An Information Flow Perspective on Access Consciousness

Michael C. Mozer
Donald W. Mathis
Jennifer Finger
Michael Colagrosso

Institute of Cognitive Science and Department of Computer Science
University of Colorado, Boulder

Consciousness Course
Spring 2004

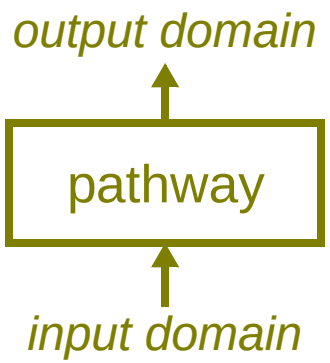
Cognitive Architecture

Cortical computation can be characterized by coarse-scale, functionally specialized pathways.

E.g., visual word-form recognition
identification of semantic features of visual objects
auditory word recognition
computation of spatial relationships
construction of motor plans

Cognition requires coordination of multiple pathways.

Pathways act as associative memories.



E.g., visual word recognition

input: visual features
output: letter strings

Past experience imposes well-formedness constraints on the output.

Letter strings must be consistent with English orthography.

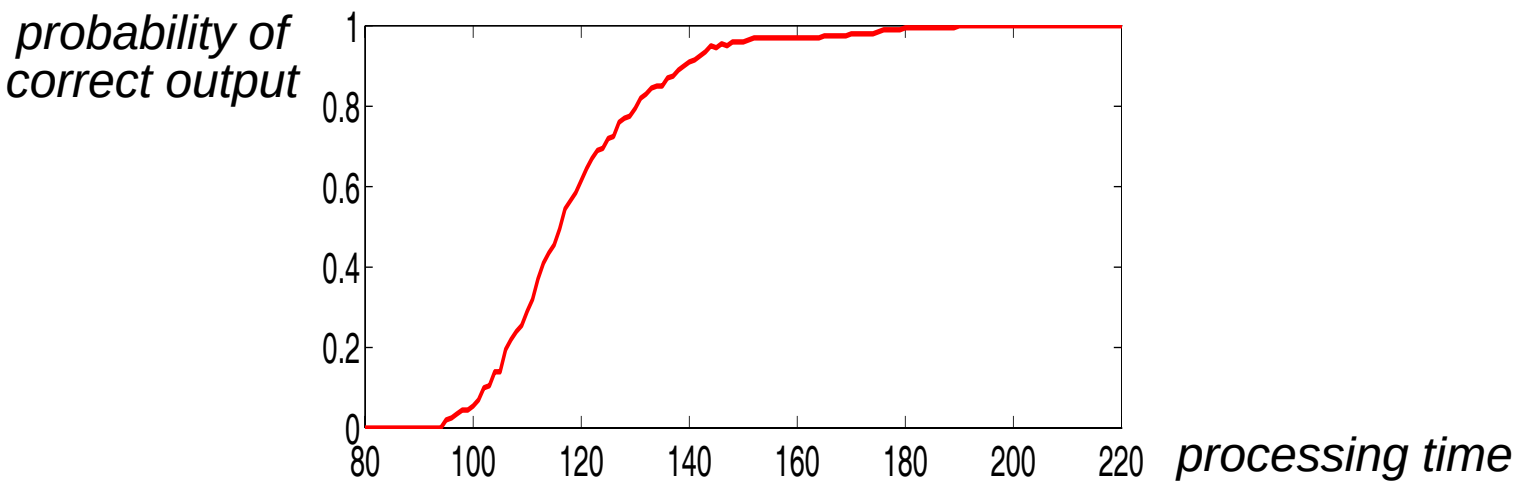
Output is the best-fitting interpretation of the input.

Cognitive Architecture

Pathway operation shows a speed-accuracy trade off.

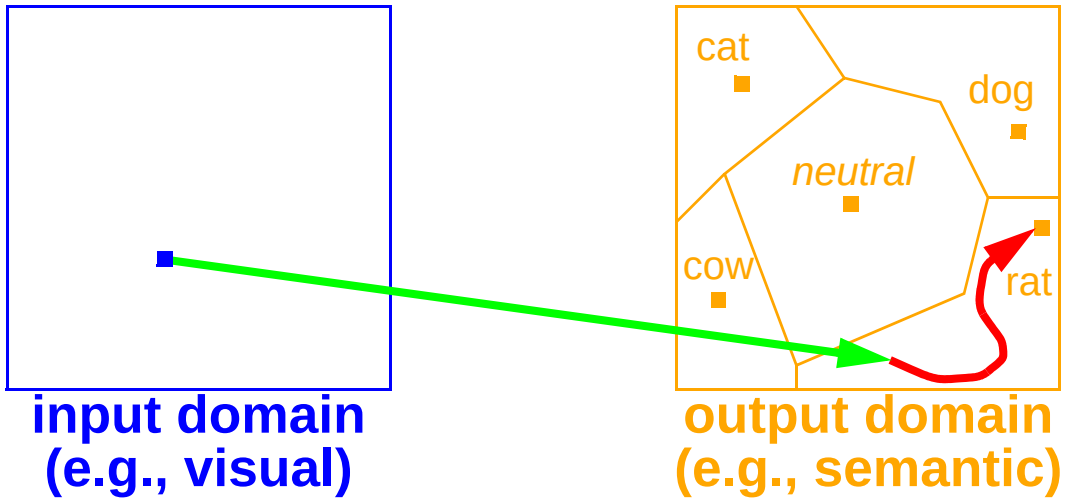
Initial output appears rapidly, but is inaccurate.

Pathway asymptotically converges on the best interpretation.

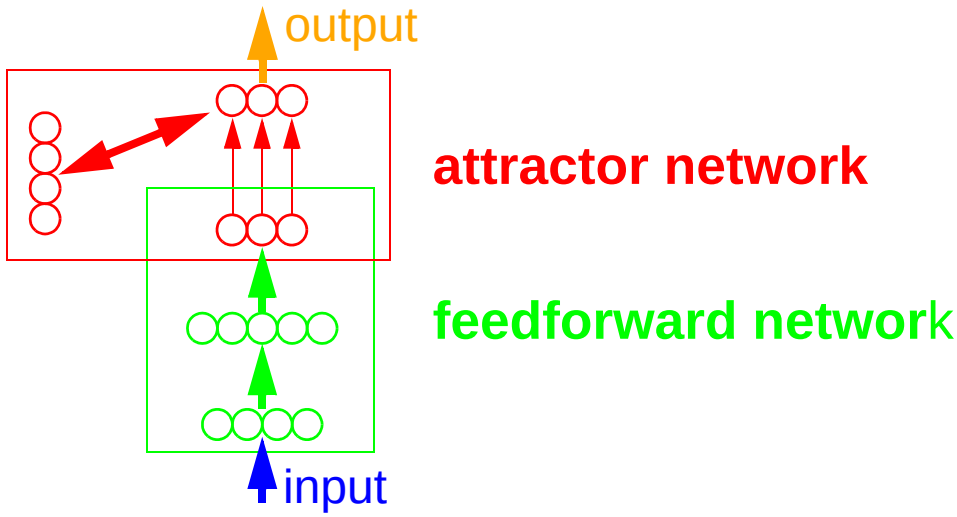


With each experience, pathway tends to produce its response more rapidly.

Pathway Operation



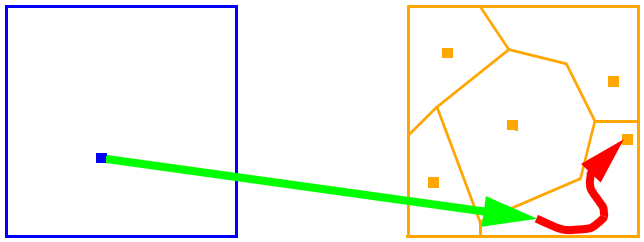
Implementation using two standard connectionist components:



