Modeling sequence classes sequentially

Rick Dale & Michael J. Spivey
Cornell University

Time

- Structure of behavior in time
  - Computational modeling; small-scale
  - Temporal analytic technique; "ecological"-scale

Sequence Classes (SC)

- Ordinal relations among stimuli
  - Train: A1 ⇒ A2 ⇒ A3 ⇒ A4
  - Substitutability (e.g., A1 ⇒ B2)
  - Transitivity (e.g., A1 ⇒ A4)
  - e.g., Lazar, 1977; Green et al., 1993
- Syntax
- Behavioral "productivity" in response sequences

Elman (1995)

"Time is the medium in which all our behaviors unfold; it is the context within which we understand the world. We recognize *causality* because causes precede effects ... Time's arrow is such a central feature of our world that it is easy to think that, having acknowledged its pervasive presence, little more needs to be said."

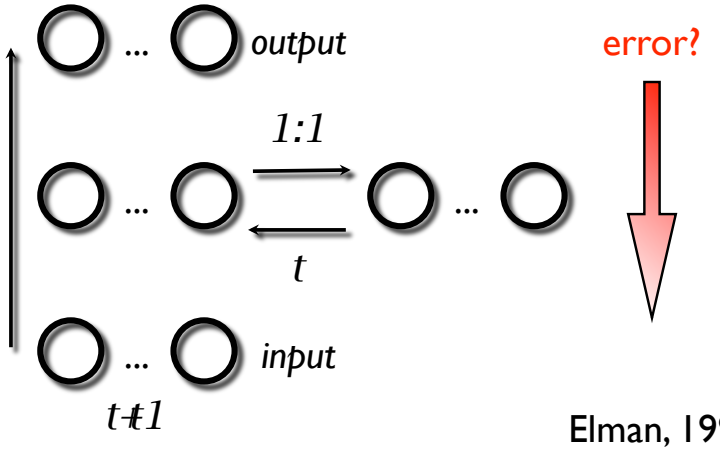
Barnes & Hampson (1993)

- RELNET: sequence and stimulus classes
- Feed-forward network
- Network learns to learn
- Transfer from stimulus to sequence classes
  - Barnes & Hampson, 1997; Cullinan et al., 1994; Lyddy et al., 2001
- "Simultaneous" performance: No time

Outline

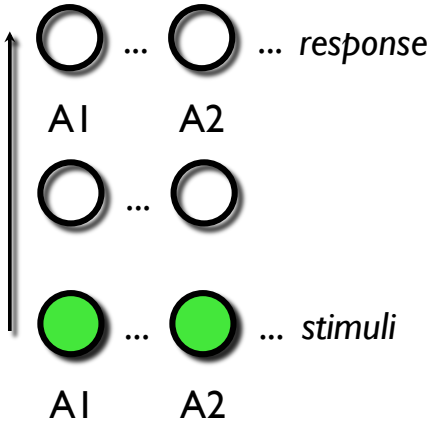
- Goal: Model with natural time properties
- Simulation 1: Substitutability
- Simulation 2: Transitivity + Substitutability
- Theoretical considerations

Simple-Recurrent Network

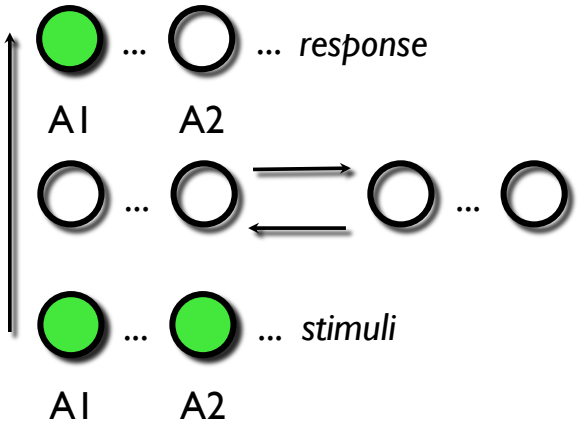


Elman, 1990

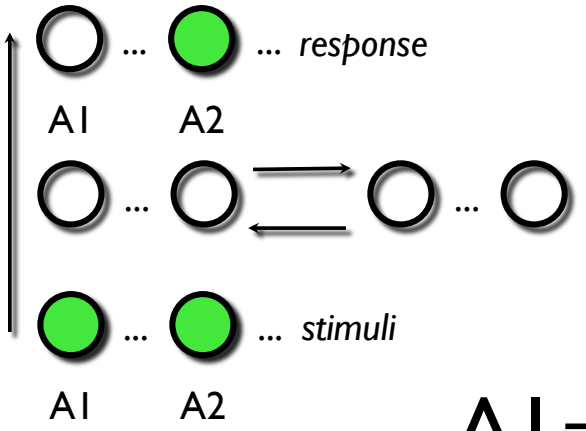
Sequence Classes in an SRN



Sequence Classes in an SRN



Sequence Classes in an SRN



A1⇒A2

Simulation 1: Substitutability

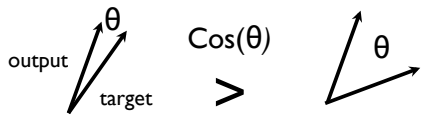
- Galy et al., 2001
- Train:
  - A1⇒A2⇒A3; B1⇒B2⇒B3; C1⇒C2⇒C3
- Untrained:
  - A1⇒B2⇒C3

