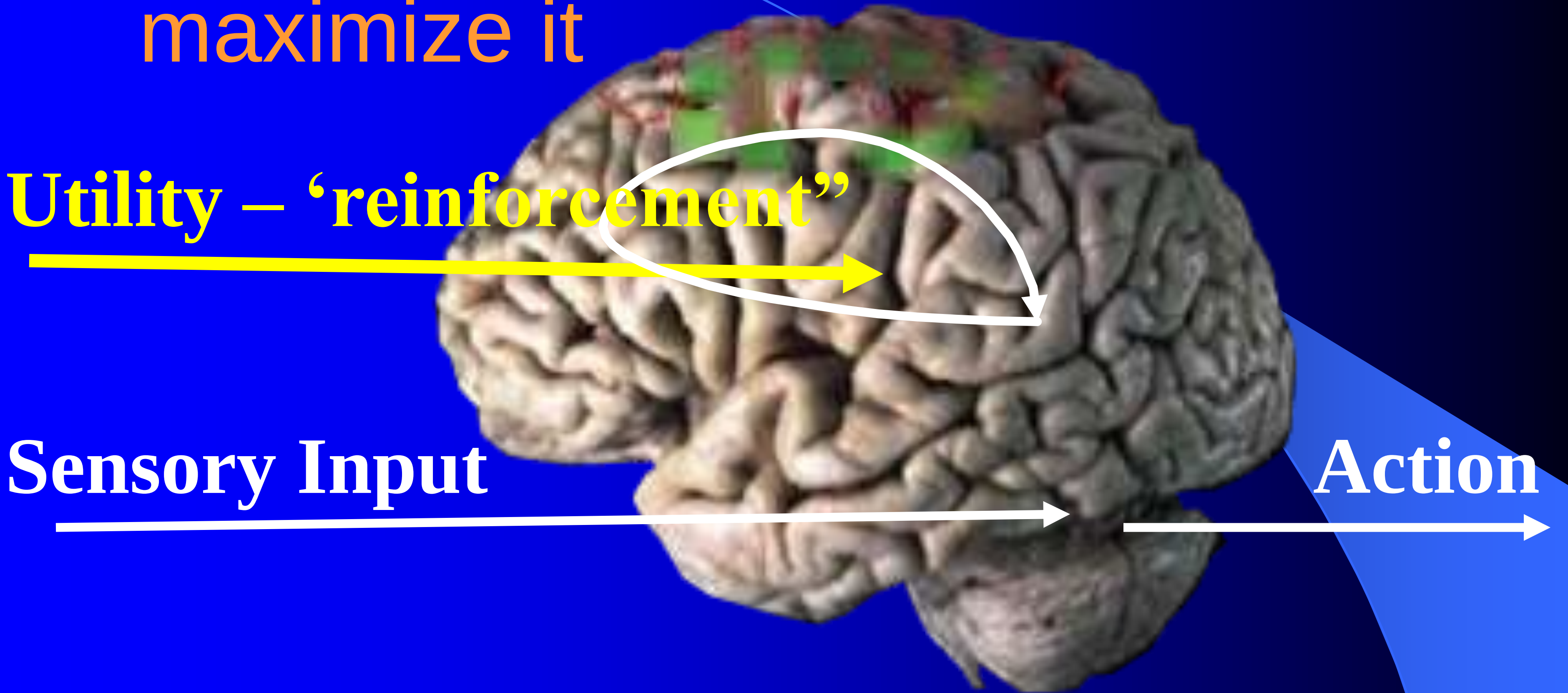
Here, just 3 possible results

- Which animals learn to maximize U(results)?
- **HOW**: general ability to maximize U in future?
- What should **MY** utility function be –goals in life?
- How can I maximize it? How can I do my best?

Personal Breakthrough at Age 15 (when studying logic from Alonzo Church)

- “What **should** I do?” is an ill-posed question in logic (and in inductive learning) – Bertrand Russell
- “What **would I** do if I were wise?” – i.e., what combination of ideas would truly satisfy me about this. Well-posed logically, because it refers to “I.”
- Confucian friend Chen: “Who are you?” Integrity: know and be true to self.

Two challenges – what is U; how to maximize it



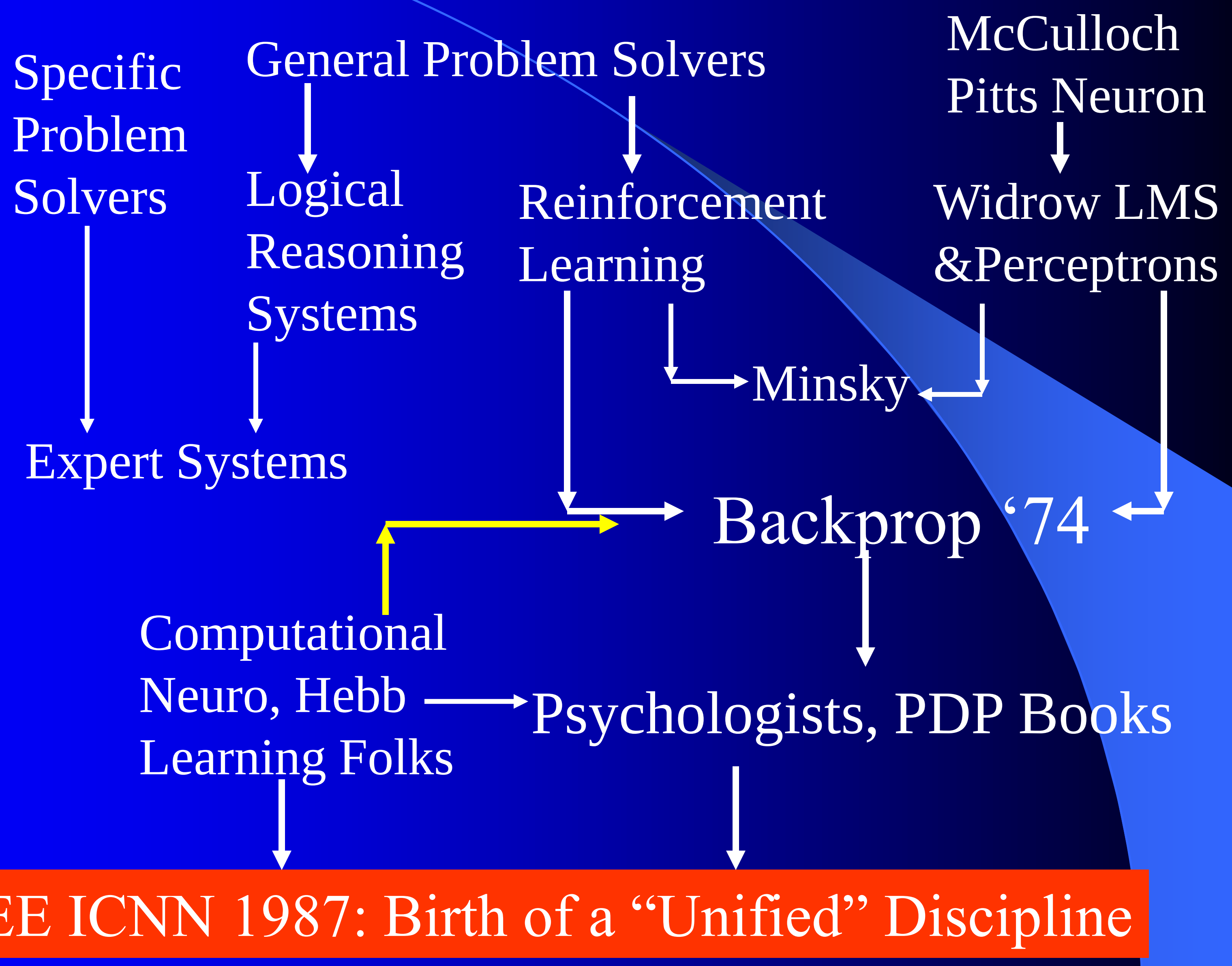
Brain As Whole System Is an Intelligent Controller

