Intelligent Artificiality

and an Economics of Mental Behavior

Mihnea Moldoveanu
University of Toronto

Brook Lecture, Institute for Cognitive Science
Ottawa
April 16, 2013

What I Am Not (A Partial List, Part I)

- I am not an economist (though I work alongside a few);
- I am not a neuroscientist, of either wet or dry kinds (though I will work with many of both, soon);
- I am not a psychologist (though I used to work with one, and even tried to impersonate one for 18 months);
- I am not 'an AI guy (or, gal)' (though my post docs/RA's come from a computer science and artificial intelligence lab);

What I Am Not (A Partial List, Part II)

- I am not a neuro-economist (I do not understand what that means);
- I am not a neuropsychologist (they don't understand what that means);
- I am not an empiricist (but, who is, really?);
- I am not a theorist (see 'I am not an empiricist');
- I am not an epistemologist or 'impartial observer of scientific practice' (an incoherent concept).

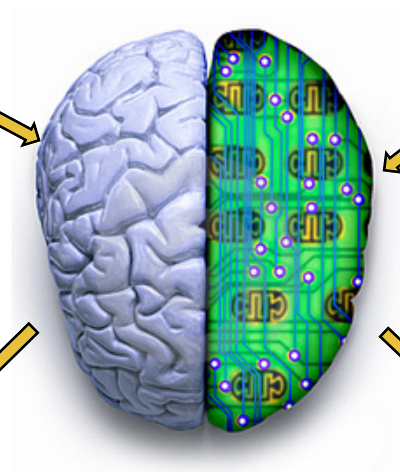
A Gap(ing Hole) in the Core of 'Rational Choice' Models

- Choice-theoretic conditions on 'rational choice' (antisymmetry, acyclicity, completeness, identity) 'guarantee' existence of objective function economic agents are said to maximize *in virtue of* choosing.
  - How are we to interpret **maximization (optimization)**? As a real process whose temporal dynamics *refer* to something?
    - If, so, what is it 'running on'?
      - Brains?
      - Researchers' desktops? Laptops? iPads?
      - 'Turing Machines'? (i.e. an imaginary process running on an imaginary device?)

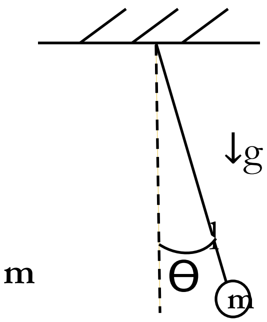
"The Predictive Apparatus Is Faulty"

Predict how this Creature will choose from among N options:

INPUTS:
Past choices among similar options
Revealed preference model
Rationality conditions



Predict behavior of this device:



INPUTS:
Newton' Laws

$$Fg = G\frac{m_1 m_2}{d^2}$$

Measured values of g, l, ϴ, m

OUTPUTS:
Prediction of choice/behavior

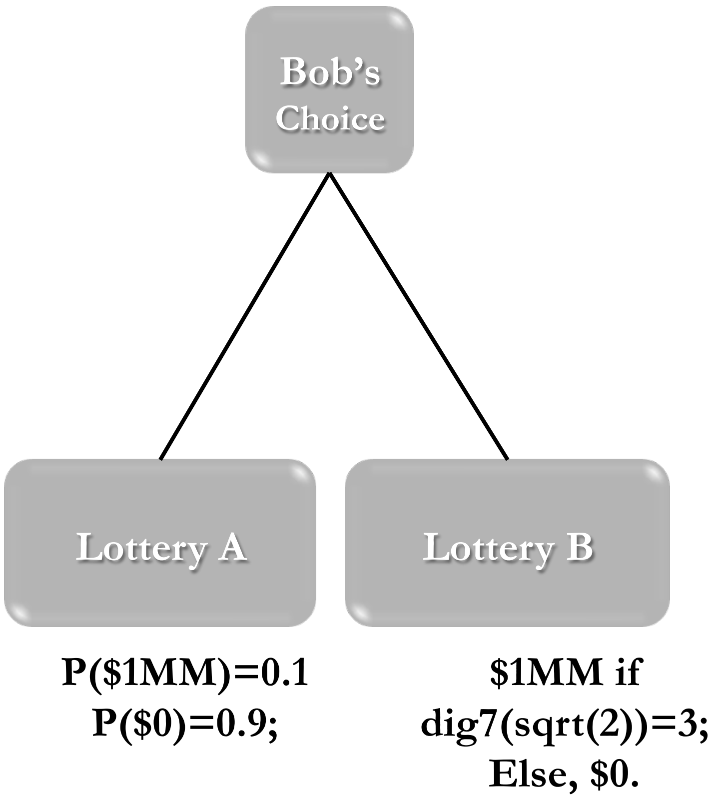
OUTPUTS:
Prediction of movement trajectory, transient and static

(Not) A Trick Question
(Illuminative of the question: 'How does optimization happen?'

- Suppose Bob must choose between two lotteries:

  - Lottery A pays *$1MM* with probability *0.1* and *$0* with probability *0.9*.
  - Lottery B pays $1MM if the 7th digit in the decimal expansion of *sqrt(2)* is an 3 and $0 otherwise.

    - **No calculator, SmartPhone or computer;**
    - **Needs to choose in 2 min.**



What If We Know Bob Knows **This**…?

- Depends on whether or not Bob sees the problem as one solvable by the algorithm;
- Depends on whether or not Bob can **correctly** perform required operations **quickly enough** to generate answer in under **2 minutes**.
- Depends on whether or not **Bob thinks** he can correctly perform the operations quickly enough to generate the answer in under 2 minutes.

Problem: Given x such that $x^2 = 2$, find x
[NEWTON'S METHOD]

➢Step 1: Form $f(x) = x^2 - 2$

➢Step 2: Compute $f^1(x) = 2x$

➢Step 3: Make first guess at x: $x_0 = 1$

➢Step 4: (Repeat as necessary) $X_{k+1} = x_k - \frac{f(x_k)}{f^1(x_k)}$

e.g. $x_1 = 1 - \frac{(1-2)}{2} = 1.5$

$x_2 = 1.416667$

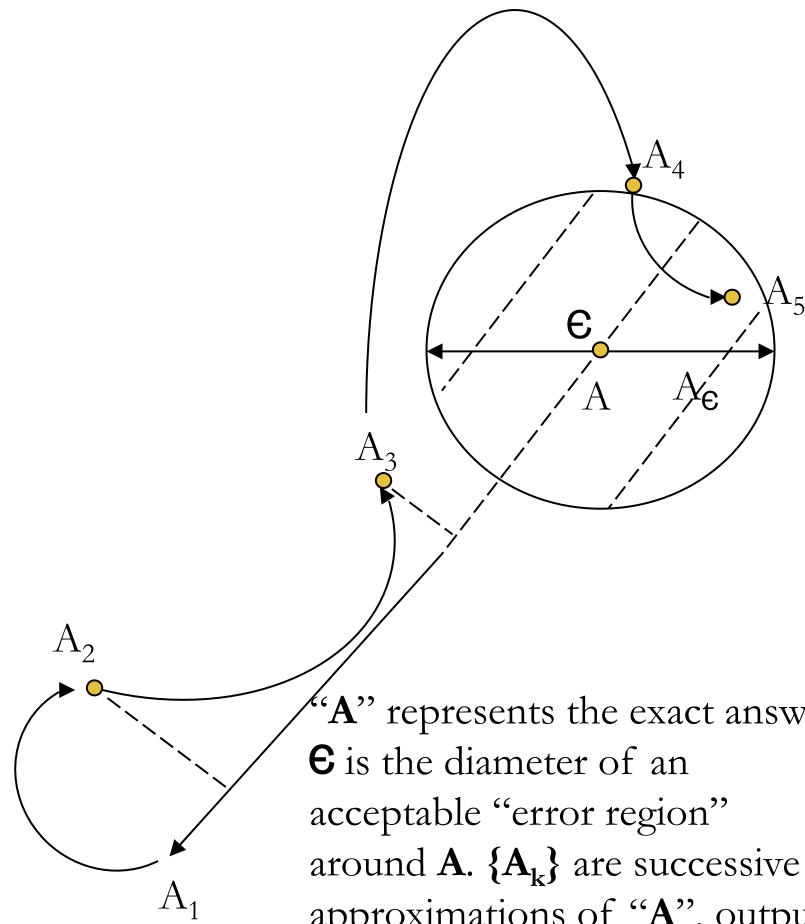
:

:

Calculator says $x = 1.4142135$. *(requires 5 steps)*

The Computational Process Model Matters to Whether We Ascribe 'Rationality' to Bob

- Each calculation generates new information (2 bits)…
- … that reduces Bob's uncertainty regarding the true value of the answer…
- …on account of the fact that it actually reduces the instantaneous search space of the problem he is trying too solve.



"**A**" represents the exact answer, **ϵ** is the diameter of an acceptable "error region" around **A**. {$\mathbf{A_k}$} are successive approximations of "**A**", outputs of successive iterations of the solution algorithm.