Simulation 1: SRN

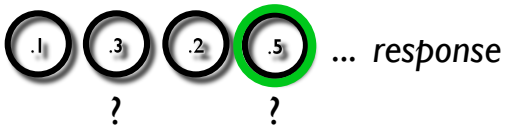
- Network training:
  - A1⇒A2⇒A3; B1⇒B2⇒B3; C1⇒C2⇒C3
- Network test:
  - A1⇒B2⇒C3; B1⇒A2⇒C3; C1⇒B2⇒A3

Simulation 1: Evaluation

- **Cosine** between target response and network output activation

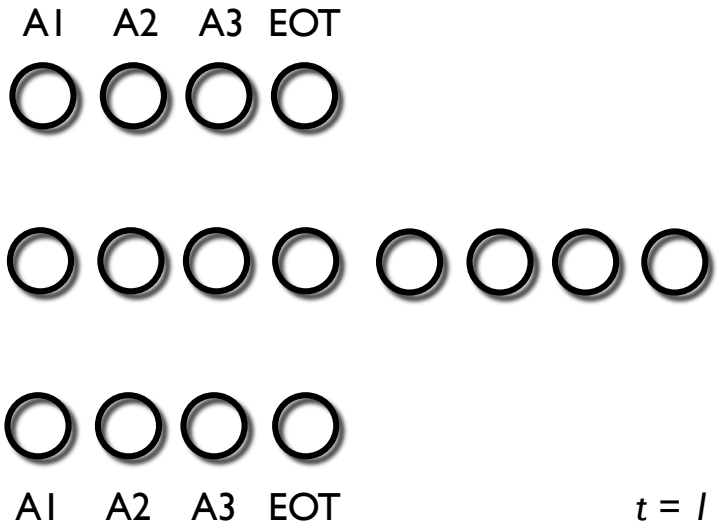


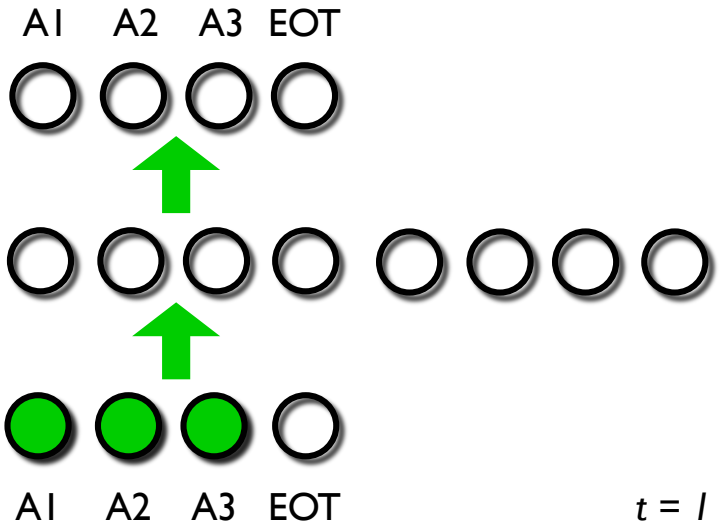
- **Response-selection** from output