-- Mouse maximize probability of survival among other things

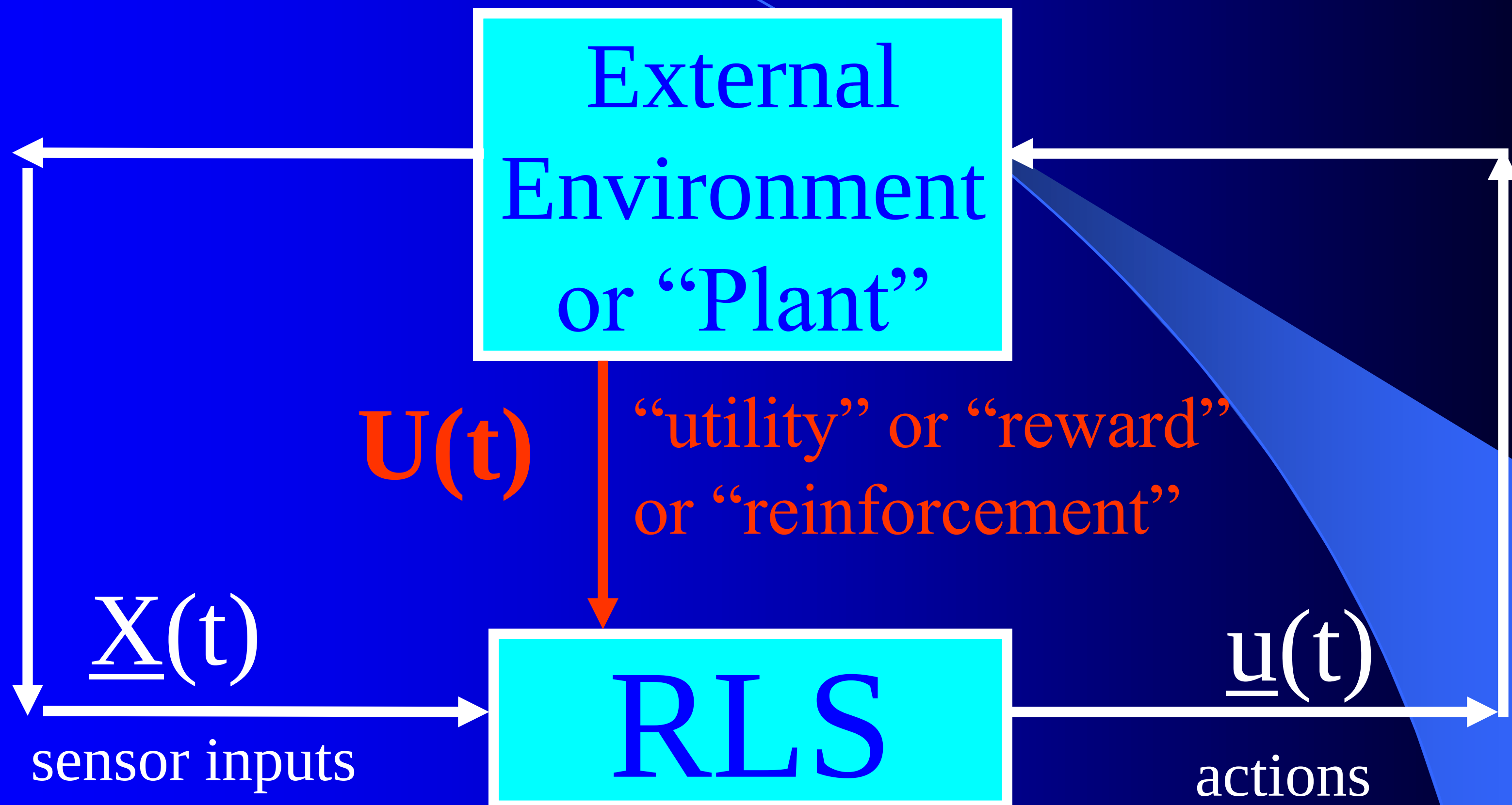
-- Lots of animal behavior research

-- Lots of recent motor control research (UCSD…)

Where Did ANNs Come From?



Reinforcement Learning Systems (RLS)



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

Maximizing utility over time

Model of reality

Utility function U

Dynamic programming

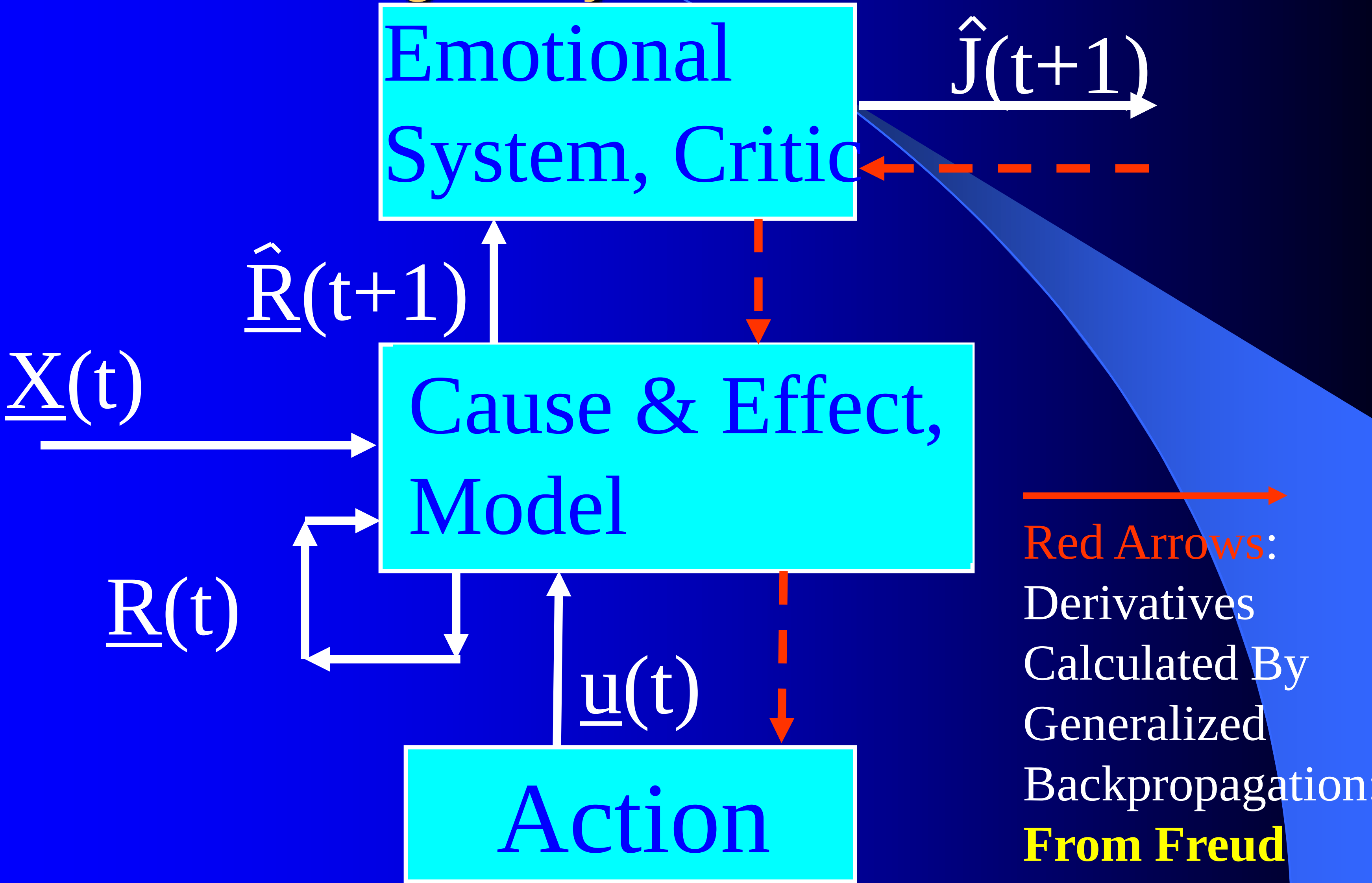
$$J(\mathbf{x}(t)) = \max_{\mathbf{u}(t)} \left\langle U(\mathbf{x}(t),\mathbf{u}(t)) + J(\mathbf{x}(t+1)) \right\rangle / (1+r)$$

Secondary, or strategic utility function J
Now often called value function V but..