… and 'the Logical Depth of Calculative Thinking' Matters to Strategic Payoffs…

[Cournot-Nash Duopoly Without Logical Omniscience]

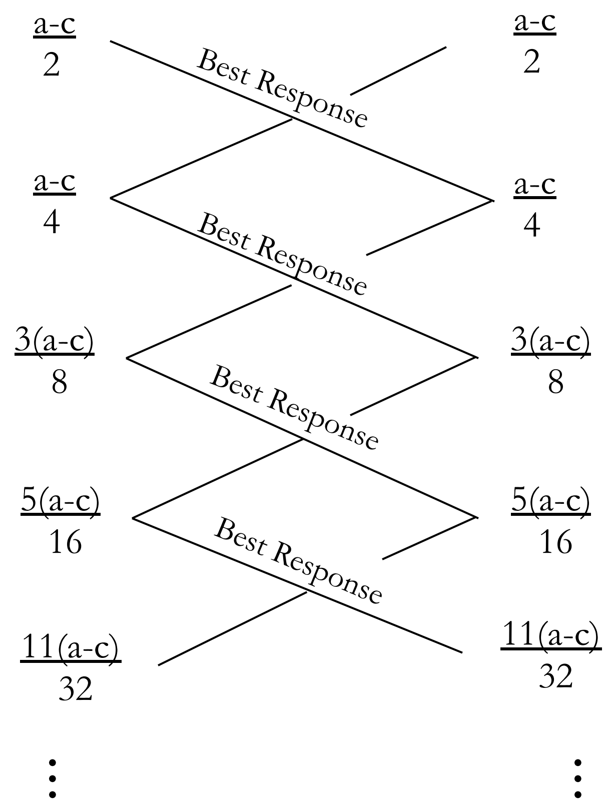
Generate using series

$q_N = \frac{a-c}{2} - \frac{q_{N-1}}{2}$;

$q_o = 0$

Which results from joint maximization of profits

$\Pi_i = (a-c-q_i-q_j)q_i$



Firm 1's quantity choice/ best response

Firm 2's quantity choice/ best response

$\frac{a-c}{3}$

NASH EQUILIBRIUM

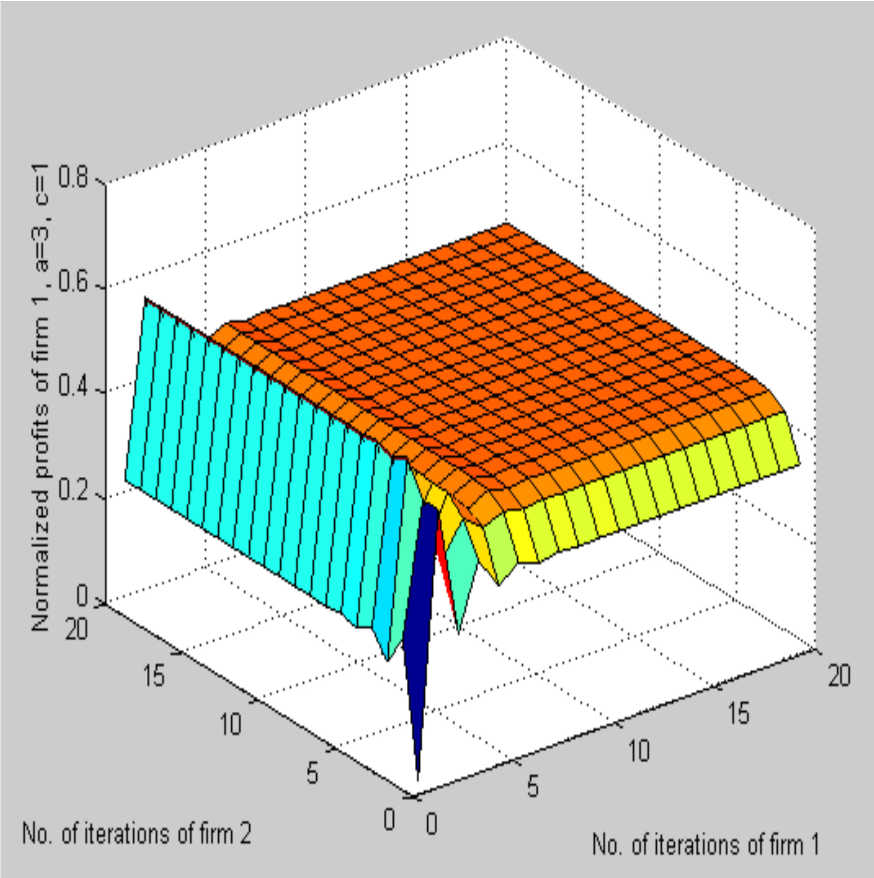
So, if firm 1 says, "I will sell $\frac{a-c}{2}$, firm 2 will credibly retort, "I will sell $\frac{a-c}{4}$ "; which would Lead to losses relative to the $\left[\frac{a-c}{3}, \frac{a-c}{3}\right]$ solution

Computational Landscapes for Interactive Problem Solving (Duopoly)

- Computational Landscape of Cournot Nash Equilibrium, 2 firms, a=3, c=1.

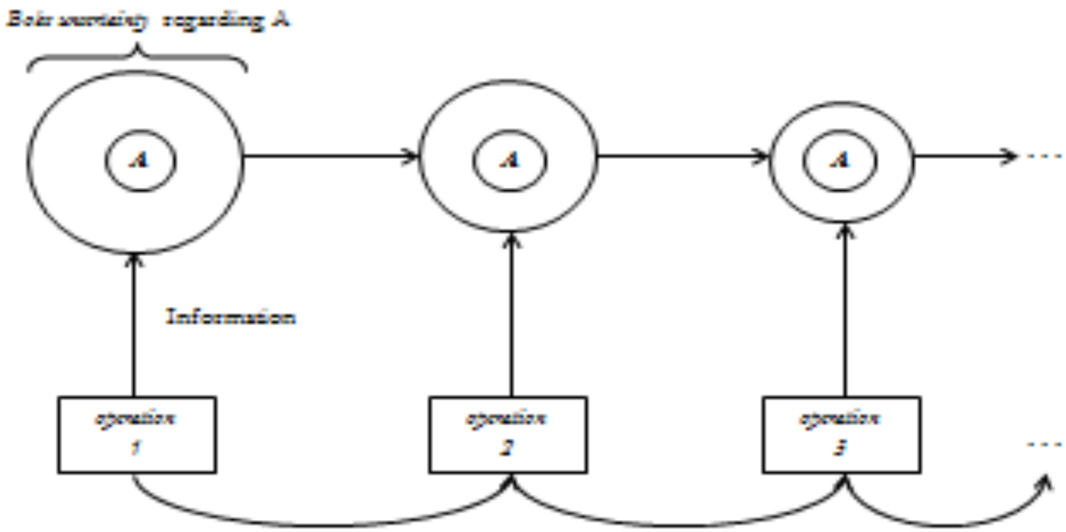
- Horizontal axes represent number of iterations for each firm. Vertical axis is the profit level of firm 1. Profit levels of firm 2 are symmetrical.

- Landscape converges to Nash Equilibrium output of (a-c)/3.



If All Problem Solving Processes Had These Dynamics, We Would Be Programming on **Brains** Right Now.

A Model of Calculation as an Information-Producing Process: Each Individual Operation Reduces the Uncertainty (Conditional Entropy of pdf(loc(answer(SearchSpace))) – Associated with Creature's Guess at an Answer, whose exact value is denoted by A.

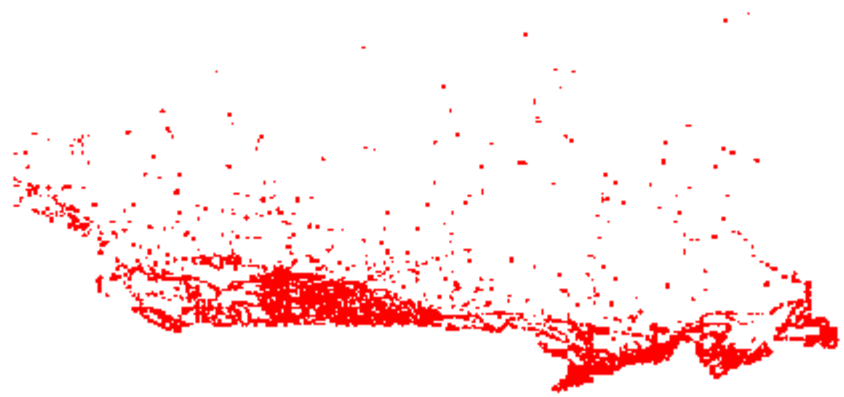


What if Bob Had to Make a Different Choice with Procedural Implications?