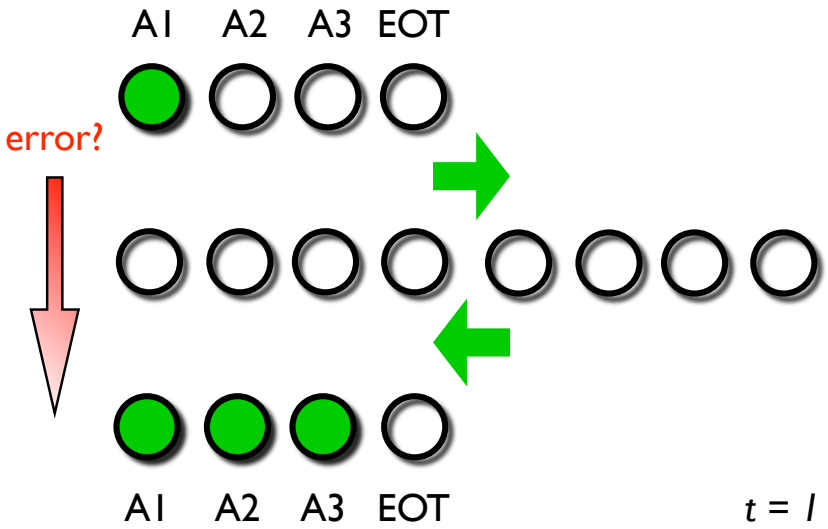


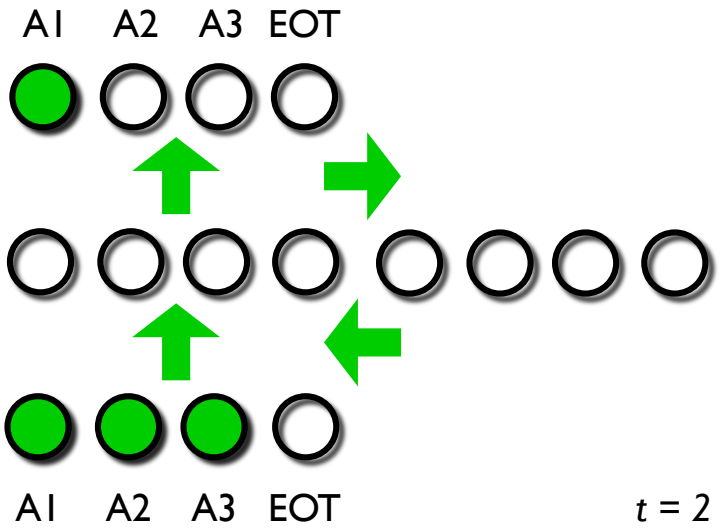
Simulation 1: Training Details

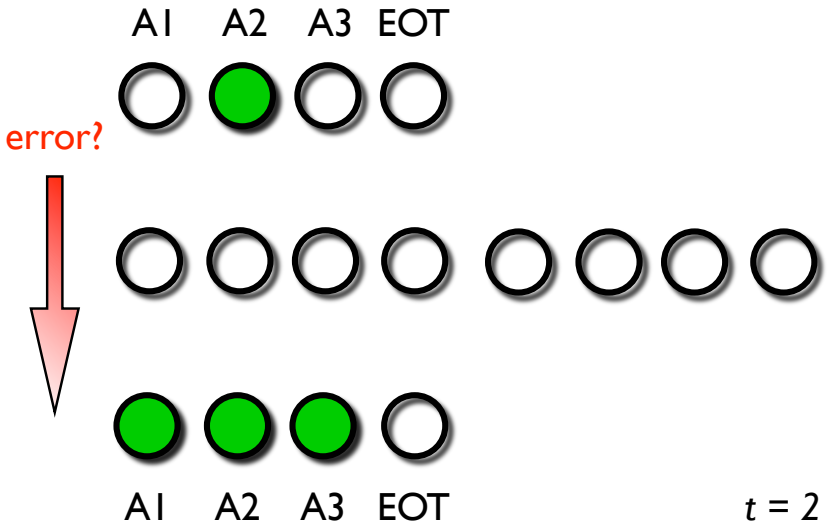
- 19 inputs/outputs
  - 3 (A/B/C) x 3 (1/2/3) x 2 (two sets)
  - 1 end-of-trial unit
- Train to criterion: 20 correct trials
- Hidden units: 20; Learning rate: .25
- 4 random network + 4 control nets
- out = sig(net + net·ε/10) (e.g., Dell, 1986)