Pathway Implementation

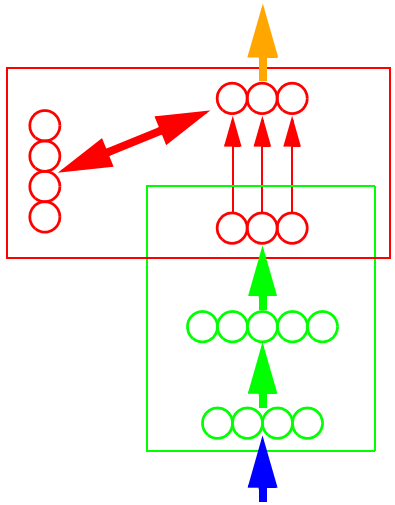
Speed-accuracy trade off due to

- inaccuracies in feedforward mapping
- Gaussian noise added to unit activities

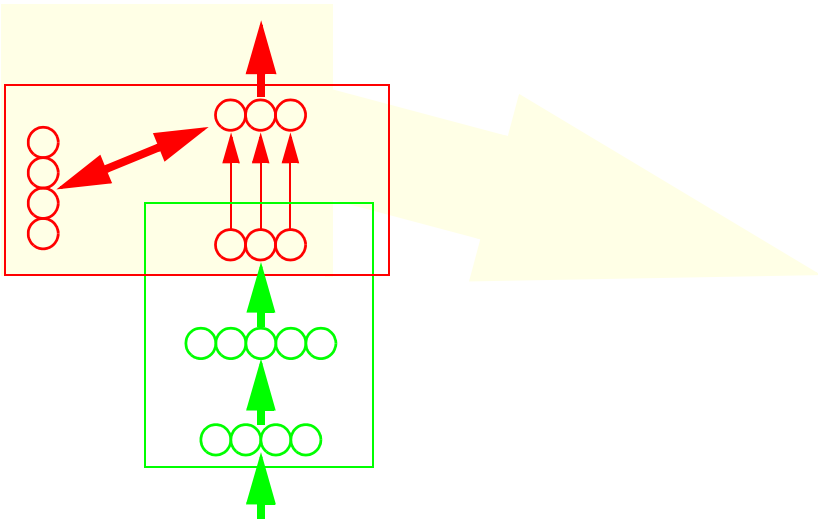


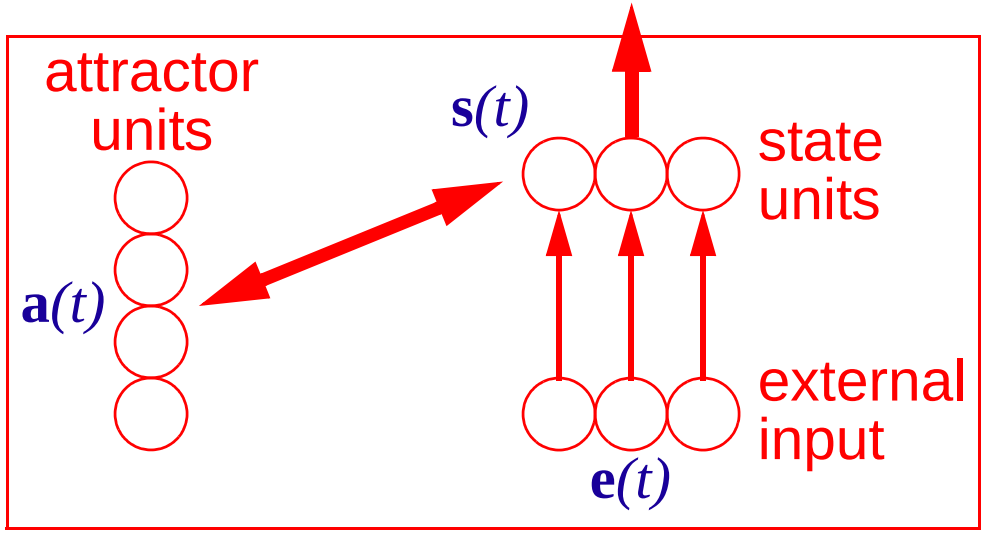
Common connectionist architecture

- Hinton & Shallice (1991); Plaut & Shallice (1993); Mozer & Behrmann (1990)
- Cascaded attractor net models incorporate feedforward mappings (e.g., McClelland & Rumelhart, 1981).
- Feedforward net models implicitly assume attractor net for read out (e.g., Sejnowski & Rosenberg, 1987).



Pathway Dynamics





attractor unit update equation:

$$\hat{a}_j(t) = \exp(-\|\mathbf{s}(t) - \mu_j\|^2 / \beta_j)$$

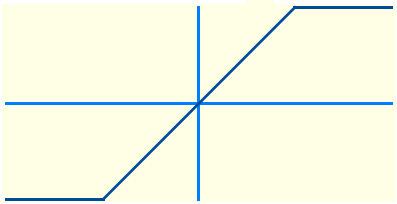
$$a_j(t) = \hat{a}_j(t) / \sum \hat{a}_i(t)$$

μ_j: center of attractor j

β_j: width of attractor j

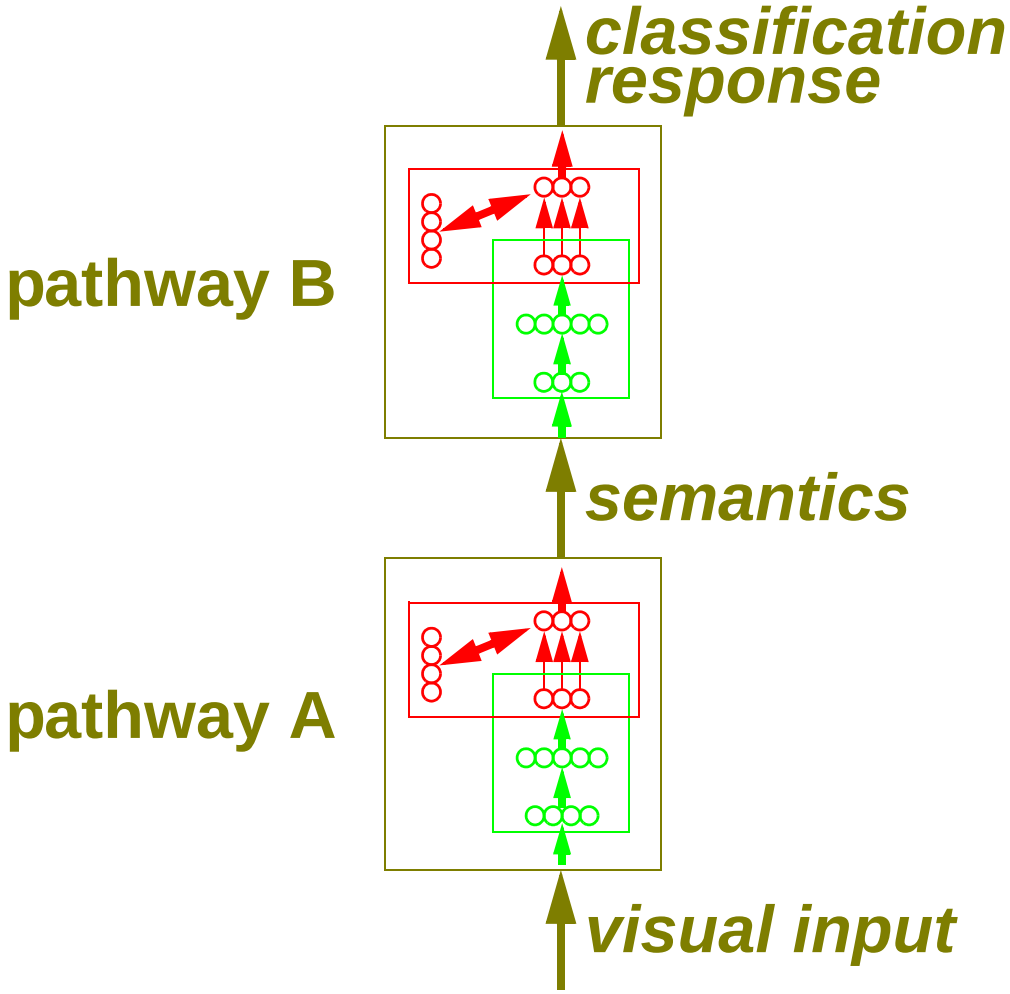
state unit update equation:

$$s_i(t) = h\left[s_i(t-1) + \omega_{ext} e_i(t) + \omega_{attr} \sum_i a_j(t-1)\mu_{ji} + \eta_i(t;\sigma) \right]$$



free parameters

Cascaded Pathways

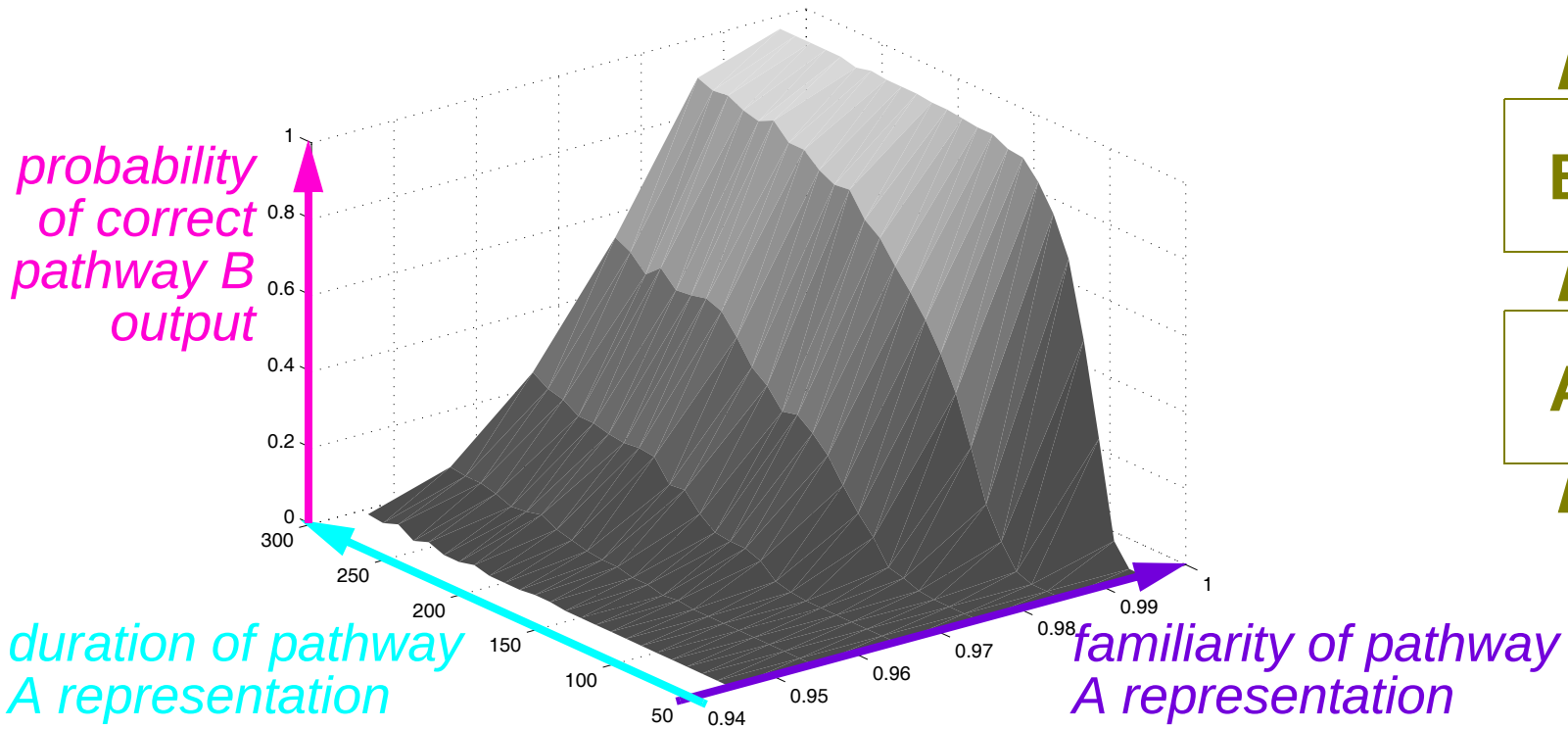


Sequential operation?