- *$1MM* for finding the shortest

Path connecting Canada's 4663 cities in 1 day of less, OR

- One day's consulting fees guaranteed



  - Total number of operations required

$$K \sim 2^{4663} \sim 5 \times 10^{1403}$$

His computational prowess R ~ 10^{12} ops/second

His computational budget

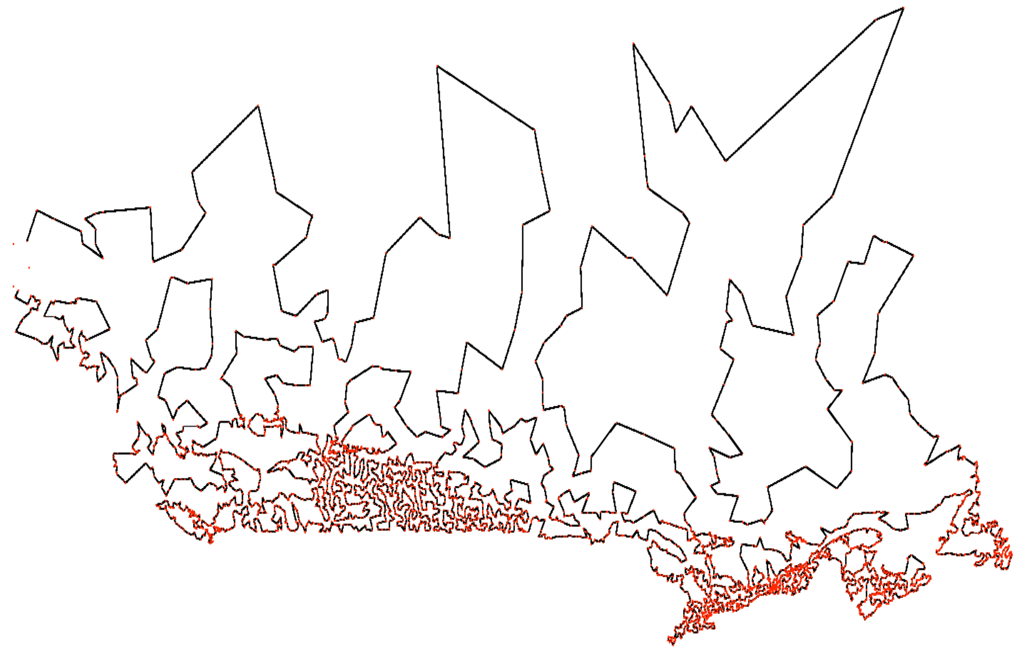
(10^{12} ops / second) (3600 sec / H) (24h / day) x(365 days / yr) ~ 3×10^{20} ops

He <u>can</u> solve this problem in 1.6×10^{1383} years

→not worth it!

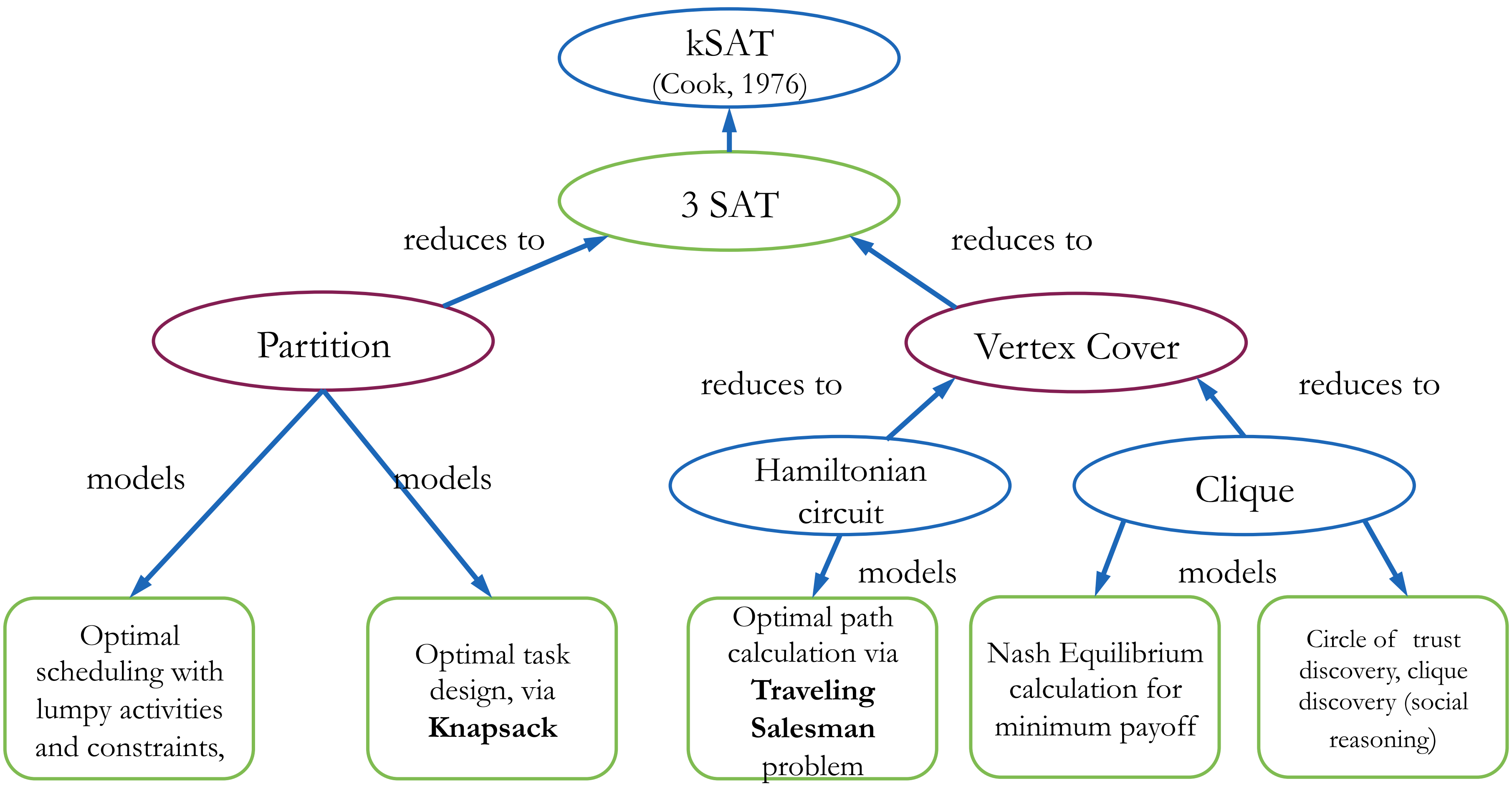
UNLESS, Bob Had Some Kind of a Short Cut

- Non-exhaustive
- Non-deterministic
- Non-universal (will not be optimal for other NP-hard optimization problems)
- Locally exportable (to other TSP's)
- Hardware-adaptable (more/less RAM, and operations per I/O cycle);