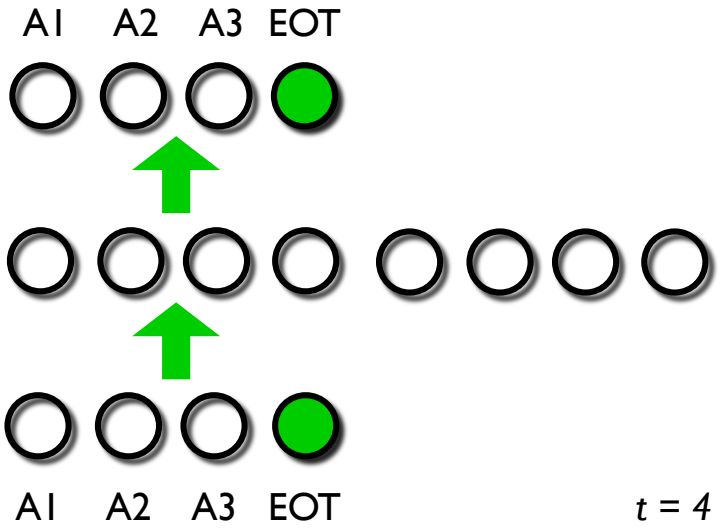




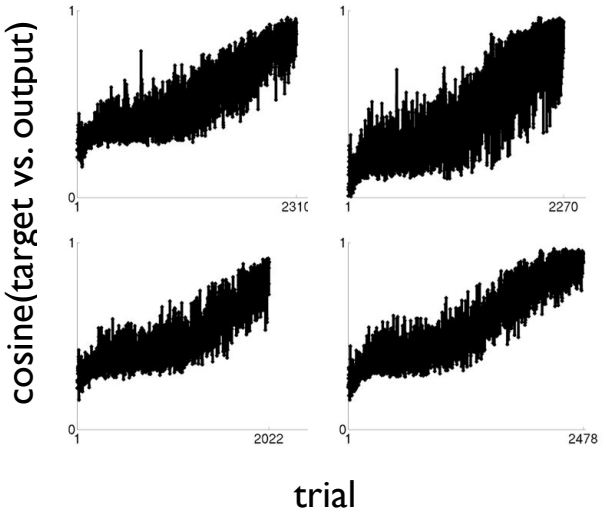

A1 A2 A3 EOT
error?
A1 A2 A3 EOT
t = 1


A1 A2 A3 EOT
A1 A2 A3 EOT
t = 2

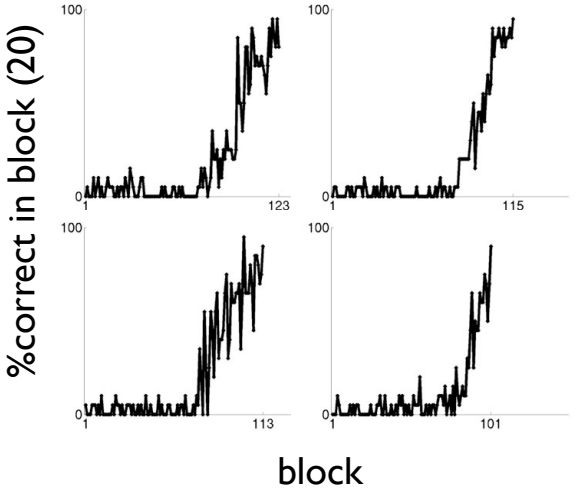

A1 A2 A3 EOT
error?
A1 A2 A3 EOT
t = 2


A1 A2 A3 EOT
A1 A2 A3 EOT
t = 4

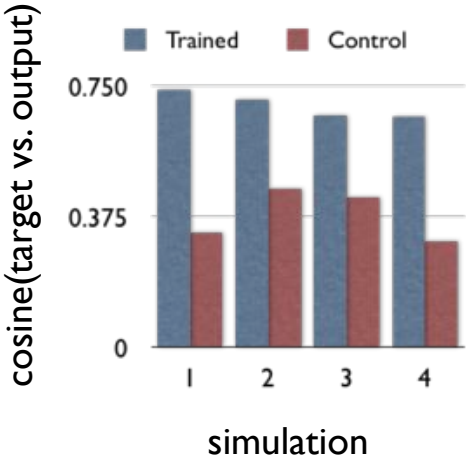
Results: Training



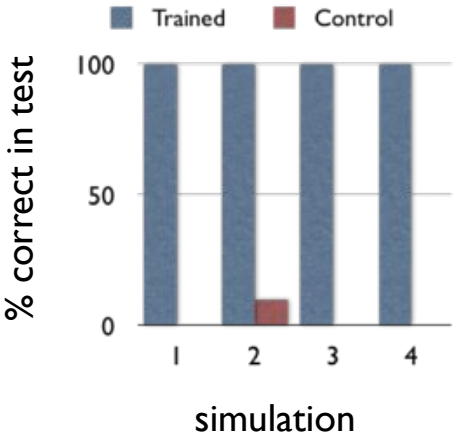
Results: Training



Results: Test



Results: Test



Summary

- SRN models substitutability in 3-term sequence classes
- Dependent upon establishment of 3-term sequence classes
- Sequence classes sequential

Simulation 2: Transitivity Too

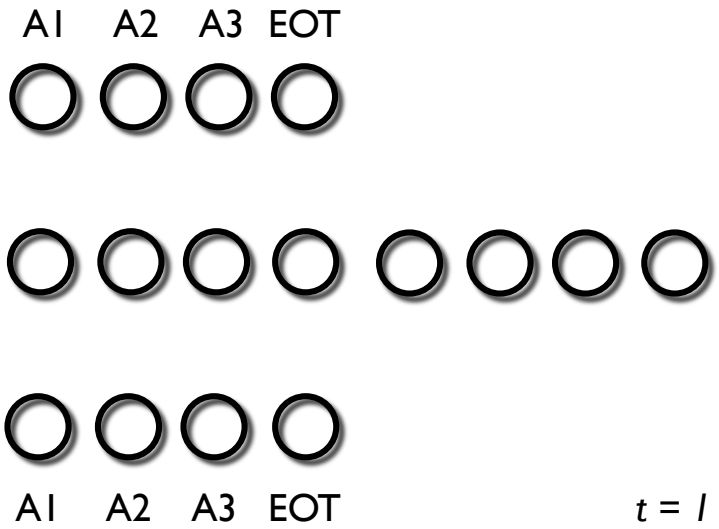
- Green et al., 1993
- Train:
  - X1⇒X2⇒X3⇒X4⇒X5
  - 3 classes (A/B/C)
- Untrained:
  - A1⇒A3⇒A5 (within-class)
  - A1⇒B4 (across-class)