- biologically implausible
- computationally inefficient

Information Flow

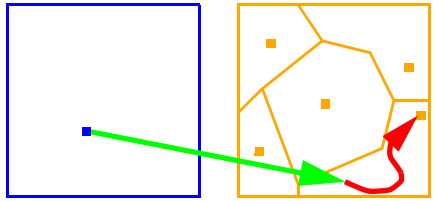
What sort of representation in the output of pathway A will support a response from pathway B?



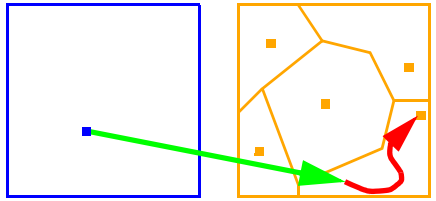
Sufficient condition: familiar and temporally persistent representation

Information Flow

Familiar and persistent representation attained when pathway reaches attractor



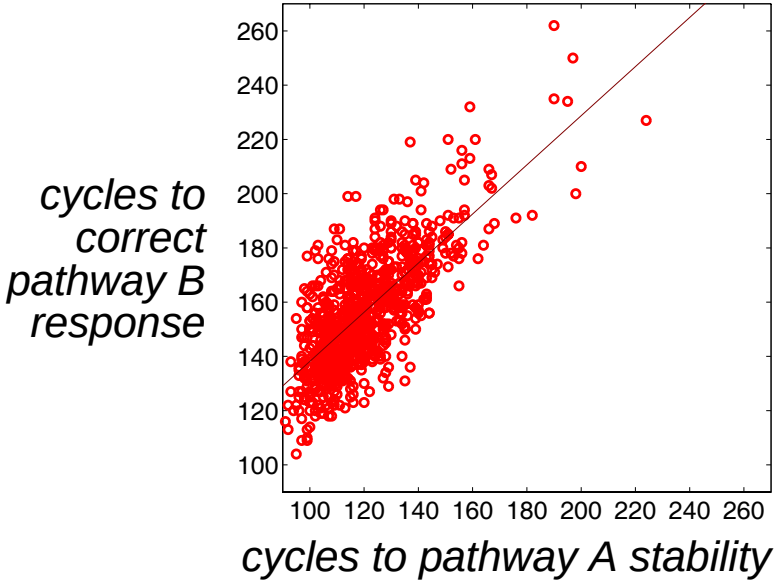
Information Flow



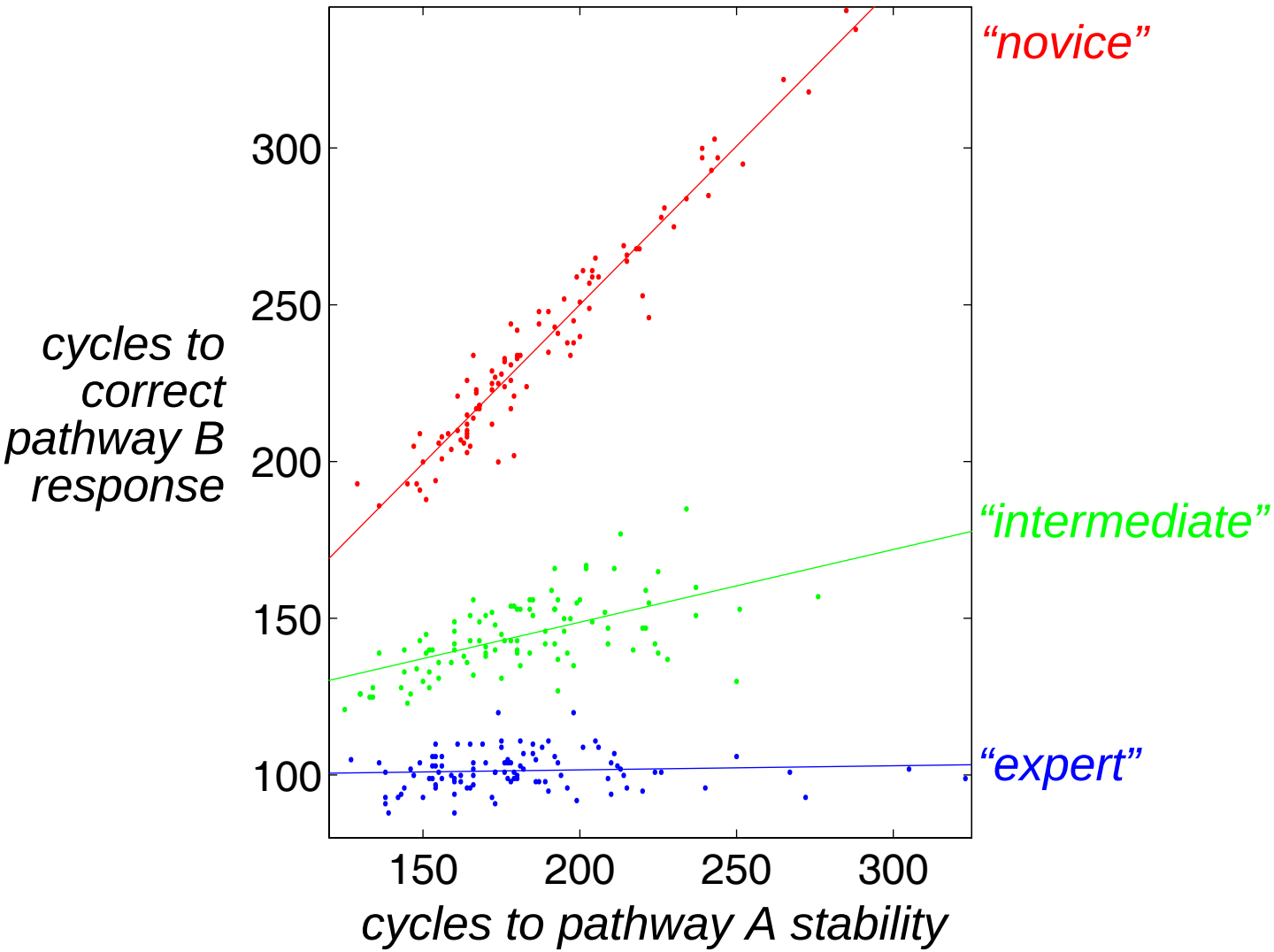
Familiar and persistent representation attained when pathway reaches attractor

At attractor, representation is *stable*.

Stability defined to occur if $\bar{d}(t) < \zeta$**,**
where $\bar{d}(t) = \alpha \bar{d}(t-1) + (1-\alpha) d(t)$ **,**
$d(t) = \|\mathbf{s}(t) - \mathbf{s}(t-1)\|$ **,**
$\zeta = 0.25$**, and** $\alpha = 0.9$