Why It Requires Artificial Neural Networks (ANNs)

- For optimal performance in the general nonlinear case (nonlinear control strategies, state estimators, predictors, etc…), we need to adaptively estimate nonlinear functions. Thus we must use universal nonlinear function approximators.
- Barron (Yale) proved basic ANNs (MLP) much better than Taylor series, RBF, etc., to approximate smooth functions of many inputs. Similar theorems for approximating dynamic systems, etc., especially with more advanced, more powerful, MLP-like ANNs.
- ANNs more "chip-friendly" by definition: Mosaix chips, CNN here today, for embedded apps, massive thruput

1971-2: Harvard thesis proposal: first universal intelligent system– note all the vectors



Examples of J and U (and ∇U, ∇ J)

DOMAIN	INTRINSIC UTILITY U	STRATEGIC UTILITY J
Chess	Win/Lose	Queen = 9 Points...
Business	Cash Flow, Profit	Net Present Value
Human Mind	Pleasure/Pain	Hope,Fear
Behavioral Psychology	Primary Reinforcement	Secondary Reinforcement
Artificial Intelligence	Utility Function	Position Evaluator
Economics (Derivatives)	Value of Product to You Now	Market Price or Shadow Price λ

Neural Networks and Science

Universal Neural Network Intelligent Systems

Technology: Algorithms & Architectures

Models to Predict Brain Circuits & Behavior

Models to Understand First Person Subjective Experience

Third Person Objective Science (Kuhn)

- ----

Learn from shared replicable experiments

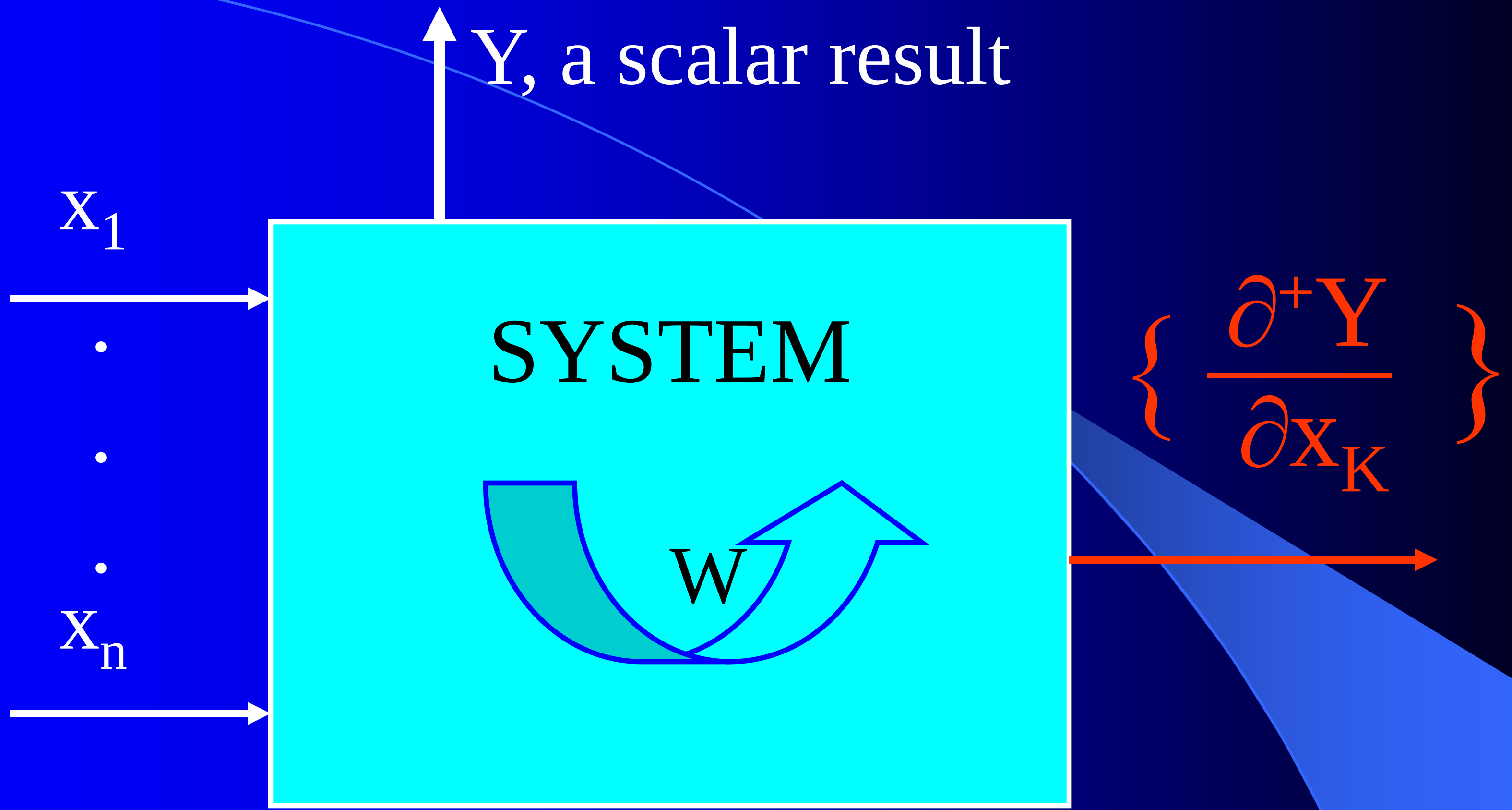
First Person Science – Learn from all experience

"Third Person" Utility Functions Are Also Important In Practical Life

- As sane individuals, we have different utility functions U(x) in our brains. To do our best, we must work with other people with different feelings. This is an extension of Von Neumann's concept of n-player games of mixed cooperation and competition, modified by T. Schelling's observations in Strategy of Conflict.
- In such games, it is crucial to seek "Pareto optimal" solutions. These are sometimes called "win-win" solutions. For example:
  - Many times we can agree on a "third person" universal kind of utility function as the basis for collective action, such as value added plus reliability to govern an electric power grid, or rational market design.
  - Since decisions are based on J(x), not U(x), we can converge to agreement as we have more foresight – and also as our sensitivity to U(x(t)) expands to better reflect terms involving other people and natural inputs from collective intelligence.