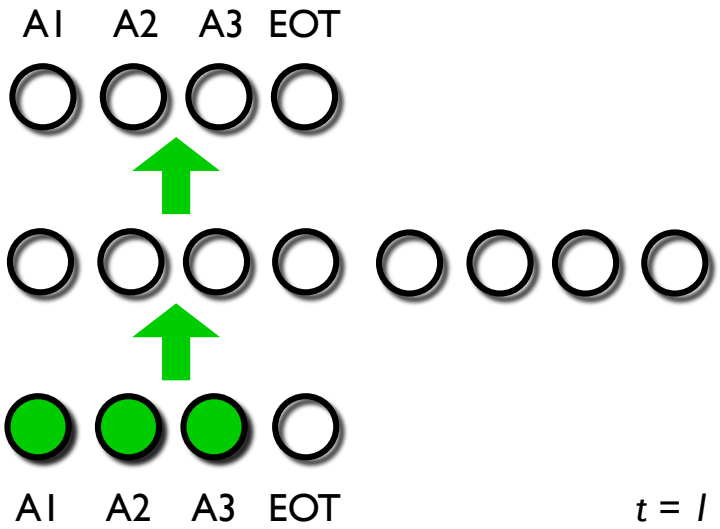
Simulation 2: SRN

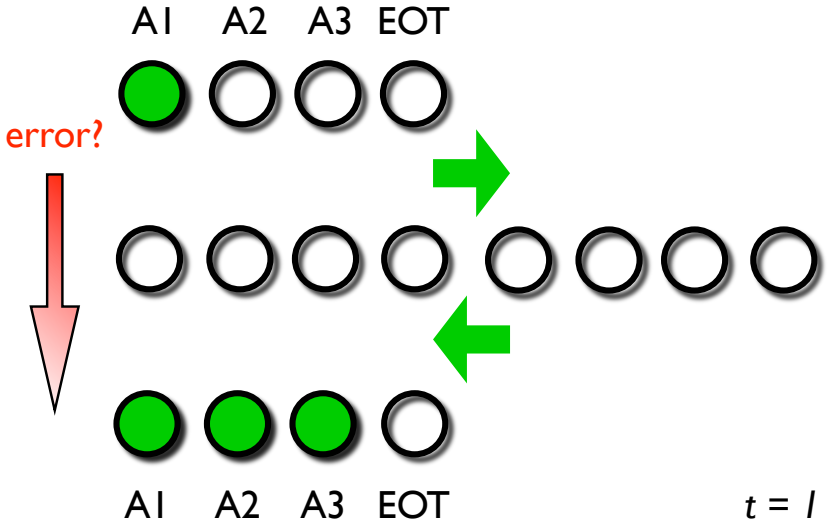
- "Faded" network training:
  - (X1⇒X2), (X3⇒X4⇒X5)
  - 3 classes (A/B/C)
- Network test:
  - A1⇒A/B/C3⇒A/B/C5