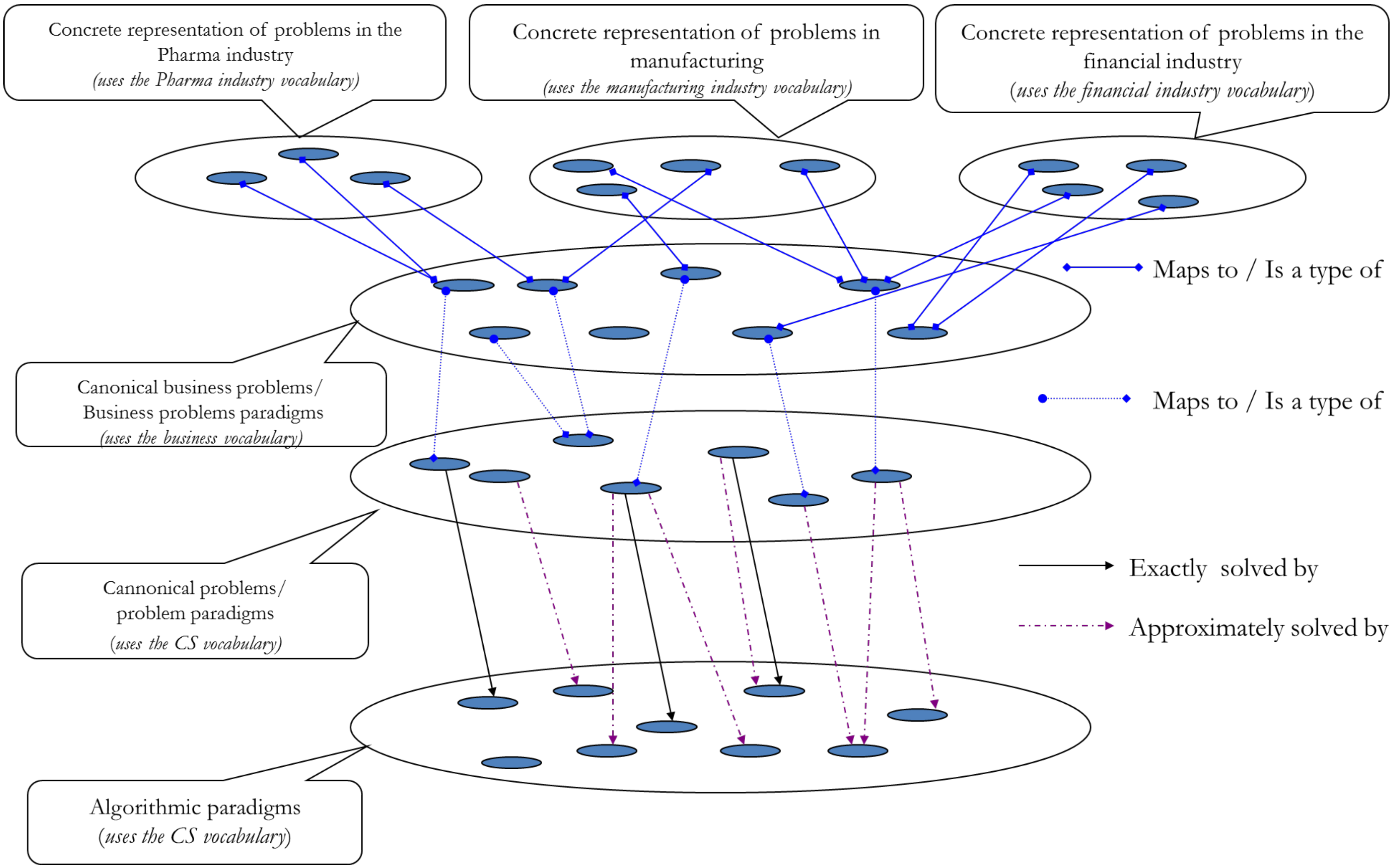


4663 city TSP, solved
using Lin-Kernighan (meta) algorithm

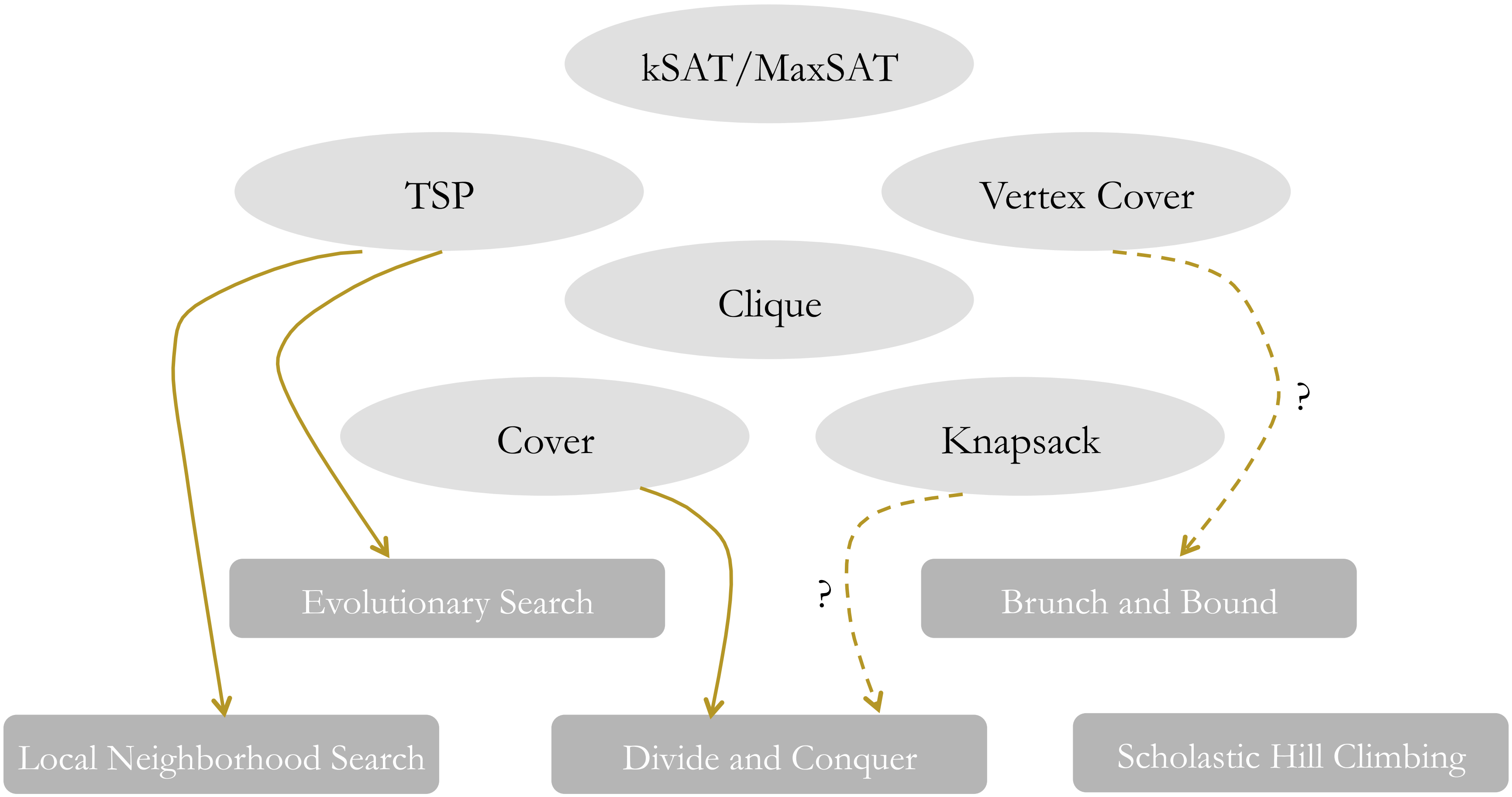
The NP Class Reads like a Who's Who of Everyday Problems (Solved by Creatures with Brains)



'Generalized Problem Solver, Version 2.X'

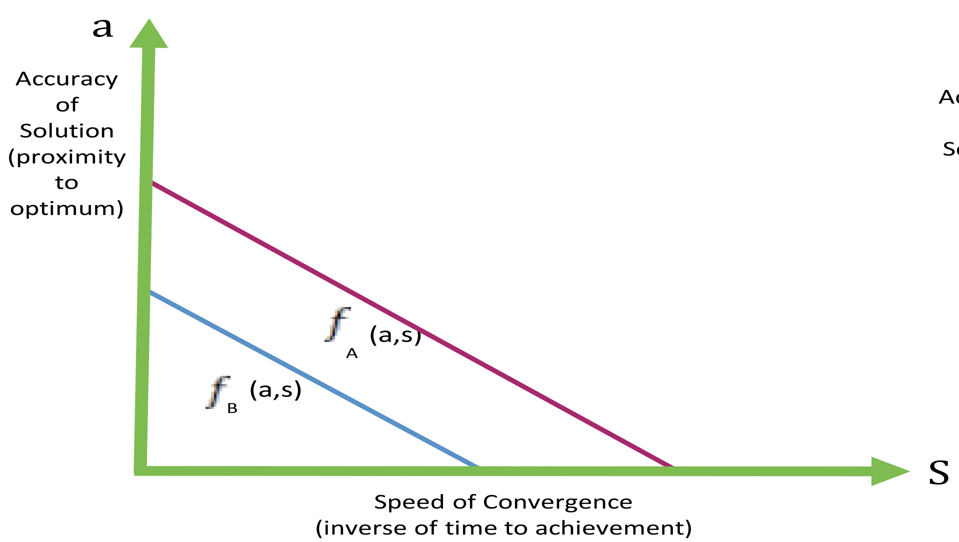


Modeling Toolkit for Problem Solving Processes: An Associative Map

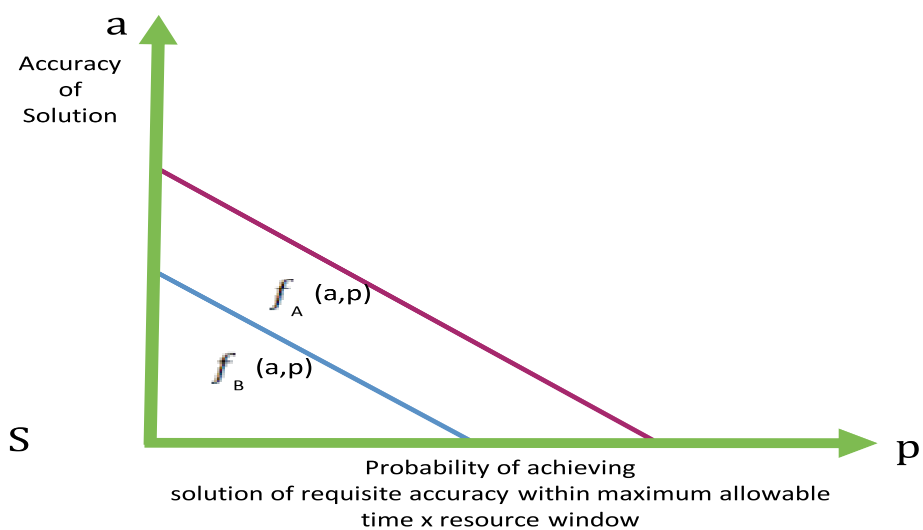


Payoffs [Problems x Procedures]