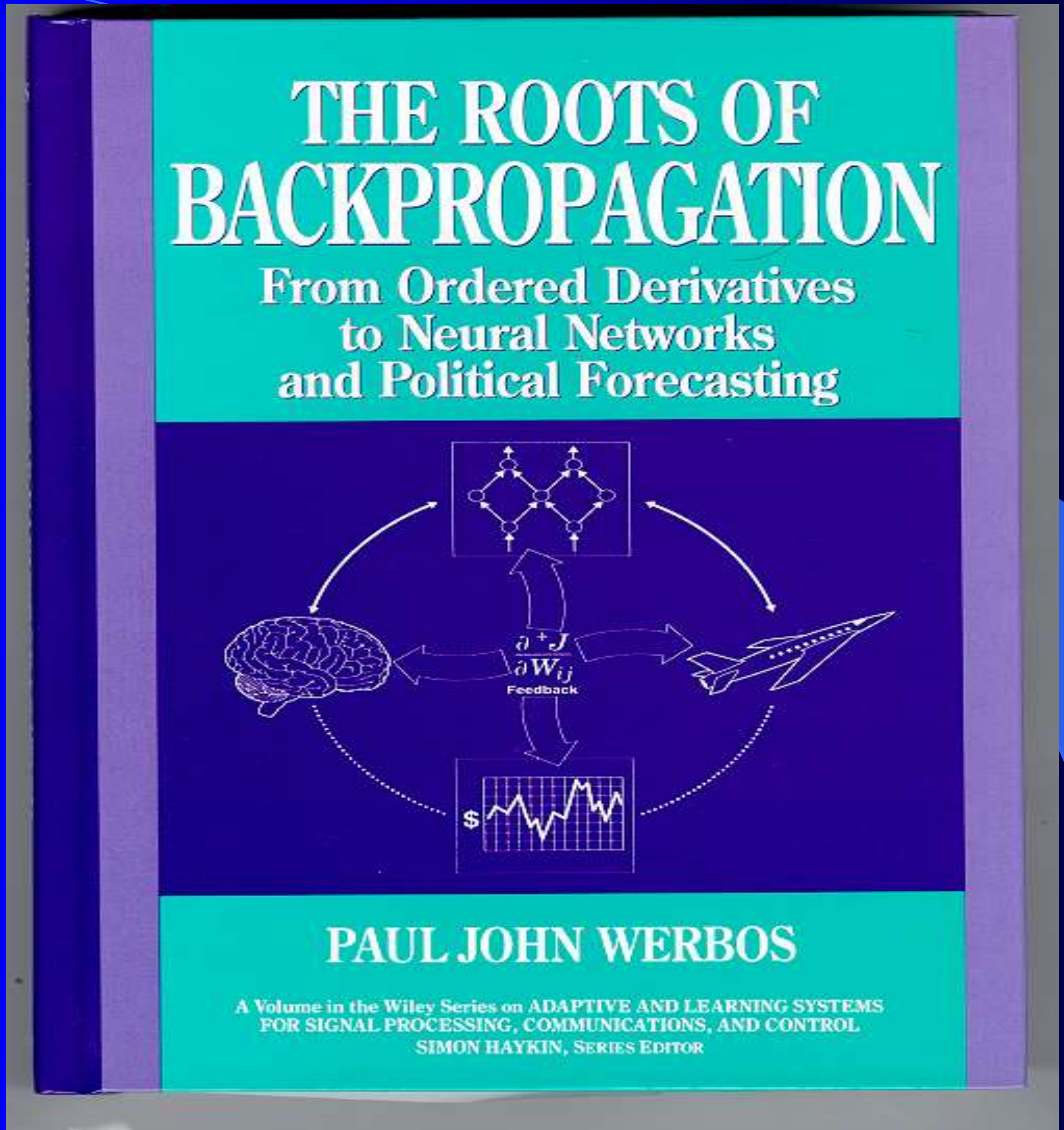
Harvard Committee Response

- 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

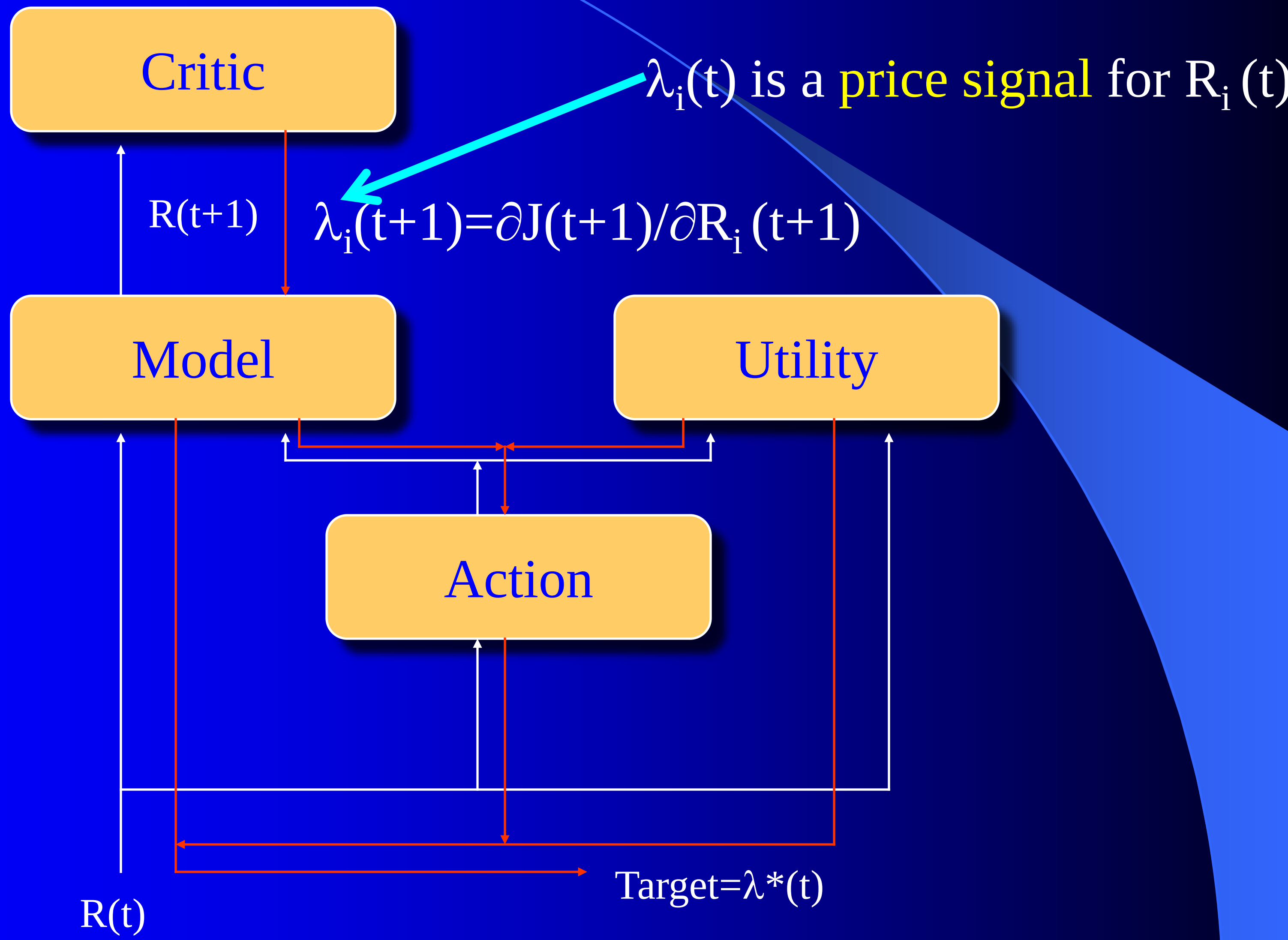


(Inputs x_k may actually come from many times)

Backwards Differentiation: But what kinds of SYSTEM can we handle? See details in AD2004 Proceedings, & my web page. Prof. Iri the leader in Japan.

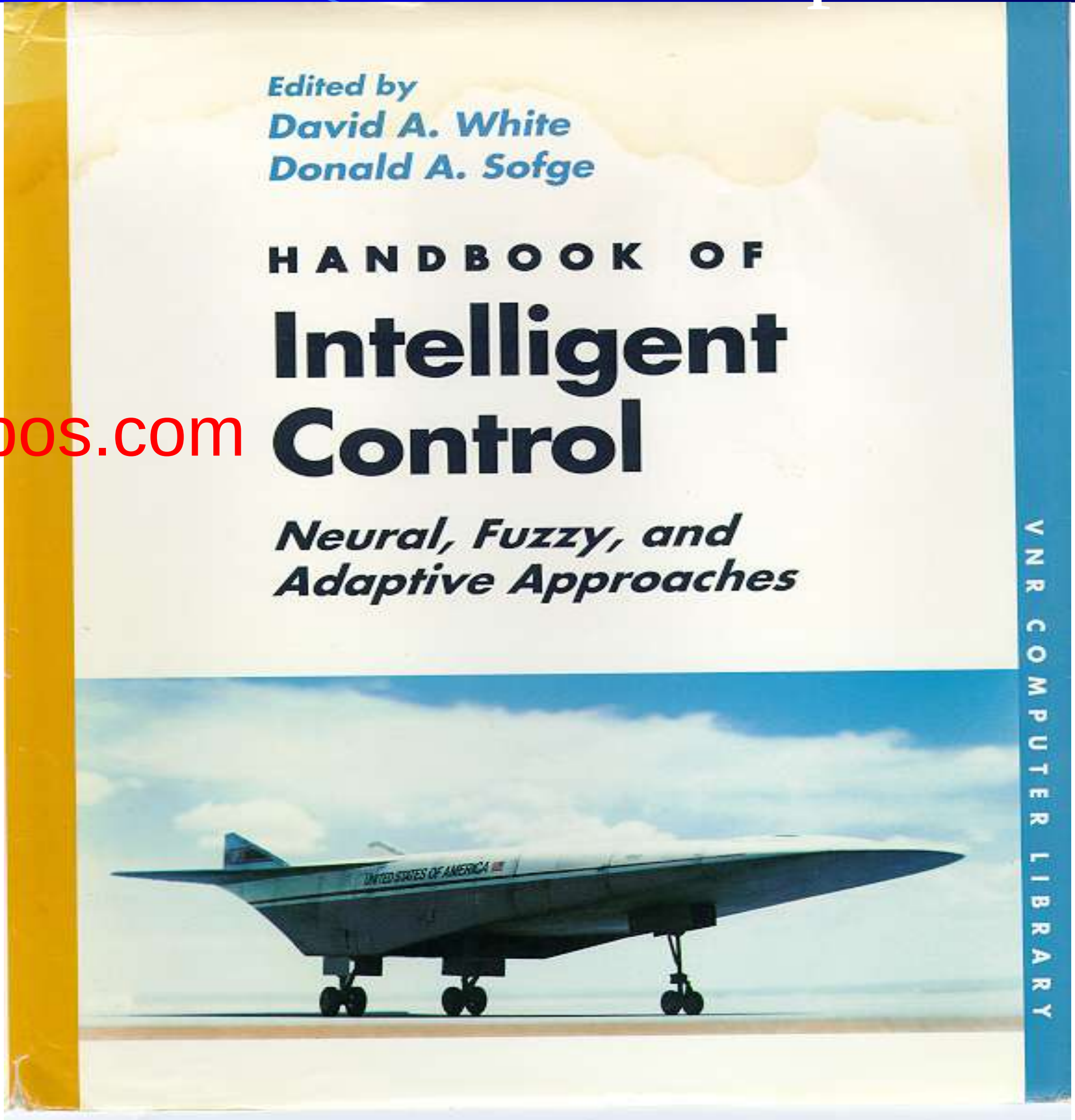

THE ROOTS OF
BACKPROPAGATION
From Ordered Derivatives
to Neural Networks
and Political Forecasting
$\frac{\partial^+ J}{\partial W_{ij}}$
Feedback
$
PAUL JOHN WERBOS
A Volume in the Wiley Series on ADAPTIVE AND LEARNING SYSTEMS
FOR SIGNAL PROCESSING, COMMUNICATIONS, AND CONTROL
SIMON HAYKIN, SERIES EDITOR