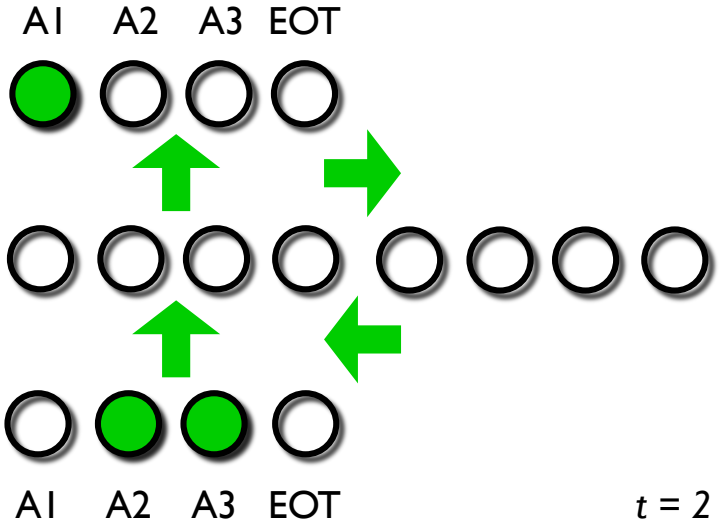
Simulation 2: Training Details

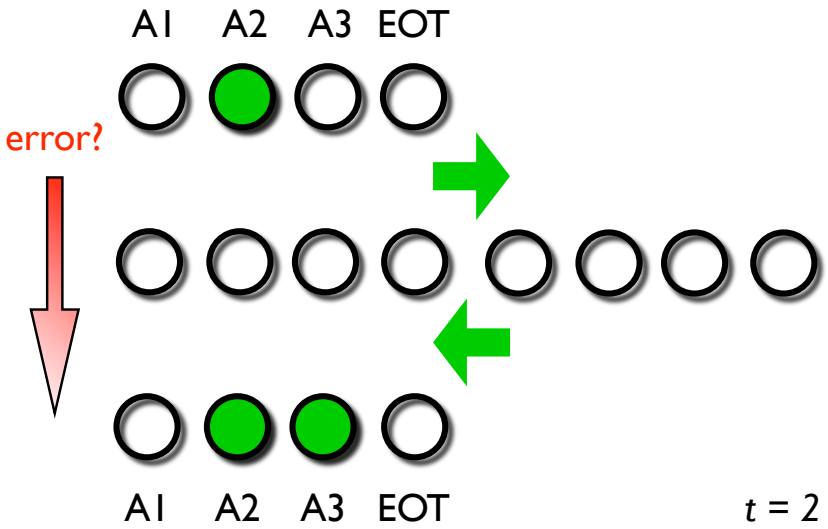
- 31 inputs/outputs
  - 3 (A/B/C) x 5 (1/2/3/4/5) x 2 (two sets)
  - 1 end-of-trial unit
- Train to criterion: 20 correct trials
- Hidden units: 40; Learning rate: .25
- 4 random network + 4 control nets
- out = sig(net + net·ε/10) (e.g., Dell, 1986)