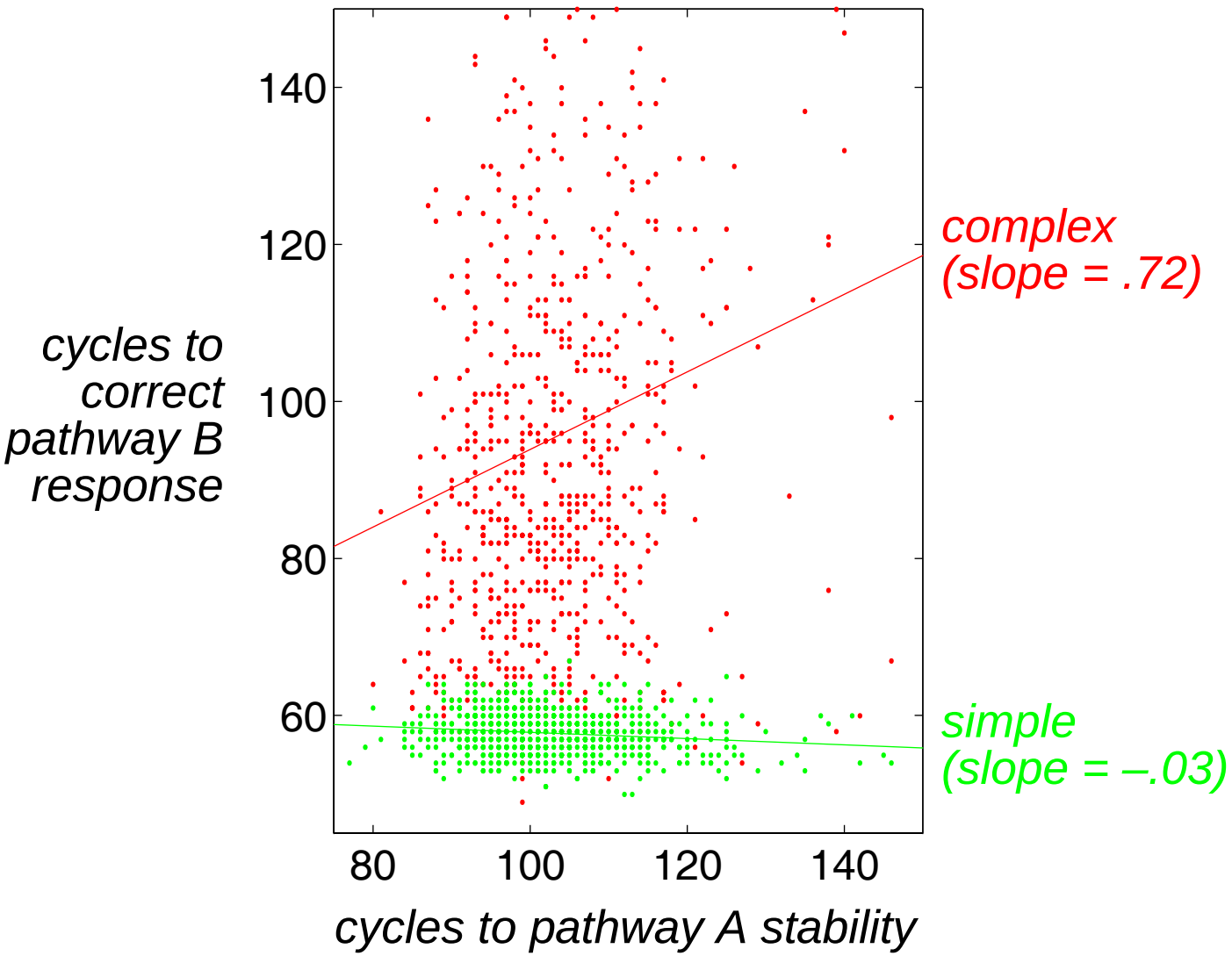


Level of Expertise



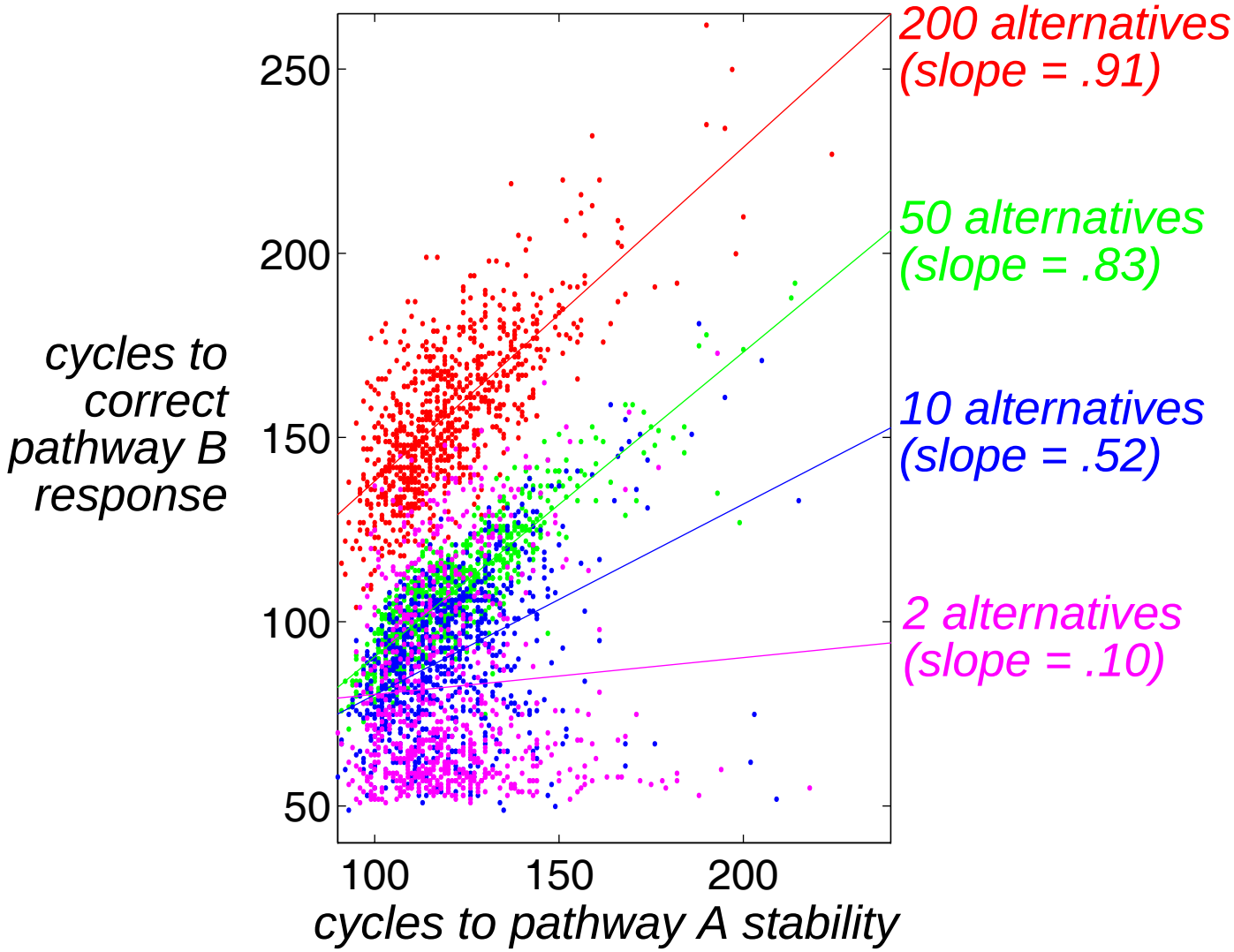
Stability required to initiate underlearned responses.

Task Difficulty



Stability required for tasks relying on finer granularity of information.

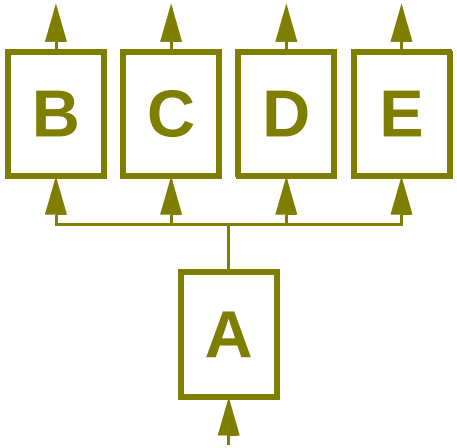
Number of Response Alternatives



Stability required for discriminations involving a large number of alternatives.

Role of Stability

Stability of a pathway's output is *sufficient* to induce the correct response from pathways to which it is connected.

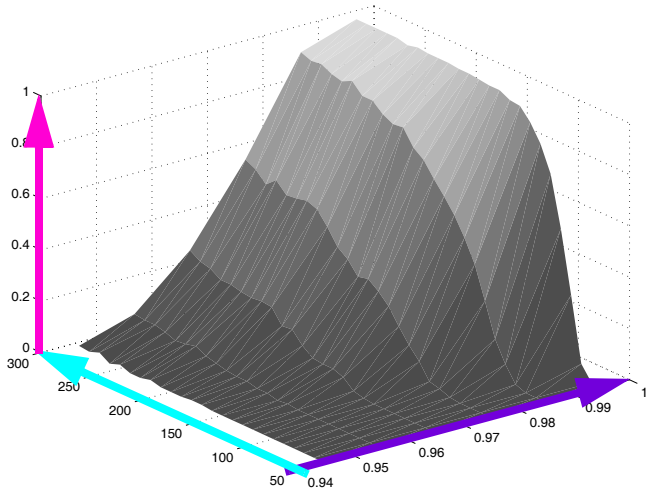


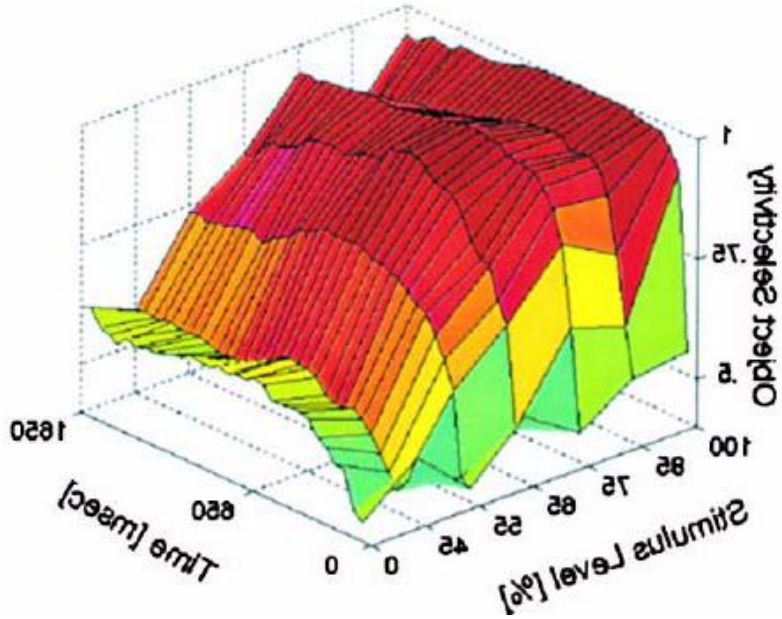
Stability of a pathway's output is *necessary* when

- decision making with limited domain expertise
- responses are arbitrary or complex
- responses involve discrimination among many alternatives

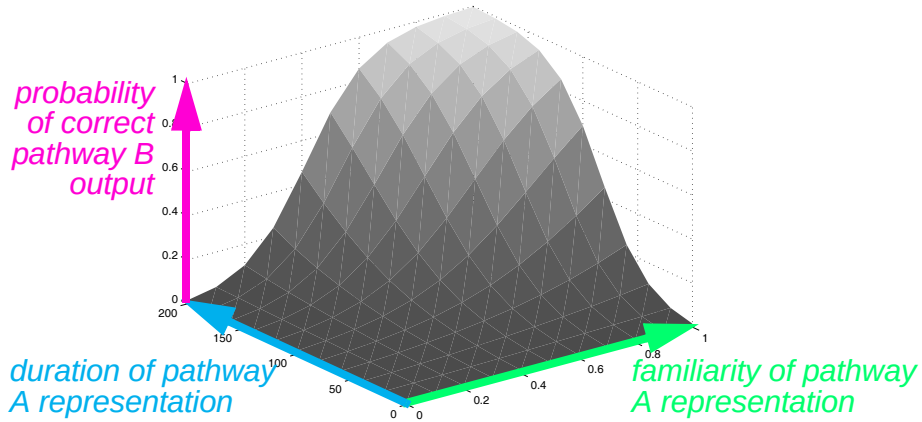
Stable representations are the most *accessible* — they allow for the flexible control of behavior (including verbal report).