Dual Heuristic Programming (DHP)



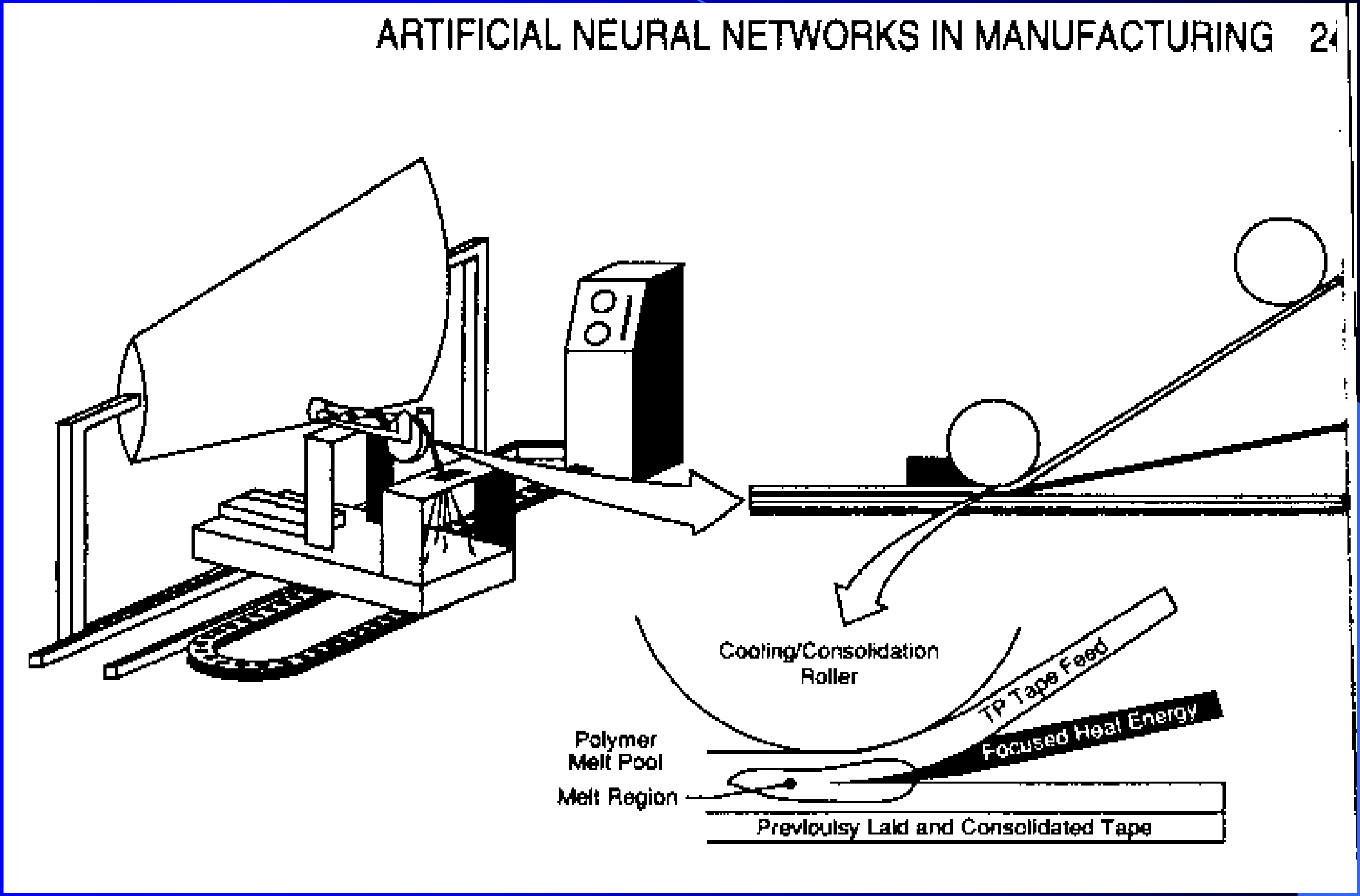
NSF/McAir Workshop 1990

See chapters posted at www.werbos.com



White and Sofge eds, Van Nostrand, 1992

McAir Process for Thermoplastic C-C Parts

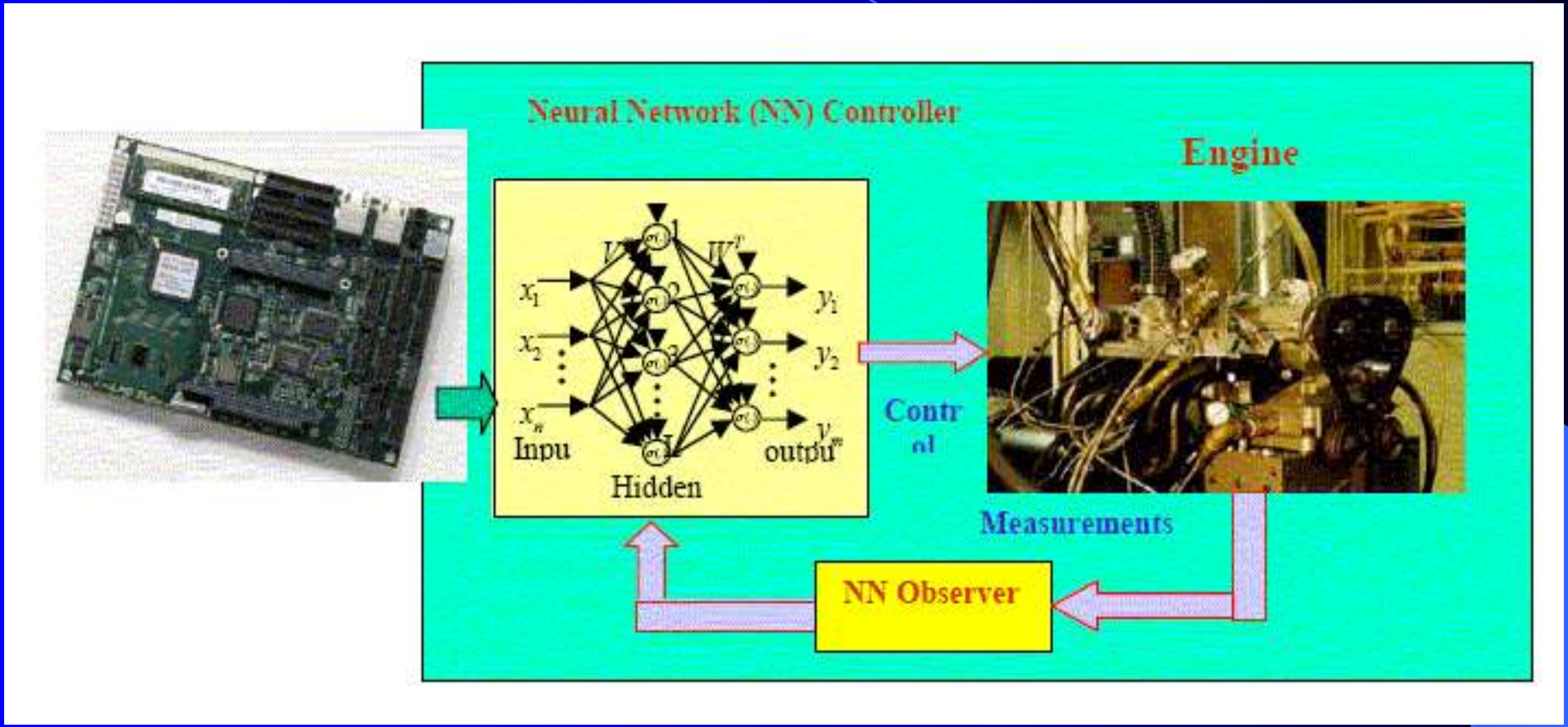


S.N. Balakrishnan: Using DHP, Reduced Error in Hit to Kill Missile Interception more than order of magnitude vs. all previous methods

* 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

ADP Controller Cuts NOx emissions from Diesel Engines by 98%



J. Sarangapani UMR 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

Human mentors robot and then robot improves skill

Schaal, Atkeson
NSF ITR project

Learning allowed robot to quickly learn to imitate human, and then improve agile movements (tennis strokes). Learning many agile movements quickly will be crucial to enabling >80% robotic assembly in space.