What Could Computational Payoffs Look Like? Two Separate Payoff Structures…



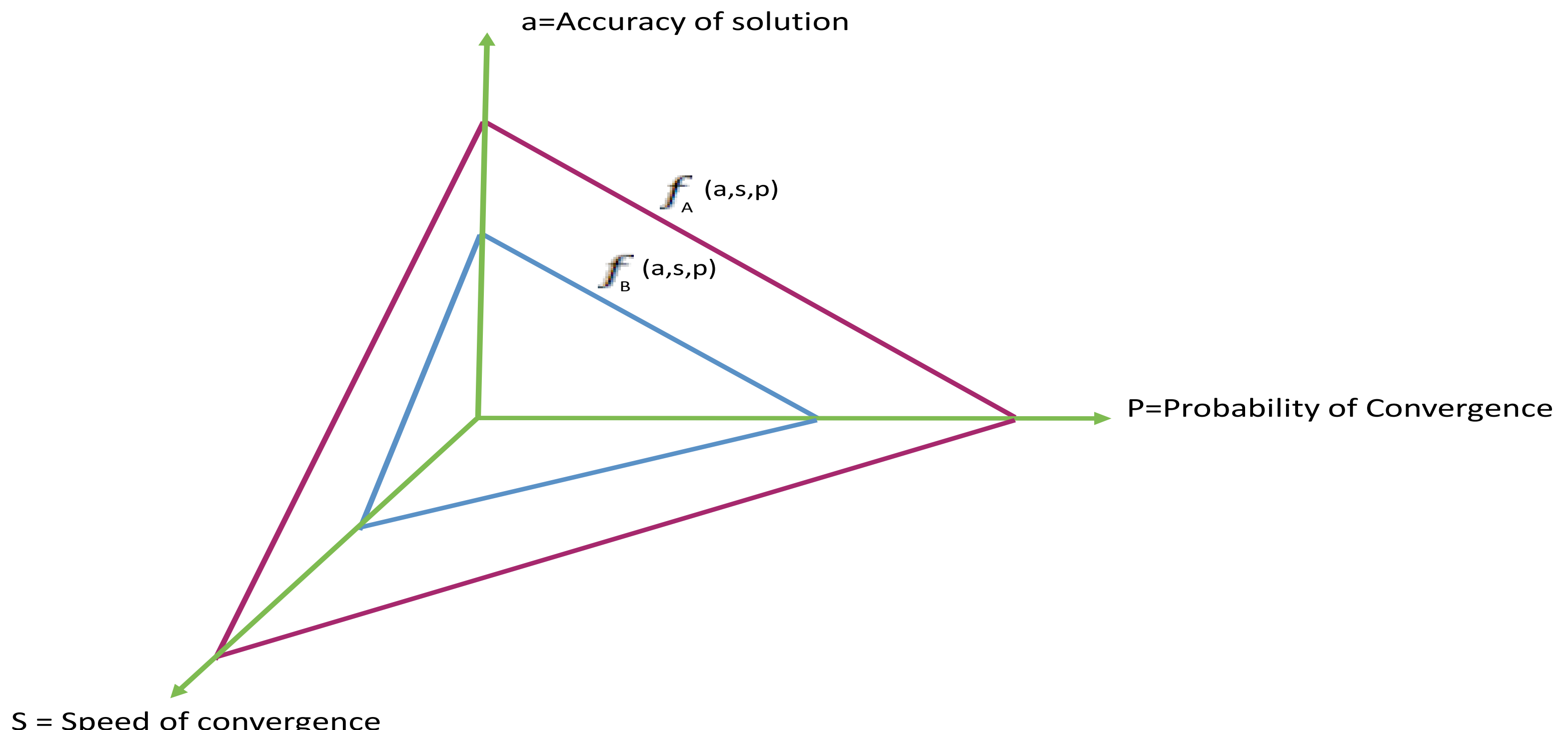
Algorithmic advantage: $f_A(a,s) \geq f_B(a,s)\ \forall a, S$



Algorithmic advantage: $f_A(a,p) \geq f_B(a,p)\ \forall a, p$

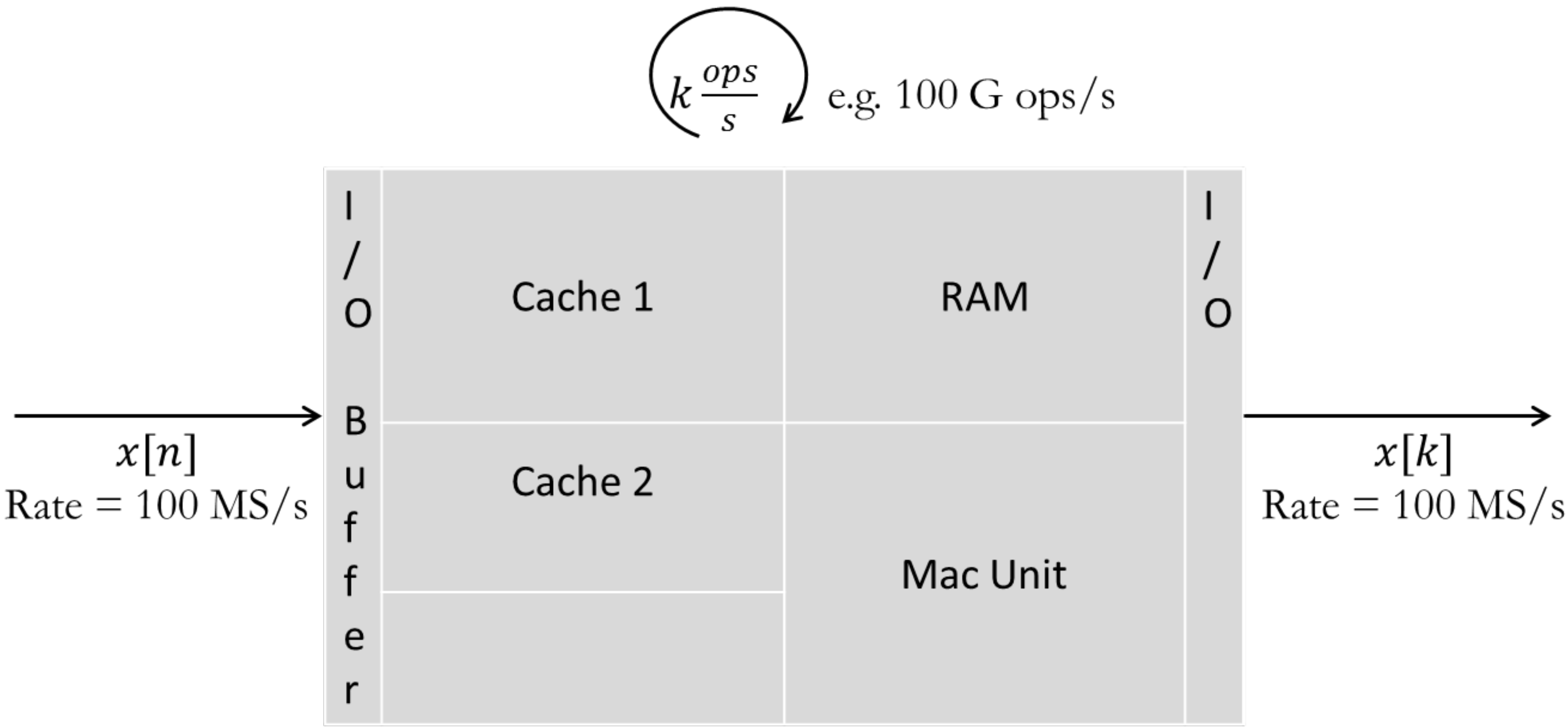
…Combine into One 3D Measure

,



Algorithmic Advantage: $f_A(a,s,p) \geq f_B(a,s,p), \forall a,s,p$

Getting Closer: How Would a Chip Designer Think About Embodied Problem Solving?

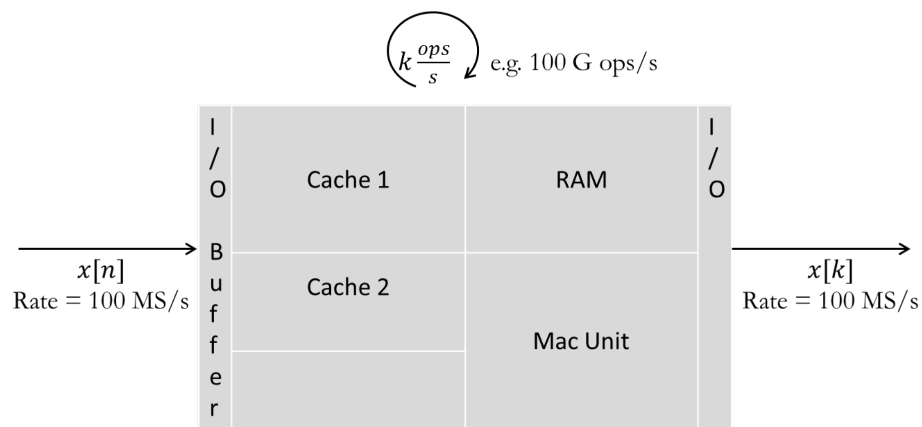


$x[k] = \frac{1}{M}\sum_{n=1}^{M} x[n]e^{j2\pi kn/M}$; M=2048

$A(x) = \text{"}Radix\ -2\ FFT\text{"}$

Using Application-Specific Chip Design as a Paradigm for Mind-Brain Investigation