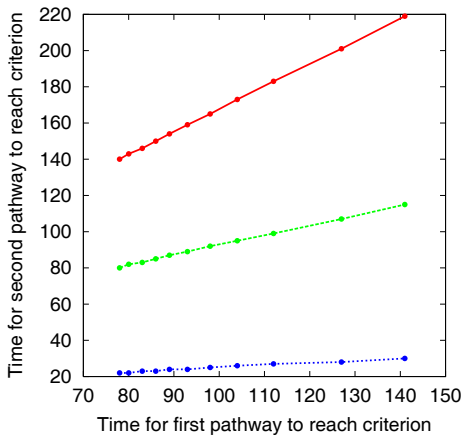
Neurobiological Correlate of Stability?

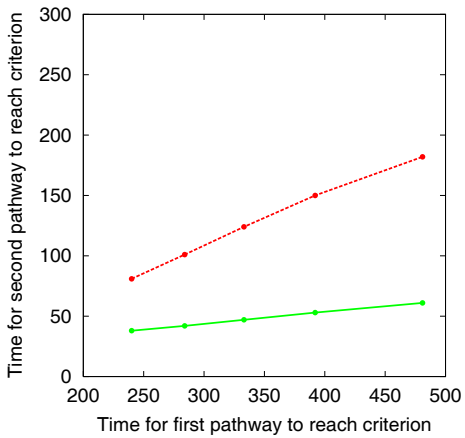


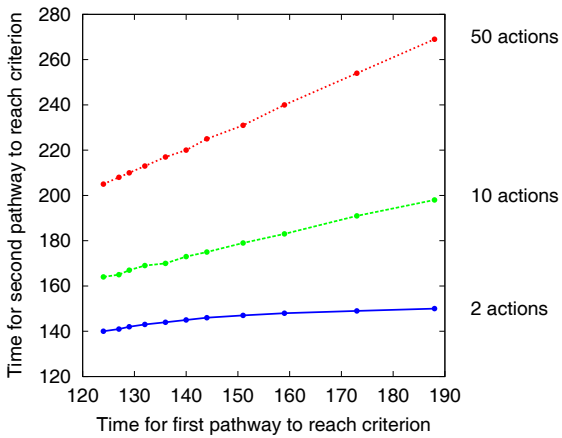


Similar Properties for Belief Net Implementation









Relation to Other Stories about Consciousness 1

Conscious states are interpretations

Gregory: ambiguous objects
Crick & Koch: high frequency flicker not perceived
Logothetis, Lumer: rivalrous displays
Marcel: subliminal semantic priming

Attractor state is an interpretation.

Relation to Other Stories about Consciousness 2

Conscious states require high quality representations

Farah, O'Reilly, & Vecera: prosopagnosia model
Munakata

Any attractor state is a high quality representation, in the sense that adding noise moves away from attractor and also lowers quality of a representation.

Probably makes more sense to focus on familiarity than quality:
Familiarity a natural way of explaining how pathway A can have an effect on pathway B.

Relation to Other Stories about Consciousness 3

Conscious states require temporal persistence

O'Brien & Opie: "stability" is a characteristic of phenomenal awareness
Crick & Koch: "coalitions" that are sustained in time; "snapshots" that must cross a threshold in time
Dehaene & Naccache: "information must be maintained over a sufficient duration"

My story is about access consciousness.

Crick & Koch coalitions are probably coarser scale (front vs. back)

Relation to Other Stories about Consciousness 4

Conscious states require explicit representations

Crick & Koch, O'Brien & Opie

Relation to Other Stories about Consciousness 5

Conscious states are attractors

Rumelhart & McClelland, Smolensky: cited in Farah et al.
Crick & Koch: dynamic competition among coalitions

Relation to Other Stories about Consciousness 6

Consciousness entails global availability of information to to widespread brain sources

Baars, Dehaene & Naccache: global workspace framework

My story requires functional connectivity between perceptual and response pathways for access to occur.

Frontal areas could provide this functional connectivity, so global workspace ideas complement my story.

Relation to Other Stories about Consciousness 7

Attention is a prerequisite for consciousness

Dehaene & Naccache; Crick & Koch

In my story, attention is not necessary, but it makes awareness more likely: Standard notion of attention is to amplify some representations. Attractor corresponding to amplified representation more likely to be chosen.

Semantic Priming Paradigm

Task: lexical decision to target word following prime word

Congruent condition

prime	LEAF
target	TREE

Unrelated condition

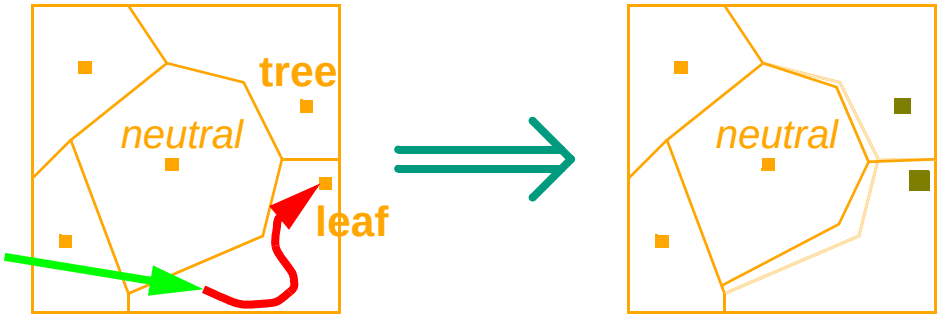
prime	AUTO
target	TREE

Facilitation on lexical decision RT for targets congruent with prime.

Modeling Semantic Priming

Presentation of stimulus *strengthens* attractors that are visited (McClelland & Rumelhart, 1986; Becker et al., 1997)

extended presentation

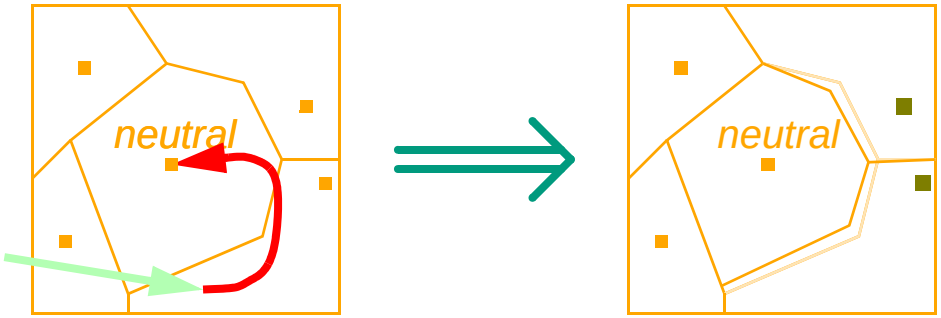


Nearby (similar) attractors are also strengthened.

Priming depends on state space trajectory.

Priming can occur without stability (awareness).

brief presentation



Semantic Priming with Polysemous Words

(Marcel, 1980; Swinney, 1979; Tanenhaus & Leiman, 1979)

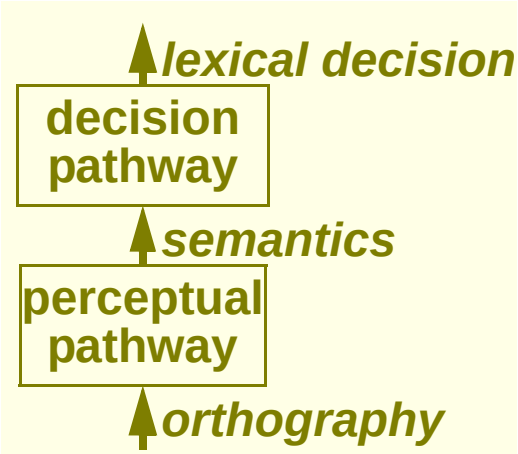
	congruent	*incongruent*	*unassociated*
context	LEAF	HAND	HAND
prime	PALM	PALM	RACE
target	TREE	TREE	TREE

Task: lexical decision to target

Modeling semantic priming

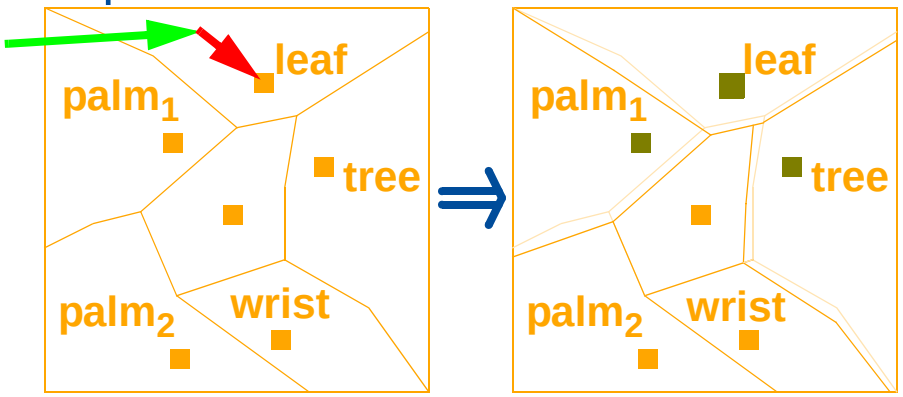
Two pathway architecture

Decision pathway responds “yes” if perceptual pathway obtains a well-formed state, “no” otherwise.