Wunsch/venayagamoorthy/Harley ADP Turbogenerator Control

- Stabilized voltage & reactance under intense disturbance where neuroadaptive & usual methods failed
- Validated on full-scale experimental grid in South Africa, Mexico project
- Best paper award IJCNN99
- Neural net control now in 20% of US coal plants – Lefebvre, related approach

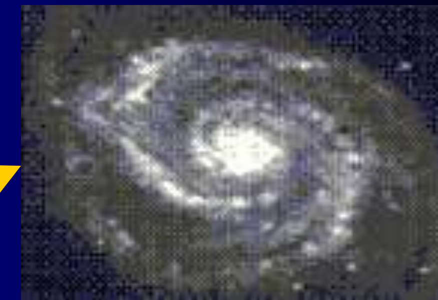
ADP Vector Intelligence Since 1992

- A Number of New Breakthrough Applications and lessons learned from them
- New Stability Results – Lewis, Werbos, Others for ADP; Suykens and others for "neural MPC"
  - But only Werbos arxiv 1998 addresses stochastic case
  - Two NSF workshops (2002, 2006) and new Handbook (2002): www.eas.asu.edu/~nsfadp. ADPRL conferences.
- Warren Powell book on ADP addresses stochastic case, big follow-up in OR, but not up on universal approximation aspect
- User-friendly general software & manuals still needed
- Comprehensive PhD program still needed

“Cognitive Optimization and Prediction”, search on “COPN” at www.nsf.gov – and international follow-on?

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

Prediction



Memory

. . .

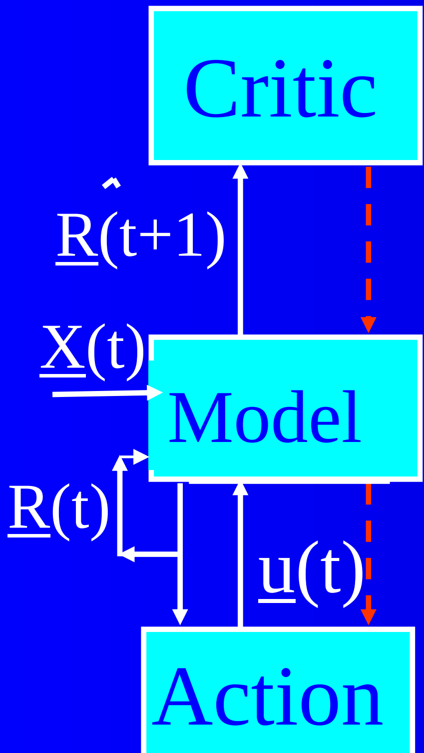
Clustering

Optimization

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

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

4^{th} gen model: From Vector to Mammal (Neural Networks 2009)



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

2. reptile

Add Creativity System

3. Mouse

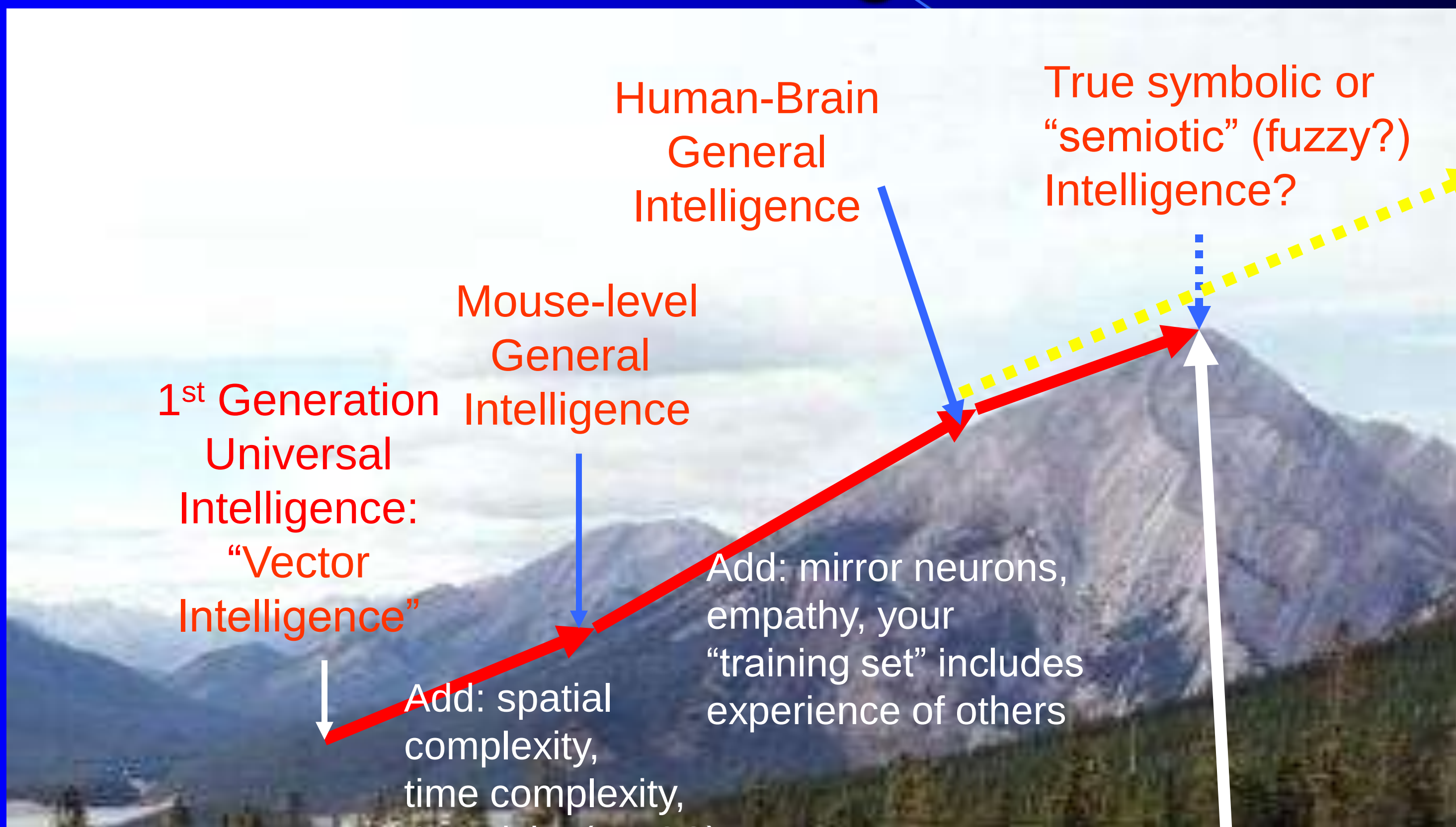
How can we go from dinosaurs to creative mice in our larger world?

4 Stages: use our inborn creativity; achieve sanity; connect with soul; train all

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?

What is the Fastest and Most Promising Path to Build and Understand the Highest Level of Intelligence??



Straight Up the Cliff
to "Human-Level" Intelligence?

New View of Language – Neural Networks 2009 & IJCNN Beijing 1992

- Mouse learns from current experience and from database of his memory of his experience
- “monkey see, monkey do”: with mirror neurons, monkey observes experience of other monkeys and people he sees, and learns from expanded memory.
- Humans expand memory more by shared experience – like Bushman dance after hunt, or ‘word movies”
- Logical propositions, logical reasoning and grammar are all new (~3000 years), learned and artificial

Scientific Basis of Confucian Integrity

- Symbolic reasoning – logic, mathematics and logical planning for the future – are very powerful, but are not inborn; they must be learned and be correct.
- If the utility function and axioms we use do not fit with the whole brain (our whole feelings), either our life is ruled by subsymbolic mind (we are as weak minded as a mouse) and hypocrisy, or conflict causes nervous breakdown and weakness.
- This is just Confucius and Freud versus Mo Tzu!
- In true integrity, we are **sensitive to deepest feelings,**

which drive ideas about U both in logic and subsymbolic intelligence, and both are totally integrated. Like Freud's idea of true sanity.

But What of Mao's 4 Olds?