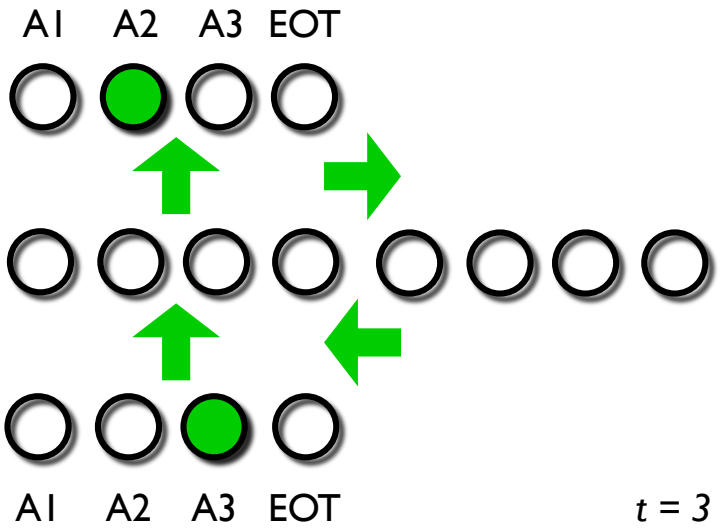

A1 A2 A3 EOT
A1 A2 A3 EOT
t = 1

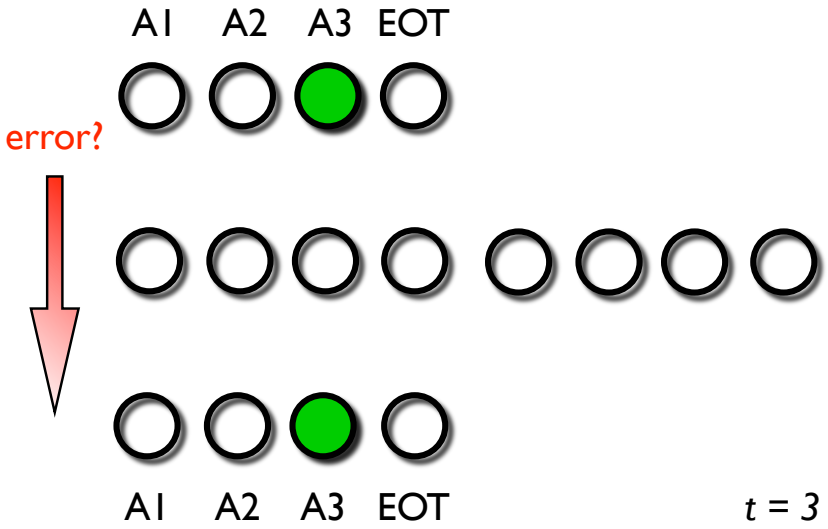

A1 A2 A3 EOT
A1 A2 A3 EOT
t = 1

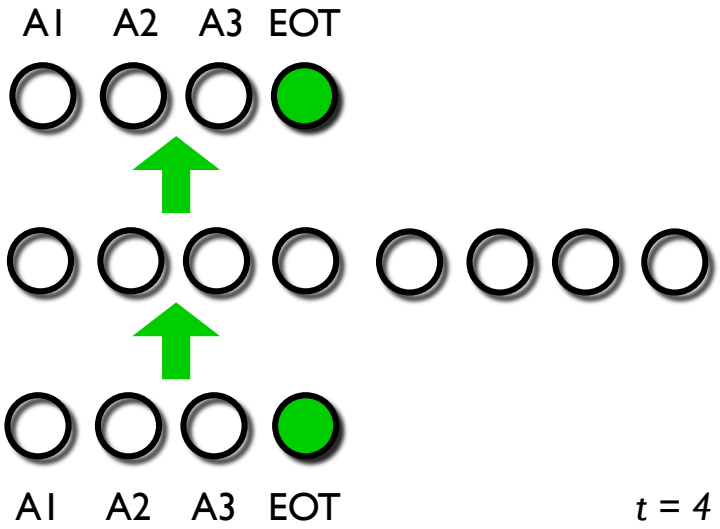

error?
A1 A2 A3 EOT
A1 A2 A3 EOT
t = 1


A1 A2 A3 EOT
A1 A2 A3 EOT
t = 2


A1 A2 A3 EOT
error?
A1 A2 A3 EOT
t = 2


A1 A2 A3 EOT
A1 A2 A3 EOT
t = 3

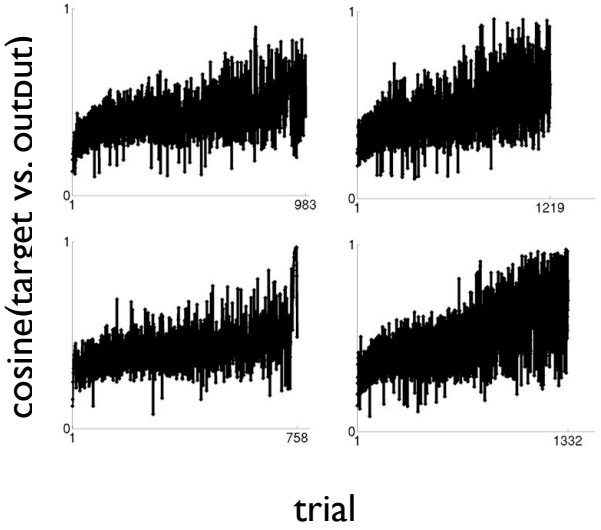

A1 A2 A3 EOT
error?
A1 A2 A3 EOT
t = 3


A1 A2 A3 EOT
A1 A2 A3 EOT
t = 4

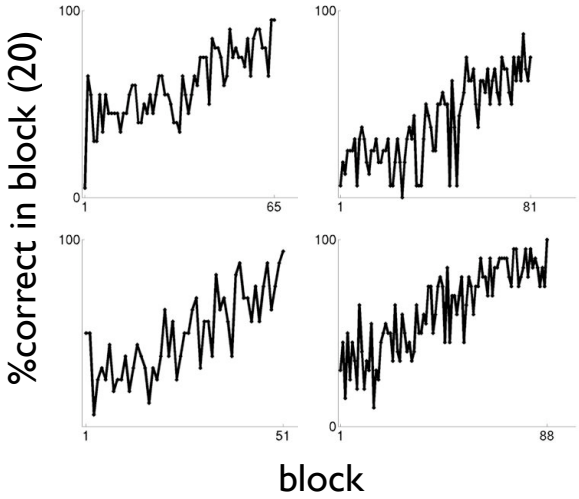
Test Evaluation

- Response-selection: Network *chooses its own* response
  - Select among highly active available responses
- This response/stimulus is removed from array at next timestep