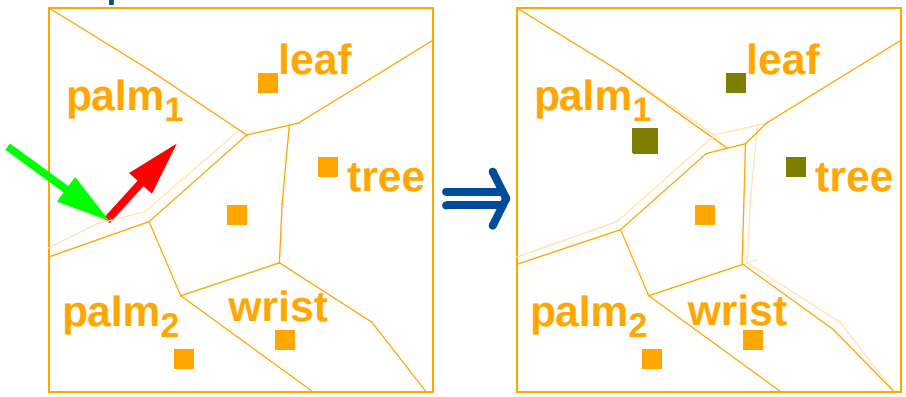


Simulation with extended presentation of prime

1. present LEAF



2. present PALM



After model has stabilized on PALM:

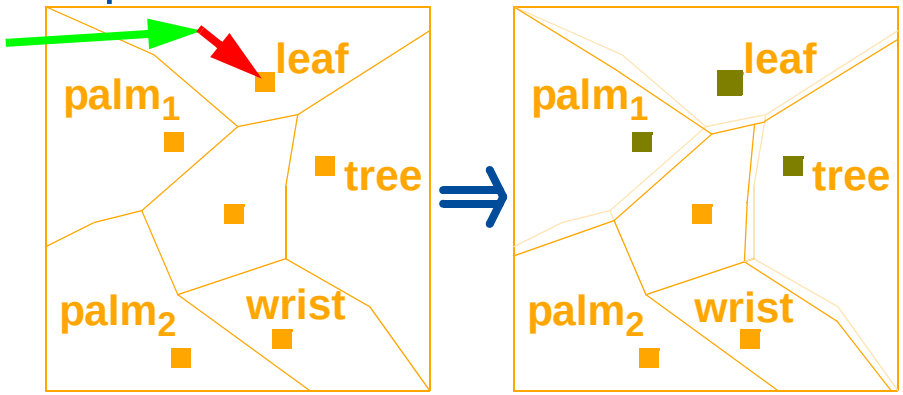
- one interpretation has been selected
- facilitation for targets congruent with that meaning
- awareness of prime

	congruent	***incongruent***	***unassociated***
human data (Marcel, 1980)	499 ms	547 ms	541 ms
simulation	224 cycles	252 cycles	257 cycles

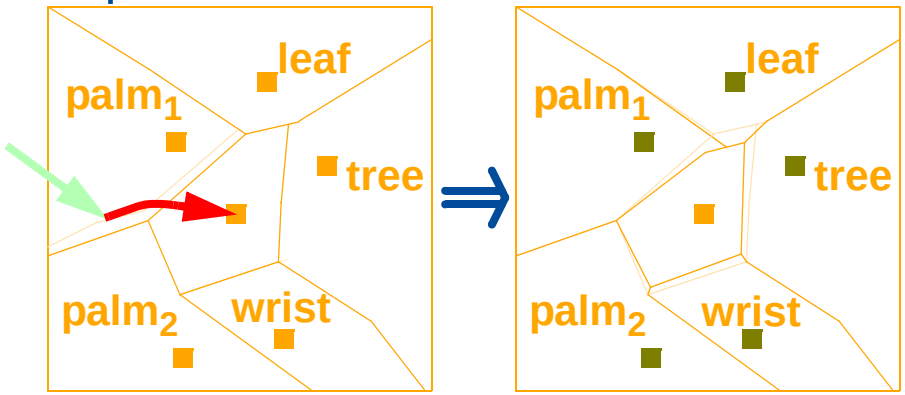
= facilitation

Simulation with brief presentation of prime

1. present LEAF



2. present PALM



Before model has settled on PALM:

- no interpretation has been selected
- facilitation for targets congruent with either meaning
- no awareness of prime

	congruent	***incongruent***	***unassociated***
human data (Marcel, 1980)	511 ms	520 ms	548 ms
simulation	238 cycles	242 cycles	254 cycles

Human data showed correlation between onset of awareness and selection.

Model suggests further that selection is a necessary precursor to awareness.

Primed Stem Completion

(Debner & Jacoby, 1994)

Procedure

Masked prime (e.g., TABLE)

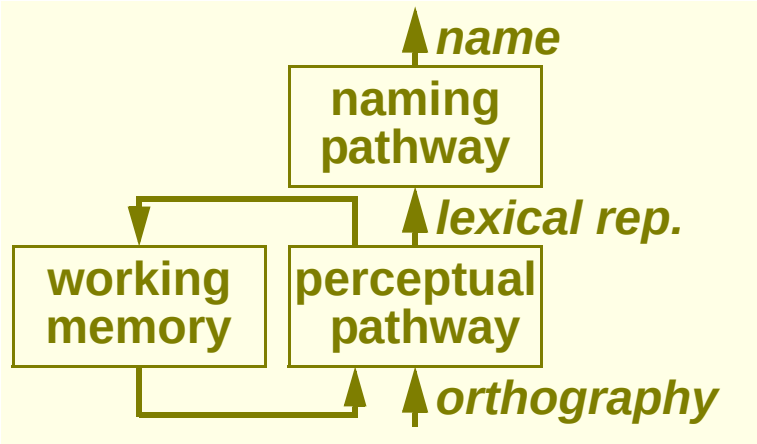
Followed by stem to be completed (e.g., TAB_ _)

Inclusion condition: Ss instructed to complete stem with prime if possible.
Exclusion condition: Ss instructed to complete stem with any word but prime.

Short (mostly unconscious) and long (mostly conscious) exposure conditions

Dependent measure: probability that stem completed with prime

Simulation



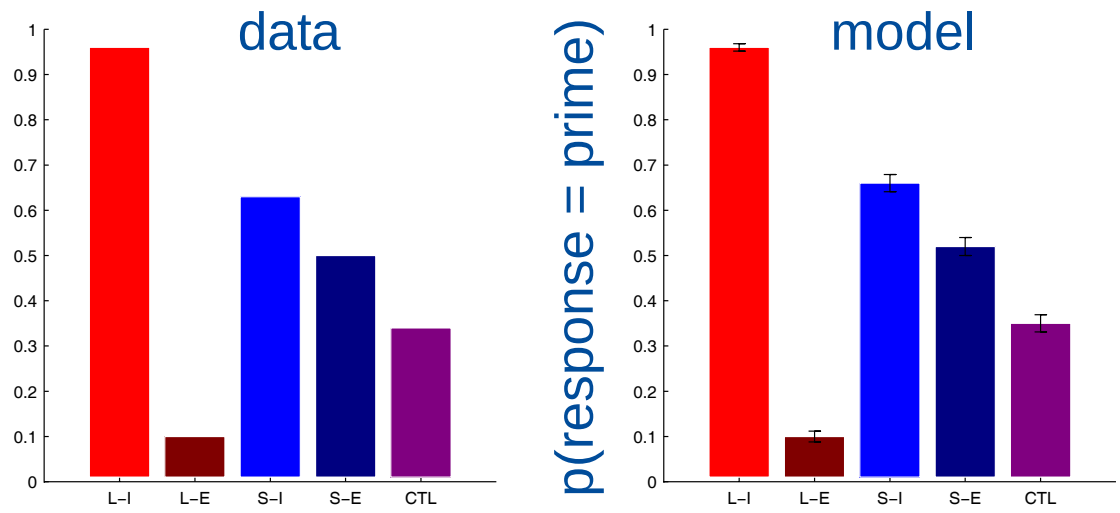
Presentation of prime (TABLE)

- Perceptual pathway produces lexical representation.
- Priming mechanism operates inside perceptual pathway.
- If input is sustained and perceptual pathway stabilizes, lexical representation copied to working memory.

Presentation of stem (TAB_ _)

- Stem representation: average of three similar orthographic representations
- Perceptual pathway finds best-fitting completion of visual input consistent with lexicon.
- Perceptual processing modulated by contents of working memory

Results



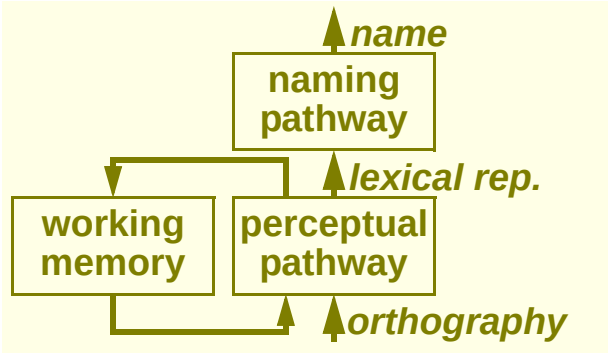
= inclusion
= exclusion
} long expos.

= inclusion
= exclusion
} short expos.

= control (no prime)

Debner and Jacoby analysis of data suggests a dissociation between conscious and unconscious processes.

No such distinction in our model: awareness is an emergent property of *representations*, not processes or pathways.



Semantic Priming and Response Windows

(Draine & Greenwald, 1996; Greenwald, Draine, & Abrams, 1996)

Masked prime word followed by target word

Semantic classification task: Is target pleasant/unpleasant?