- Mao-Tse Tung rightly criticized the terrible effects which conservative, unscientific Confucianism (as taught at Yuelu Academy) had on China in his time. I can imagine him saying "zhengqi is like a pure, strong beautiful woman, but if you marry her, beware of the criminals in her family."
- But Confucius is not alone in having some bad followers. Jesus even denounced ahead of time those who would come in his name and distort his fundamental principles.
- The science of slave owners and feudalists was not true science; there is better, true science. Likewise, there can be better, truer and more scientific zhengqi, if we work at it.

Beyond the Mundane Brain

- Human history has many stories of life which do not fit the model of individual brain as the only intelligence. But they seem impossible in a simple model.
- Personal experience forced me to become open-minded in March 1967: I remembered and quoted a speech of Mao Tse Tung the day before it was given! (See Greeley 1969: 70% of PhDs have had personal experience they do not discuss.)
- Two intense new directions:
  - Scour literature from “schools” all over the earth for ways to expand my first person database, because more and richer experience is needed to explain all this. Also use neural network model to guide this.
  - Re-examine basic physics and what is really possible, with the smallest possible change in physics from what we already know.

A Few Conclusions

- What is possible in physics: see my paper in International Journal of Theoretical Physics (IJTP 2008), at http://arXiv.org. Backwards time causal effects not only possible, but necessary to explain paradoxes in laboratory physics. Also, dark matter and energy and other things prove we do not know all the important fields of the universe.
- Collective intelligence effects (ultimately based on emergent properties of such fields) are possible and necessary to explain the full range of esoteric and spiritual experience. Similar to "Gaia hypothesis", Tao Te Ching
- Collective intelligence is still governed by mathematics, and is held together by some kinds of "qi", i.e. backpropagation.

Beyond the Individual Human Brain – A Few Quick Thoughts Based on What I Have Seen

- Quantum level – search on Werbos at http://arxiv.org
- Multimodular – just more symmetry, extends spatial complexity for a kind of collective intelligence
- $\partial J / \partial R_i$, a backpropagation signal, represents "how much an increase in R_i makes to happier. It represents the value of R_i to you. It fits Freud's idea of "psychic energy." Backpropagation of J dervatives represents a kind of flow of emotional energy in the brain. It drives the learning of all that we do.
- Collective intelligence requires similar flows of derivative information between people and connecting to the larger world. Is "ki" or 'tao" (or "tama") really just a flow of the same kind of derivative signals in our larger Mind, driving the larger course of our culture and society and world? Is it governed by the same mathematics, the chain rule for ordered derivatives?

The Force be With You!!!

Neural Networks as a Path to Self-Awareness: IJCNN 2011

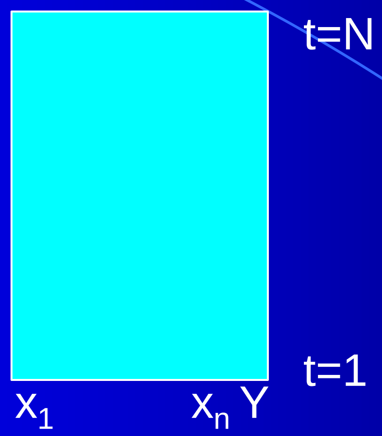
- Bootstrapping of natural image prediction to create awareness of more variables
- Mirror neurons: we have “third party dreams.” If we remember and analyze them before we even leave bed, we can see more, develop more empathy for others, and build conscious connections. (These slides at 3AM!)
- Apply integrity and awareness to all levels of our experience.
- Many others; world dialogue for human survival

The End – but some extras

Example of Freud and Syncretism

- A Freudian story:
  - Nazi hurts child, a traumatic memory
  - For years, he is terrified when anyone in black shirt appears (precedent based prediction/expectation) – the kernel-based “id” is at work!
  - Later he learns about Nazis in subjective model of world (f), “ego”
  - After that learning, if he relives that memory (trains on memory), f error on the memory is low; memory loses power to cause irrational bias
- Key corollaries:
  - False hope from memory is as dangerous as false fear
  - We still need id when exploring new realms we can’t yet reliably predict

Model-Based Versus Precedent-Based: Which Is Better?



- Model-based: Pick W to fit Y=g(x,W) across examples t. Given a new x(T), predict Y(T) as g(x(T),W). Exploit Barron's Theorem that smooth (low C) functions f are well approximated by simple MLP neural nets – though not by Taylor series. Also add penalty function to error measure, ala empirical Bayes, Phatak – min e+f(W).
- Precedent-Based: Find t whose x(t) is closest to x(T). Predict Y(T) as Y(t). Kernel is similar, weighted sum of near values.
- Best is optimal hybrid, needed by brain. "Syncretism" – chapter 3 of HIC…. Next 2 slides

"Syncretism" Design

Basic Idea: $$\hat{Y}(t) = \tilde{f}(x(t)) + \sum_{\tau} K(x(t) - x(\tau))(Y(\tau) - \tilde{f}(x(\tau))$$

Practical Implementation/Approximation:

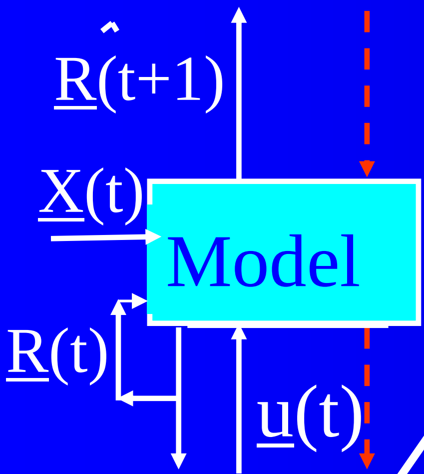
-- Associative Memory of Prototype $x(\tau), Y(\tau), Y(\tau) - f^*(x(\tau))$

-- Update $Y(\tau) - f^*(x(\tau))$ on occasion as f^* is changed

In other words: Keep training f* to match examples or prototypes in memory, especially high-error examples.
Predict Y(t) by f* plus adjustment for errors of f in nearby memory.
Closest so far: Principe kernel applied to model residuals;
Atkeson's memory-based learning.
Exactly fits Freud's description of ego versus id in neurodynamics.

From Vector to Mammal

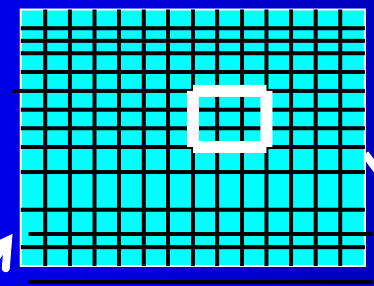
Reward direct simplicity



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

Reward symmetry

1. AT&T winning ZIP code recognizer and then CLION



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

2. reptile

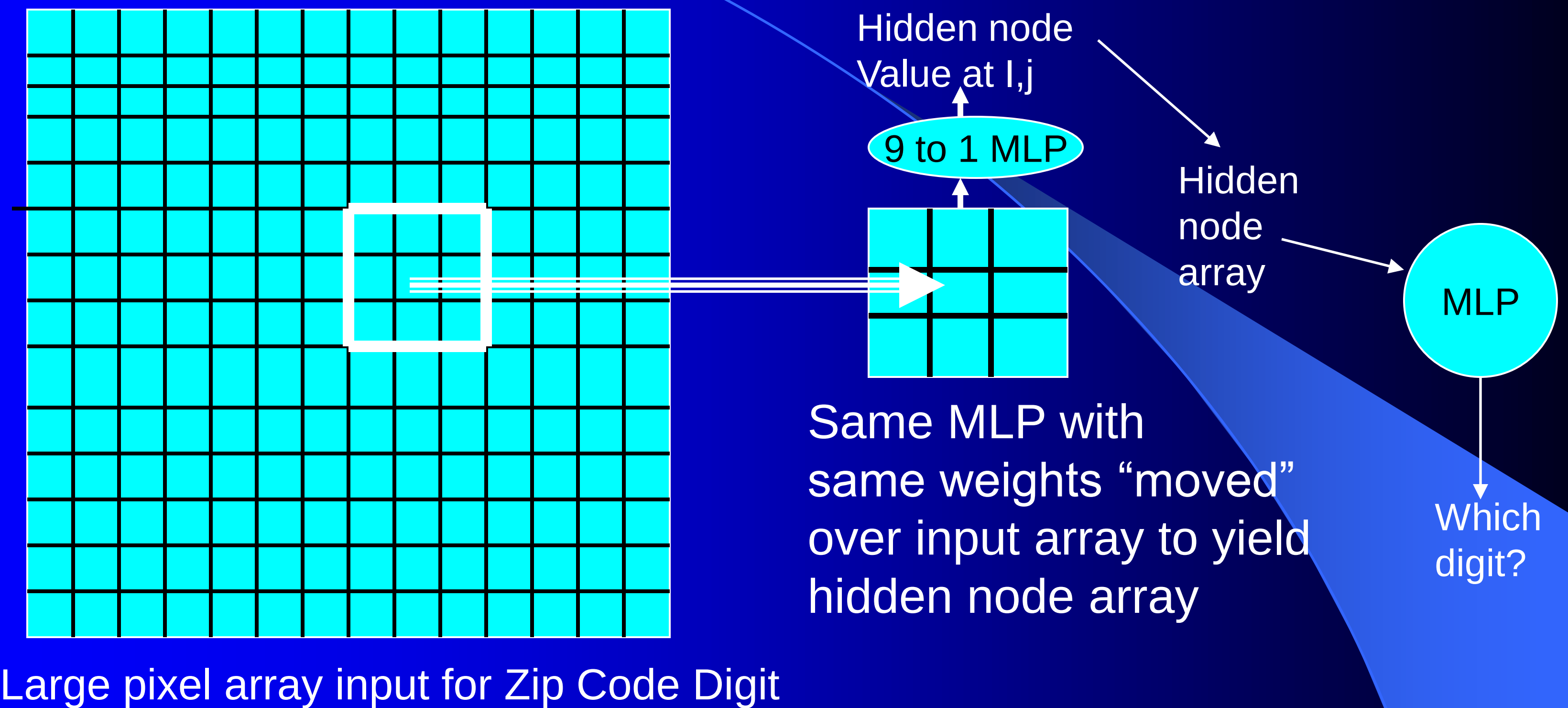
Predicts What **Will Happen** Over Multiple Time Intervals Harmonized