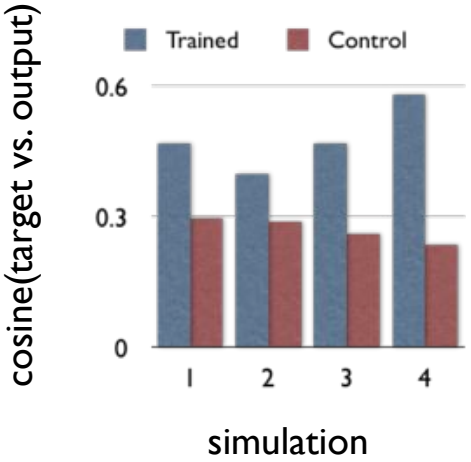
Results: Training



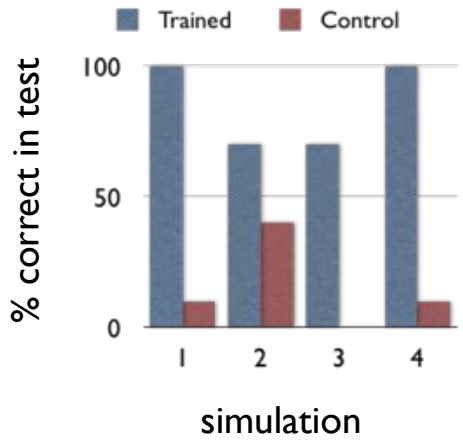
Results: Training



Results: Test



Results: Test



Summary

- SRN does transitivity in sequence classes
- "Task validity" in sequential model (Christiansen & Chater, 2001)
- More difficult task

Why?

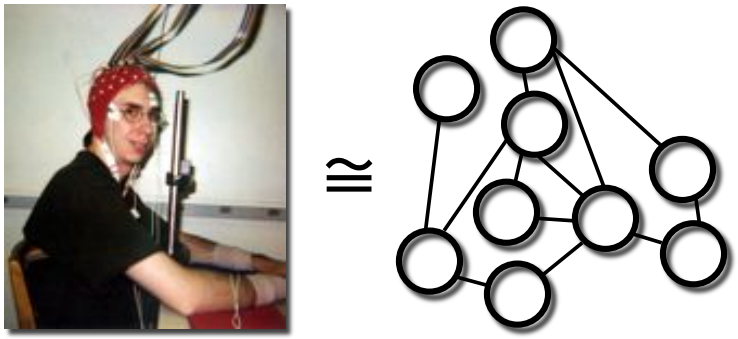
"As long as there are impredicativities sitting around in here you can not run an algorithm, so you cannot simulate them in the sense that we mean by simulation today."

Turvey

"But can your network grow a chicken wing?"

Chomsky

Why?



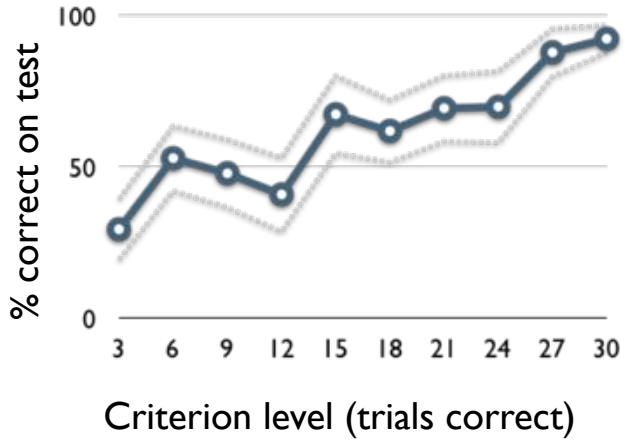
Why?

- Model exhibits proposed properties of given behavior and/or hypothetical states (e.g., "representations")
- Rigorous exploration of systems that satisfy these properties
- Model as quantitative tool
  - SRN "psychophysical" exploration of behavior in time

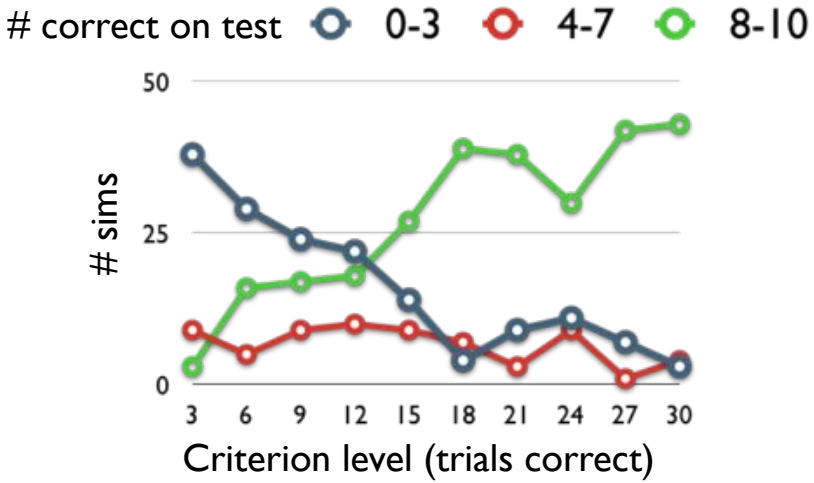
Parameter Exploration

- Explore relationships:
  - *Test = f(Training criterion)*
  - *Test = f(Sequence length)*
  - *Test = f(Number of sets)*
- SRN as function approximator, embodying exemplar-based statistical learner; discrete dynamical system in *h*-dimensional space

Test = *f*(Criterion)



Test = *f*(Criterion)



Thanks