Conditions

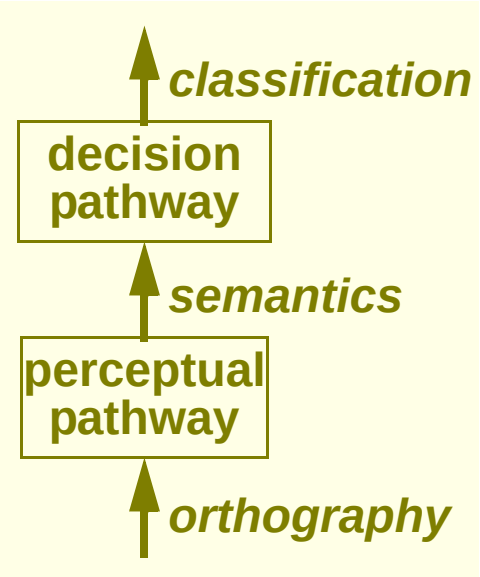
- *congruent* (FUN-PLAY) vs. *incongruent* (SICK-PLAY)
- *short* vs. *long* prime exposure ("subliminal" vs. "supraliminal")
- RT to target *constrained* to fall within temporal response window vs. *unconstrained*

Simulation

Decision pathway makes two-alternative forced-choice classification

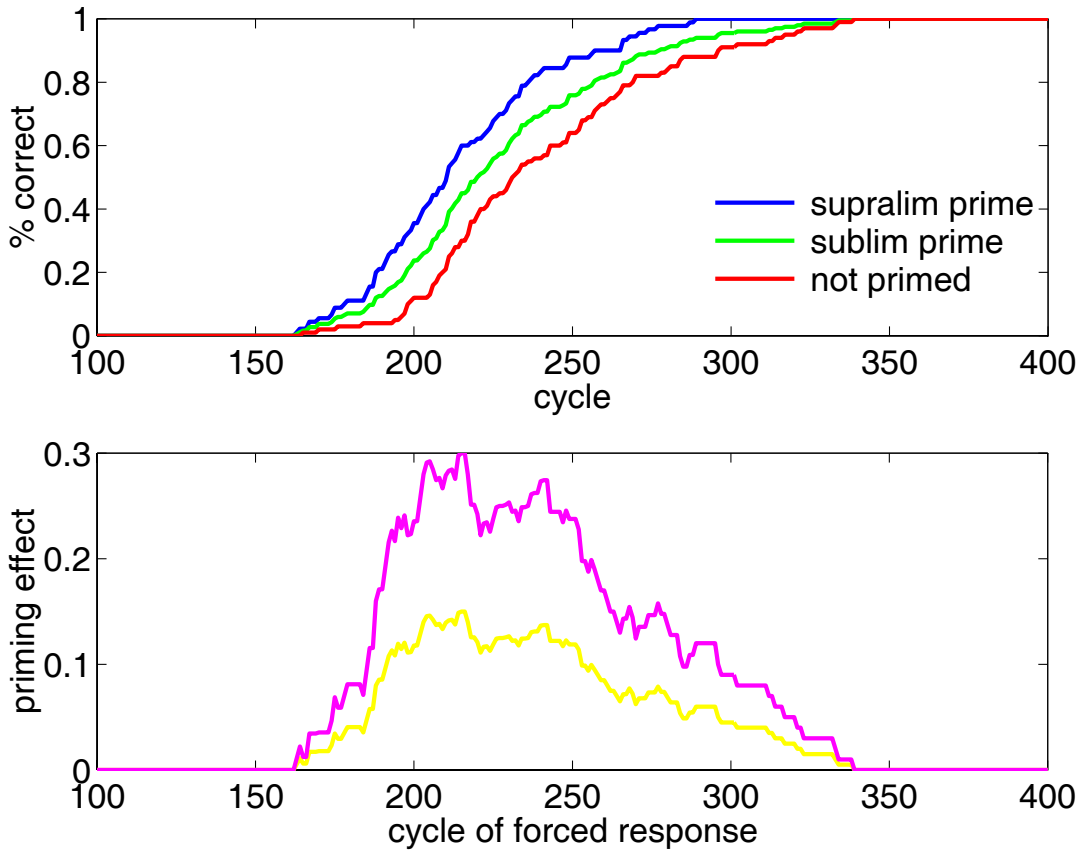
At brief exposure, prime is subliminal

Response window simulated by forcing decision pathway to closest attractor at time response is required.



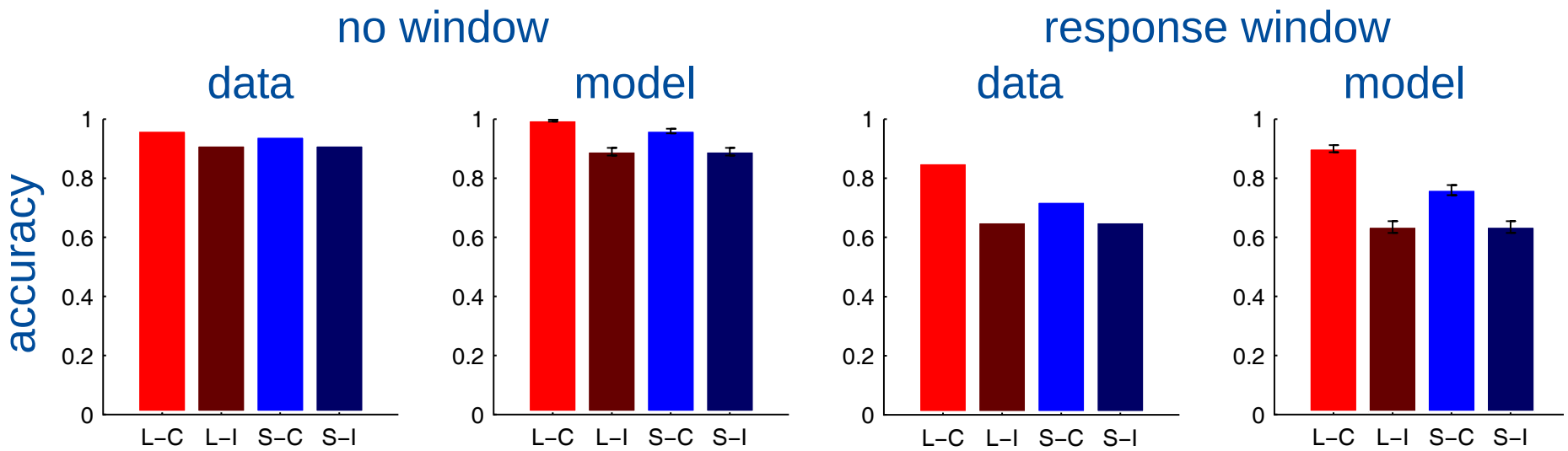
Prediction

Magnitude of priming effect is larger near center of speed-accuracy curve than near the extremes.



When a forced response is made at a time near the center of the curve, accuracy priming effects will be largest.