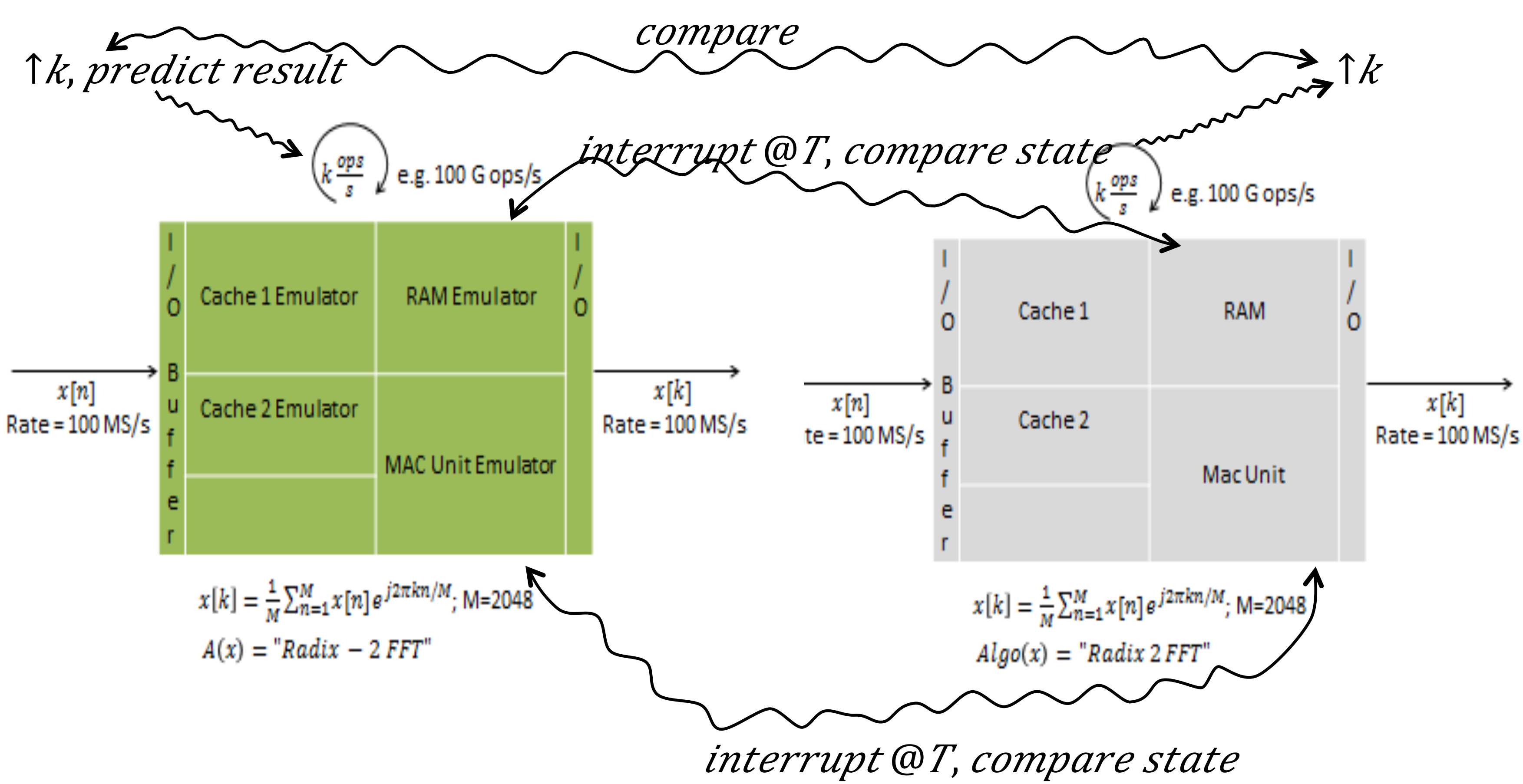
- No operation without implementation;
- No algorithmic change without architectural consequence;
- Capacity limit (Ops/sec, *M*) part of every hardware decision;
- Hardware Adaptable to Algorithm/I-O requirements (more/less RAM, operations per I/O cycle, precision of internal representation of coefficients);
- Average-case performance far more important than worst case performance(e.g. dynamic range extremes of the input *x[n]*).



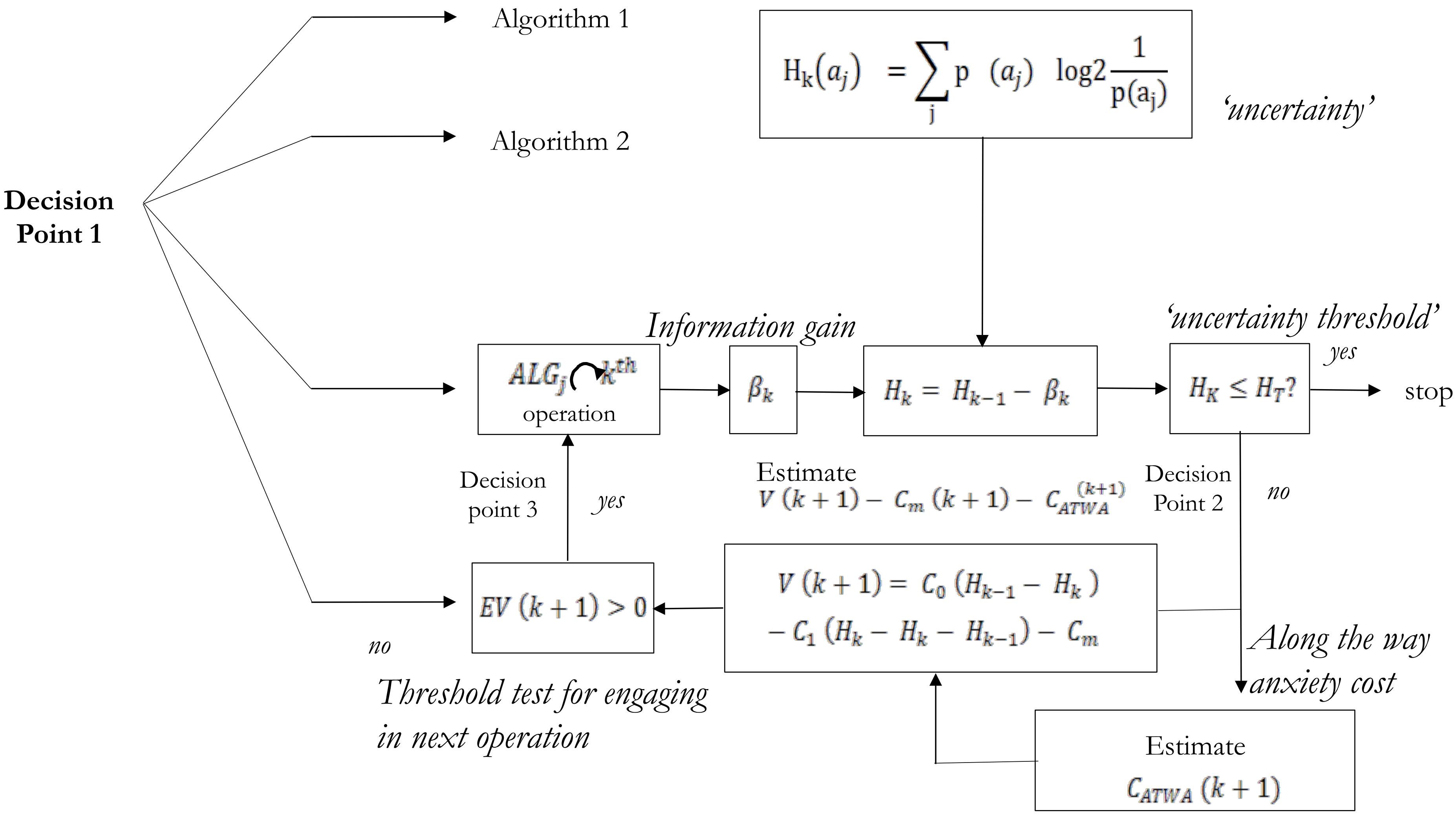
$x[k] = \frac{1}{M}\sum_{n=1}^{M} x[n]e^{j2\pi kn/M}$; M=2048

$A(x) = \text{"Radix} - 2\ FFT\text{"}$

"Simulation" Is Not Just "Modeling": It Has Bite, Which Is Why We Call It EMULATION



Of Course, Humans Can Choose Whether or Not to Proceed with an Algorithmic Computation at Many Points…



A Goal for Intelligent Artificiality: A Brain Emulator/Co-Processor

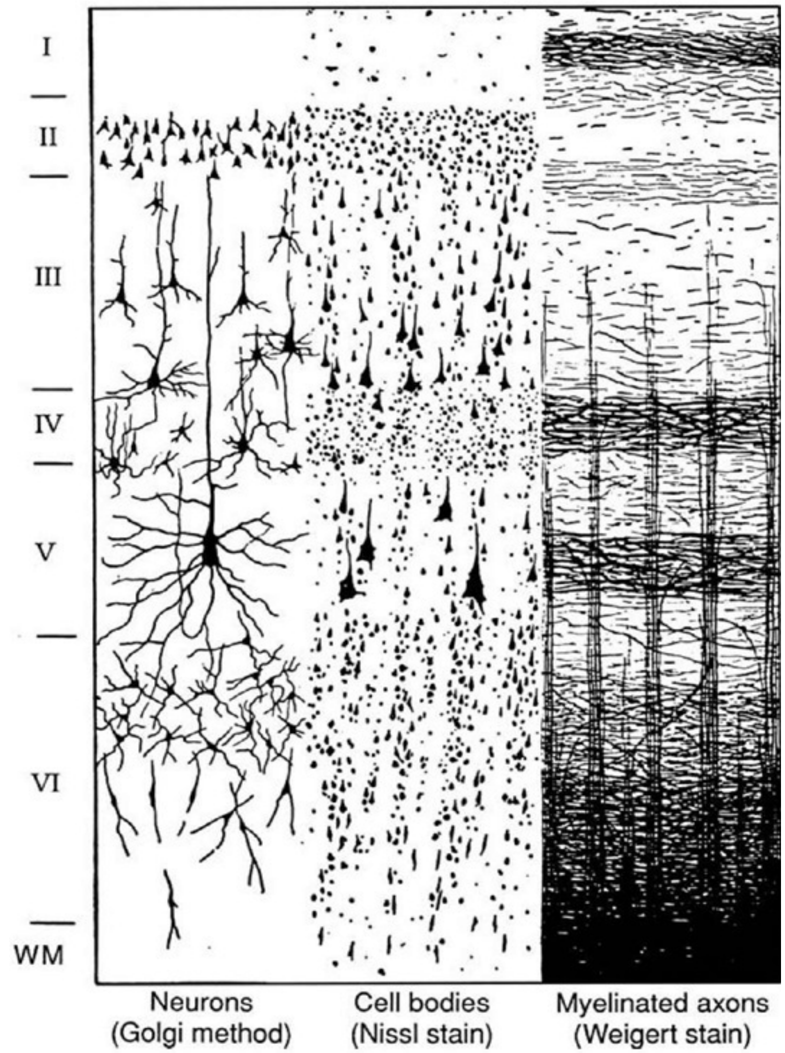
- No ‘model’ of mental behavior without architectural and behavioral consequences;

- Brain states on which mental states supervene can be tracked, not only ‘modelled’: prediction/control supersedes ‘explanation as regulative goal.