3. Mouse

Predicts and "Imagines The **Possibilities**" (Stochastic y=**f**(**X**,**e**). HIC Chapter 13 on web.)

M. E. Bitterman (Scientific American 1969; Science later):
Mouse learns to predict better in stochastic pattern recognition tasks, where turtles just slowly "go crazy." Cut mouse cortex, get turtle behavior. But reptile probably has stochastic capability, just not well-integrated.

Moving Window Net: Clue Re Complexity



- 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!

Cellular SRN: The Recurrent (SRN) Generalization of “Conformal MLP”

GENERALIZED MAZE PROBLEM

Jhat(ix,iy) for all 0<ix,iy<N+1

(an N by N array)

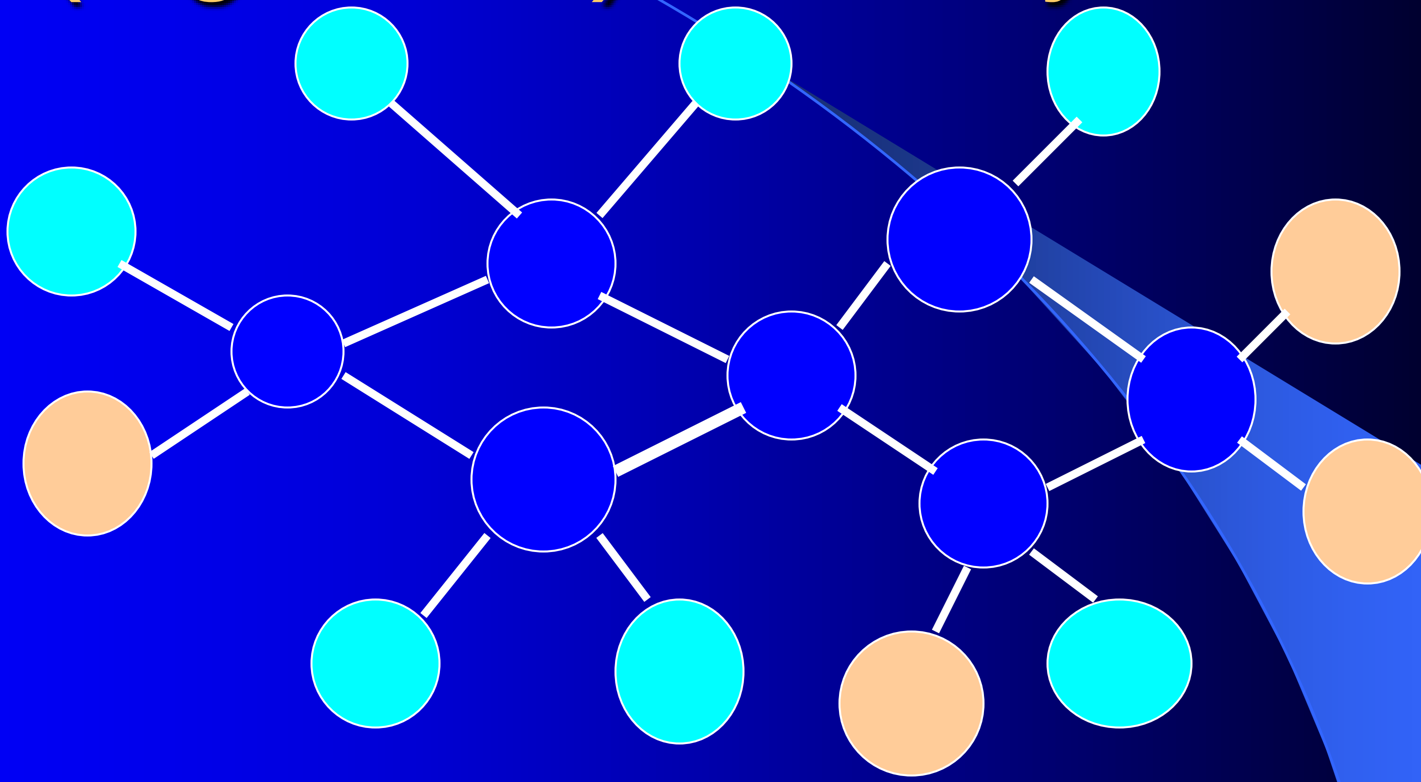
NETWORK

Maze Description

- Obstacle (ix,iy) all ix,iy
- Goal (ix,iy) all ix,iy

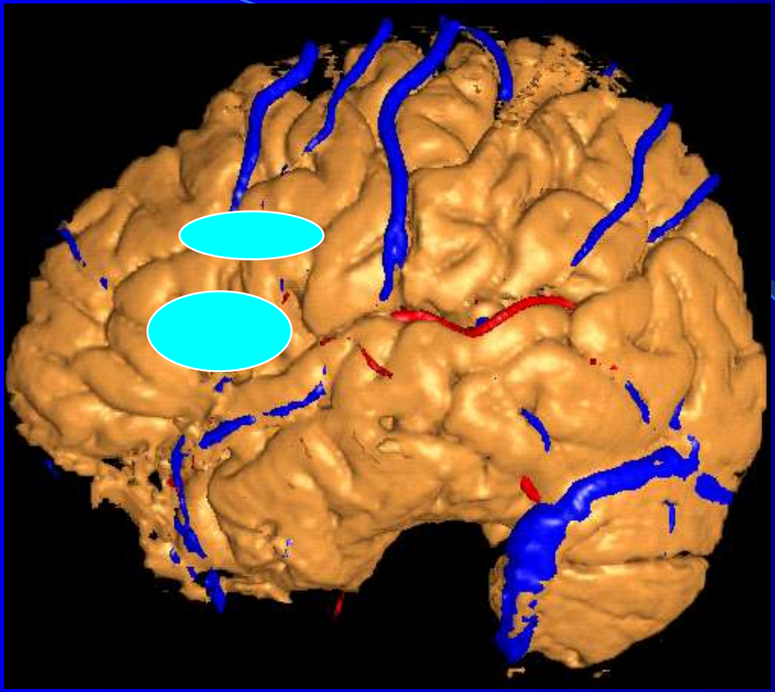
Rapid learning algorithm by Kozma, Ilin, Werbos: IEEE Transactions on Neural networks, June 2008

Spatial Symmetry in the General Case (e.g. Grids): the Object Net



- 4 General Object Types (busbar, wire, G, L)
- Net should allow arbitrary number of the 4 objects
- How design ANN to input and output FIELDS -- variables like the SET of values for current ACROSS all objects?
- Great preliminary success (Fogel’s Master Class Chess player; U. Mo. Power)
- But how learn the objects and the symmetry transformations???? (Brain and images!!)

An Interesting Reminder from Neuroscience



Petrides (IJCNN06) shows that dorsolateral (DL) and orbitofrontal (OF) prefrontal cortex – our “highest” brain centers – answer two basic questions:

OF: Where did I leave my car this time in the parking lot? (space?)

DL: What was I trying to do anyway? (time?)