Results

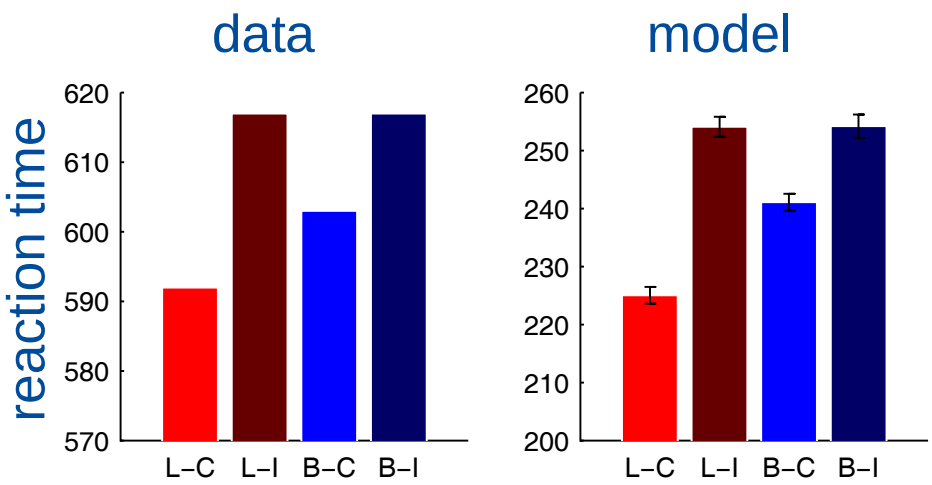


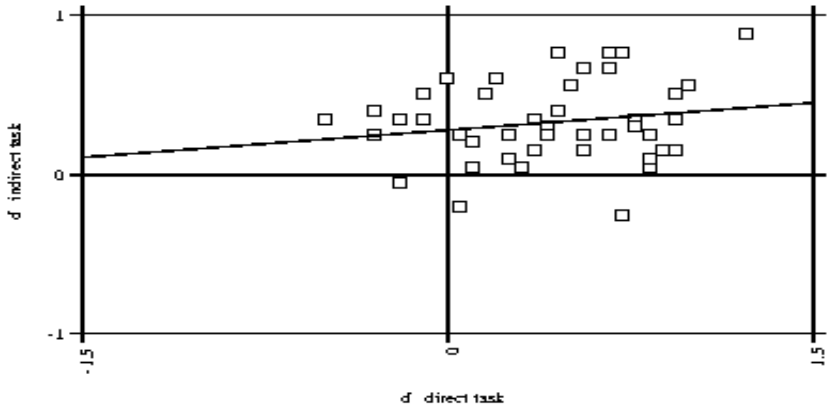
= long exposure, congruent

= long exposure, incongruent

= short exposure, congruent

= short exposure, incongruent



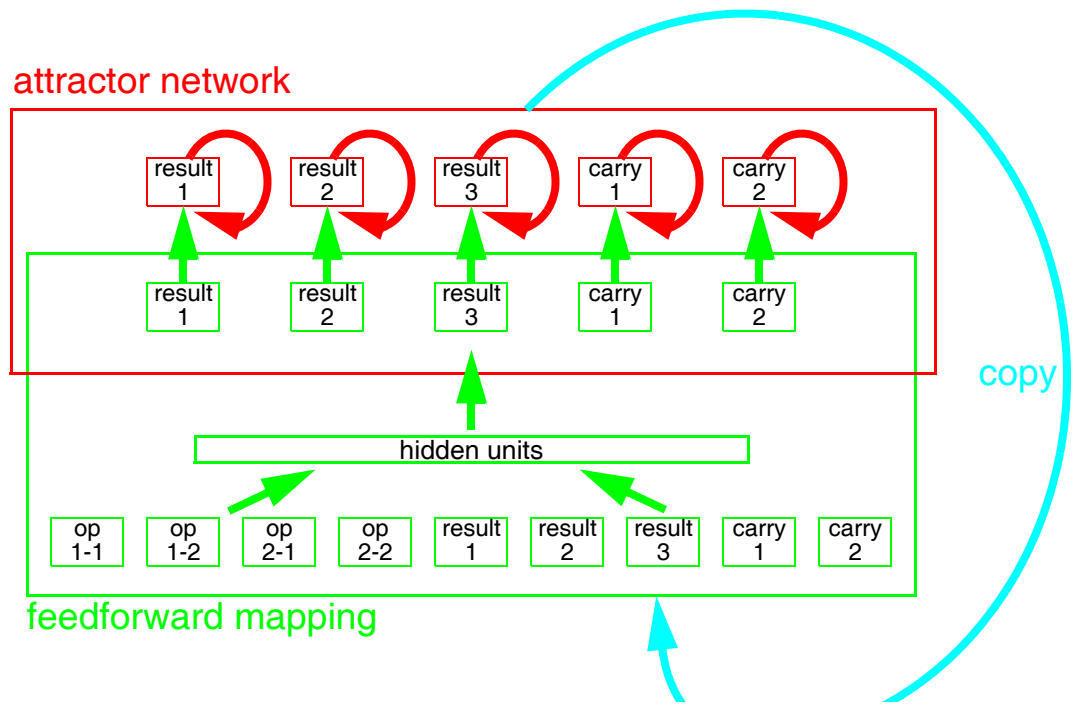


Automatization and Multistep Tasks

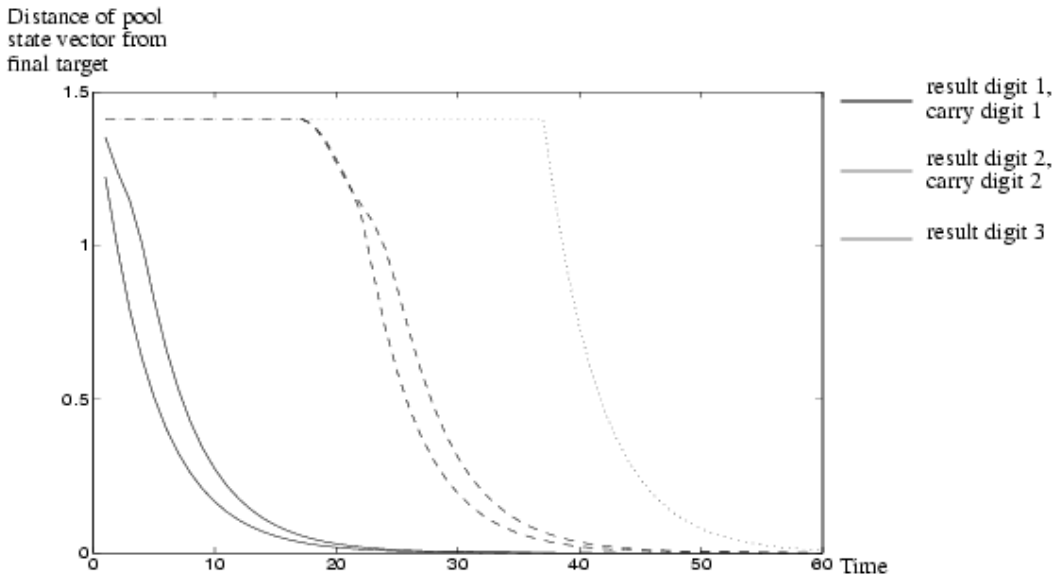
Trained network to add two 2-digit numbers in three steps

? ?		? 1		1 1		1 1
4 8	*step 1* →	4 8	*step 2* →	4 8	*step 3* →	4 8
+6 2		+6 2		+6 2		+6 2
? ? ?		? ? 0		? 1 0		1 1 0

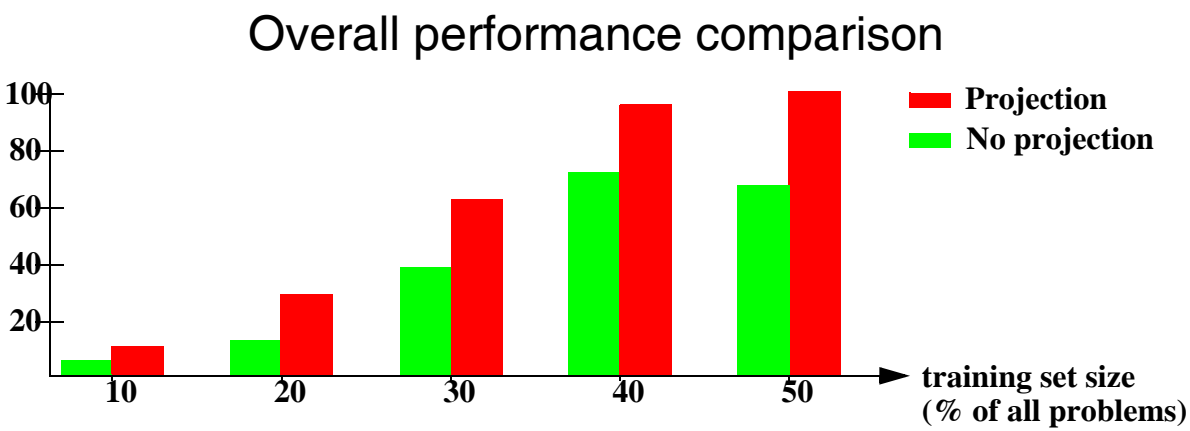
Each digit represented by pool of units.



Sequence of stable states reach awareness; correspond to *results* of computation

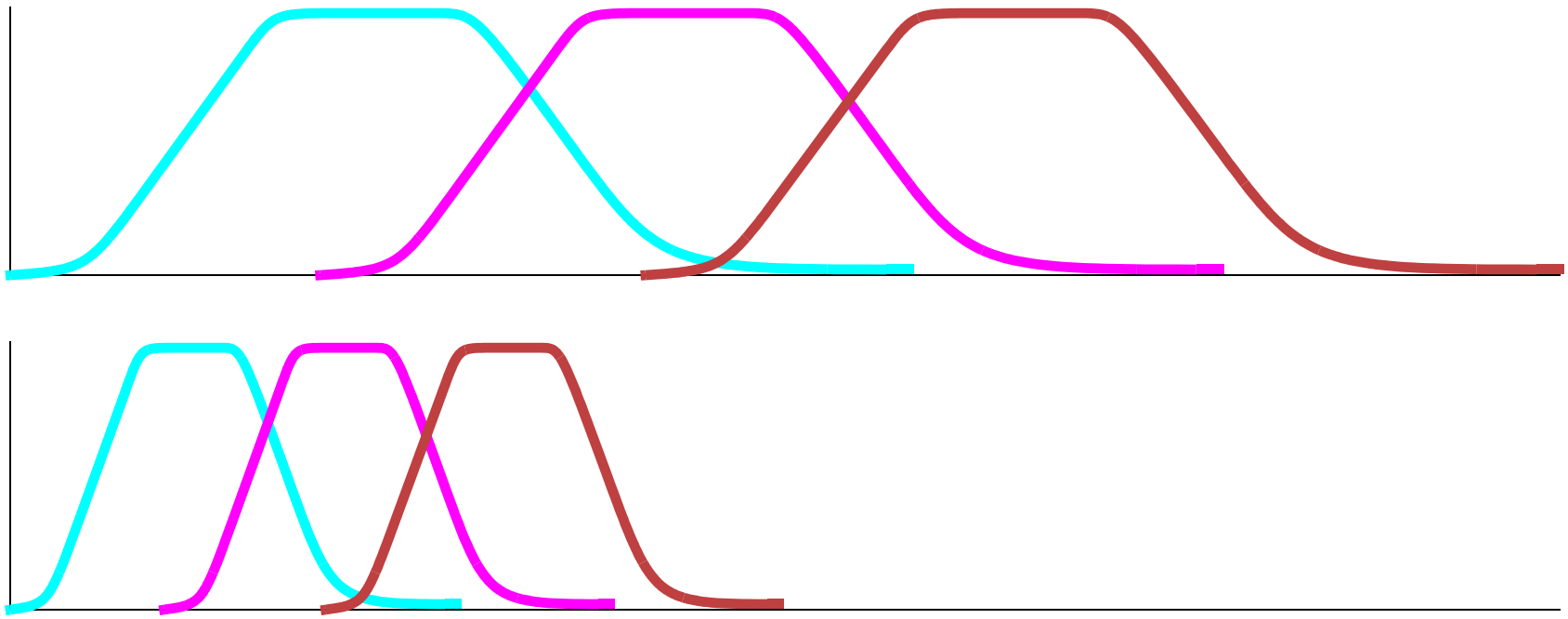


Benefit of attractor net clean up ("projection")



Prediction from the Theory: Automatization

Conscious awareness can drop out simply by speeding up processing; doesn't *require* a qualitative reorganization of knowledge



Might expect to find evidence for the application of intermediate steps, even without awareness.

E.g., Crutcher (1992) `doronico – door – leopard`

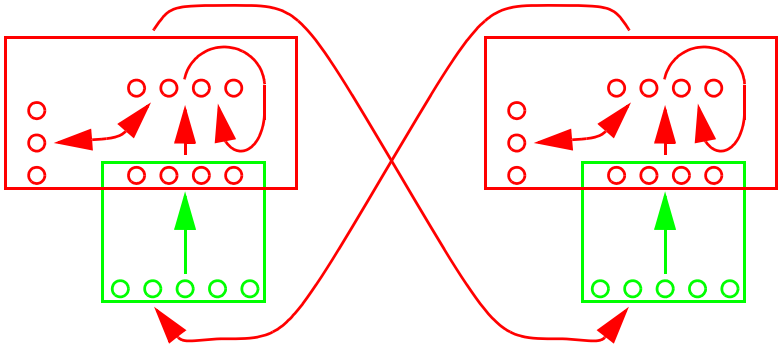
Explaining Unitary Nature of Consciousness

Often claimed that people can only be aware of one thing at a time.

According to framework, strictly speaking this is false.

Within the output domain of a pathway, a single, coherent representation is selected.

Across pathways, no explicit competition; e.g., stable patterns can be achieved in somatosensory and visual modules simultaneously. But strongly linked domains require mutual consistency.



The subjective impression of unity also arises due to the equating of awareness with verbalization. If we assume there is a module responsible for the verbal expression of ideas, then it is limited to one expression at a time.