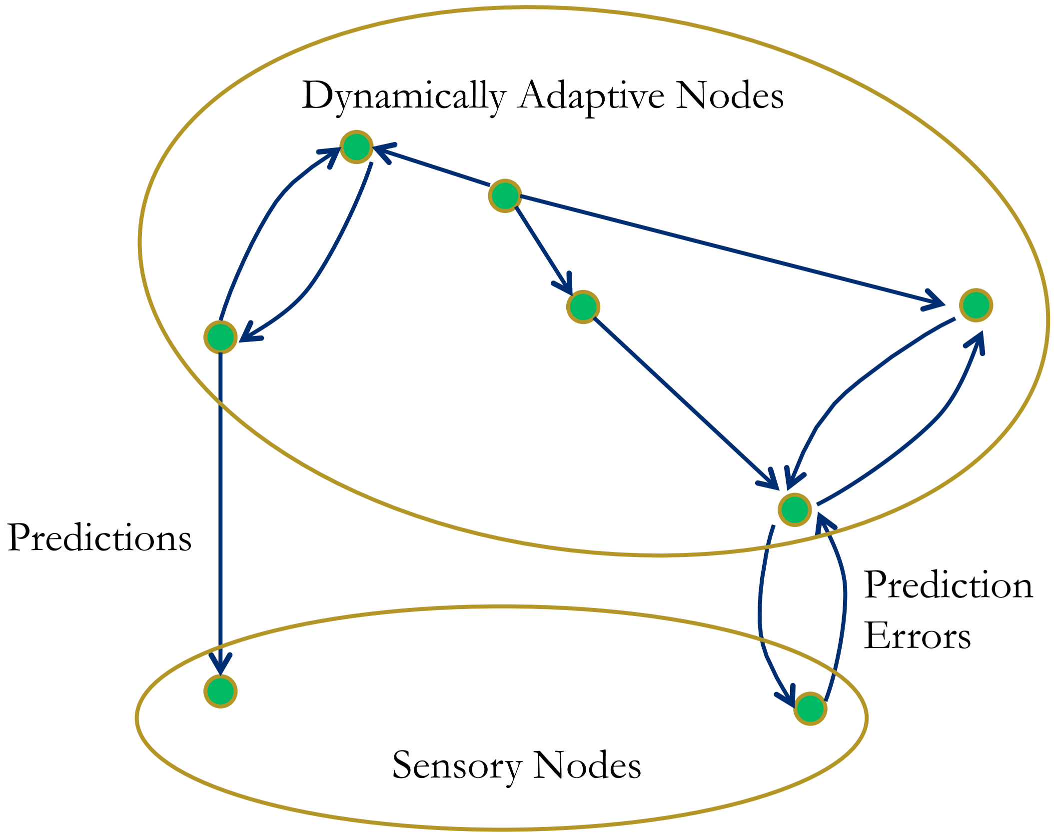
- ‘Hardware changes’ (TMS, ECT, stimulus protocols, psycho-pharm) can be emulated, enabling point predictions about mental behavior.

We Need an Anatomically Informed Model of 'Brainware'…

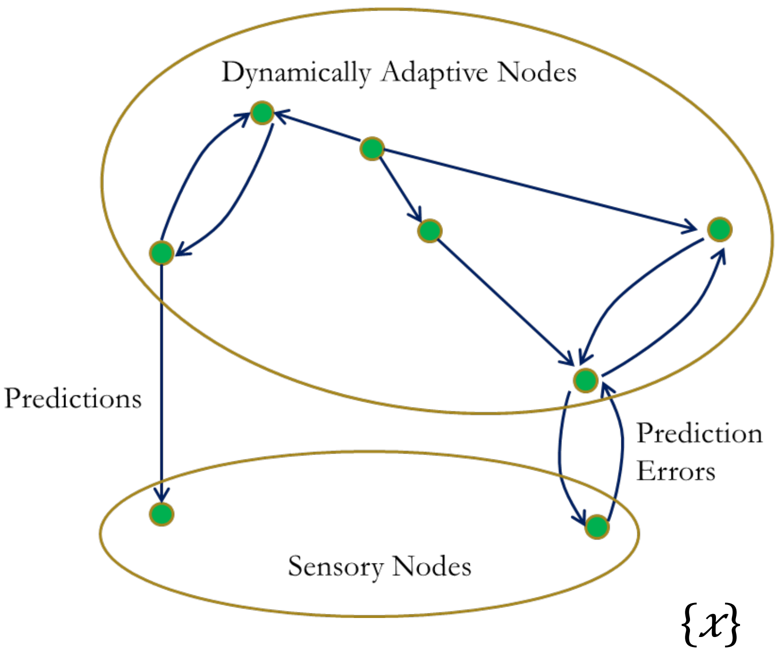
- Layered connectivity for the associative cortices;
- Cross-layer forward and backward connections (sparser), intra-layer connections (denser);
- Some (parametrizable) asymmetry between forward and backward connections;
- Architectural levers include strength of synaptic connections, 'plastic' formation of new circuits.



… That Is 'Emulable' via a Well Understood Structure (Recurrent Neural Network)



… Which Extremizes an Objective Function Familiar to Self-Organizing (Entropy-Increase-Defying) Systems…



Kullback-Leibler divergence of p,q

"Spread of q" (ENTROPY)

$$\min_{\{\alpha, s\}} F(x,s/\alpha) = -\langle \ln p(x,\theta/\alpha) \rangle_q + \langle \ln q(\theta,s) \rangle_q$$

actions

causes

internal states

Sensory states

Recognition problems (inputs, causes)

… to Provide an Extremisand That 'Works' at Different Space-Time Scales and in Different Domains of Being.

$$\min F(x,s/\alpha) = \min -\langle \ln p(x,\theta/\alpha)\rangle_q + \langle \ln q(\theta,s)\rangle_q$$

$\underbrace{\min_T\{\alpha\} F}$ actions

$\underbrace{\min_T\{S,(s)\} F}$ sampling

$\underbrace{\min_T\{S_I,(s)\} F}$ inference

$$q = \underbrace{q(\theta_p, S_p)}_{\text{"pico": synaptic weight changes}} * \underbrace{q(\theta_m, S_m)}_{\text{"micro": attentional shifts}} * \underbrace{q(\theta_M, S_M)}_{\text{"macro"}}$$

Now, If We Could Only Explain Away 'Complexity Mismatches' – Which We Can!

$Encoding\ (p,q)\ via\ Kolm\ (p,g)$:

$Kolm_M(x) = length(A_M, min(x))$

"Efficient coding":

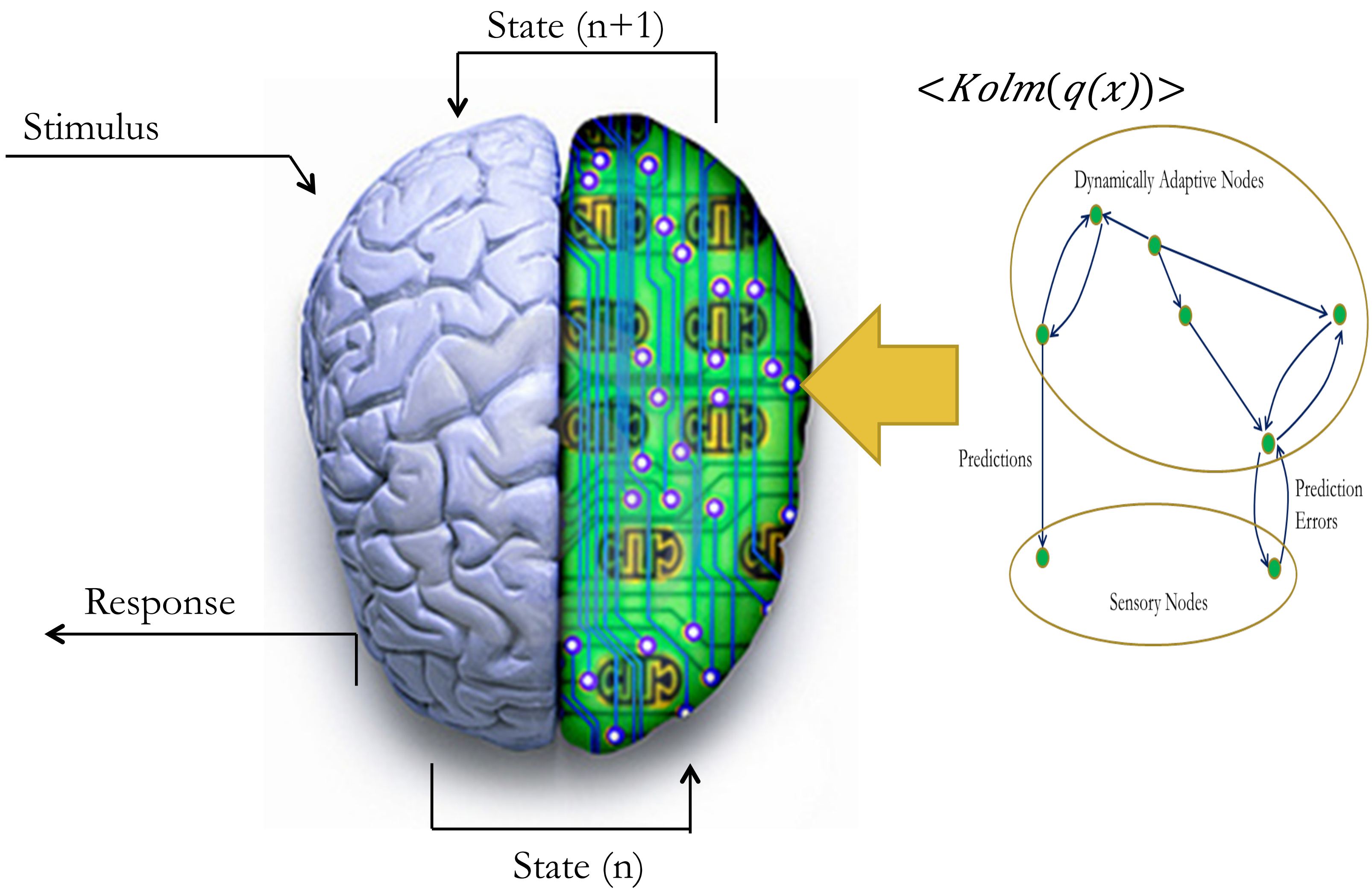
$p(x) = 2^{-Kolm_M(x)}$ [$Kraft-McMillan\ coding$]

$Information\ content\ of\ p(x)$:

$-\sum^{K} p(x) \log_2 p(x) = \sum 2^{-Kolm(x)} Kolm_M(x) = \langle Kolm_M(x) \rangle_p$

$Let\ M = CORT \rightarrow K_M(\cdot) = K_{CORT}(\cdot)$

We Can Rebuild a 'Theory of Computation' Using 'Brainware' as the Computational Substrate



… and Fill in the Gaps of Both Symbolic Representation and 'Rational Choice' Approaches

not $\max_{x,y,z,\ldots} U(x,y,z,\ldots) \; s.t \; B(x,y,z,\ldots;t) \leq B^{*}$:

PROCEDURALLY OPAQUE;
ARCHITECTURALLY INDETERMINATE;
PHYSICALLY UNREALIZABLE IN MANY CASES OF INTEREST

not $\max_{\{P\},\{A\}} V(P_1 , P_2 , \ldots P_m \mid A_{11} , A_{12} \ldots A_{mn}) \; s.t \; Comp\,(A_{jk} / P_k) \leq Con$

PROCEDURALLY UNREALISTIC;
ARCHITECTURALLY INAPPLICABLE;
WORST CASE EMPHASIS UNREASONABLE IN MOST CASES OF INTEREST

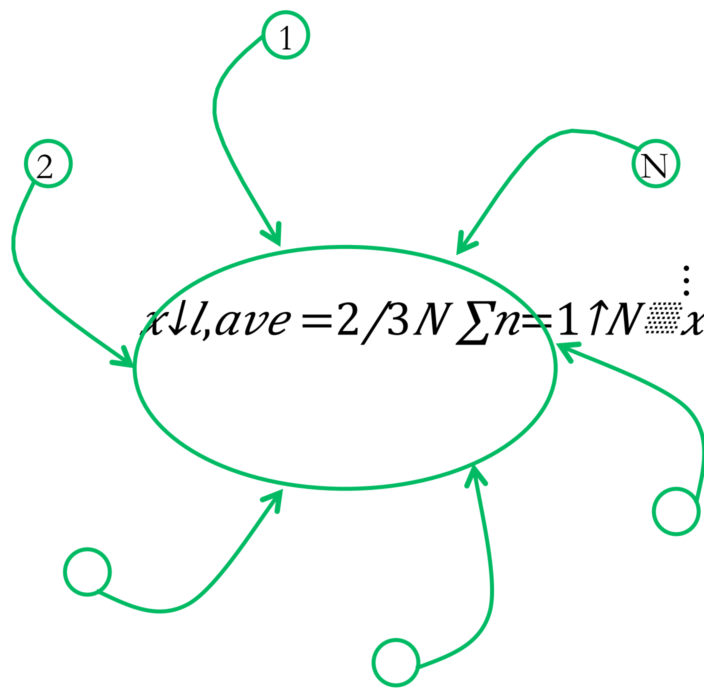
Circumventing Logically Deep Equilibrium Calculations: Beauty Contest Example

- *N* players, 1 period game;
- Each player submits number from *0 to N* to a(n honest) clearing house.
- Winner (gets \$*N x \$1000*) of the game is the player that submits the number that is closest to *2/3* of the average of all the other numbers.
- Iterated dominance reasoning:

if I submit **x** *and others submit (y,z,w,…) then winner would have* **had to have** *submitted* **z,** so **I should have submitted** *y.*

- Equilibrium submission ('strategy') is

0: **(***2/3)(0)=0*



$$x_{l,ave} = \frac{2}{3N}\sum_{n=1}^{N} x_n$$

e.g. $N=100$, $x_n \sim [0,100]$

$x_{ave}=50$, $x_{l,ave}=33.33\ldots$

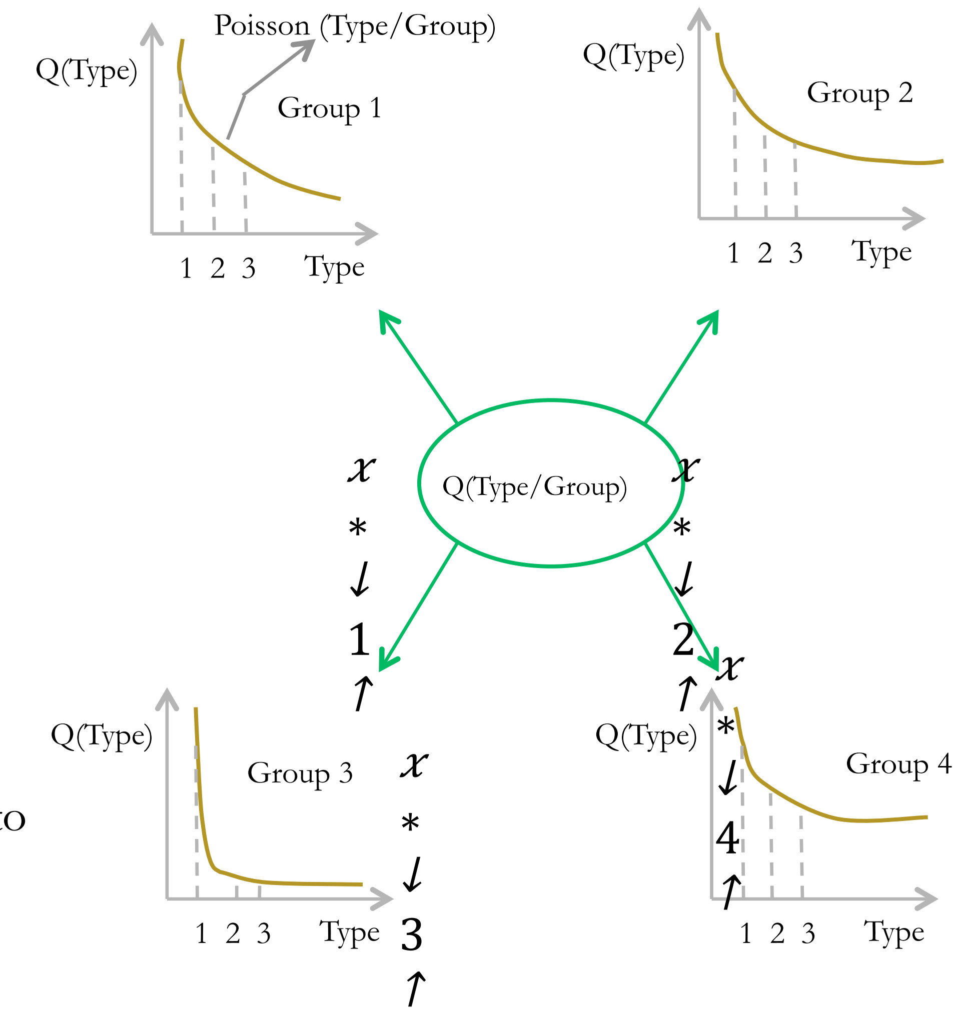
$x_{l+1,ave}=2/3\; x_{l,ave} \rightarrow x^* = 0$

Circumventing Logically Deep Equilibrium Calculations: Beauty Contest Example

- Encode *others* via Types (Ho, 2004)
- **Type 0** players do not think of what others think;
- **Type 1** players think only of what others think;
- **Type 2** players think of what **Type 0** and Type 1 players think only;
- **Type** *k* players think of what

Type (k-1, k-2,…) players think only.

- Define **Q(***this group type set)* as estimate of density of **Type** *k* players in **this group**.
- Refine *Q(types) (mode, spread)* according to cues.



Intelligent Artificiality

A Foundation for Mind-Brain Design, Diagnostics and Development