Benchmarks in category learning

Andy J. Wills

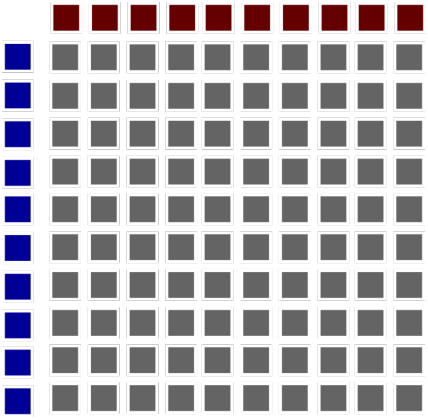
Plymouth University, UK

Why build a model?

- Unambiguous specification
  - "Theory X can accommodate phenomenon Y"
- But...
  - Many models, narrow comparisons.

Wills & Pothos (2012, Psychological Bulletin)

1. Scale



Wills et al. (2017, Psychology of Learning and Motivation)

2. Model availability

- Implementing from scratch
- Requesting from the author

Wills et al. (2017, Psychology of Learning and Motivation)

catlearn

- Package for R
  - Modular, resuable, fast (C++ for speed critical operations)
  - Documented; free and open-source; stable release every 6 months
  - 9 models (ALCOVE, Bush-Mosteller, COVIS, DIVA, EXIT, GCM, MB/MF, Rescorla-Wagner, SUSTAIN), more on the way.
  - Archival
  - >11k downloads (as of 2019-06-22)
- Helping simulation at scale with simulation archiving
- General schema applicable to other modeling domains.

Wills et al. (2017, Psychology of Learning and Motivation)

3. Benchmarks

- Common models implementations $\rightarrow$ common datasets
- This has been hard!
- Interesting discussions with:
  - Charlotte Edmunds, Rob Goldstone, Ken Kurtz, Mark Johansen, Brad Love, Todd Maddox, Doug Medin, Fraser Milton, Greg Murphy, Rob Nosofsky, Emmanuel Pothos, Bob Rehder, Vladimir Sloutsky
- Little agreement!

Issues with category-learning benchmarks

- Lack of overlap
- Lack of agreement on thresholds (independent replication?)
- Lack of agreement with “elevation” principle.
- Difficulty in communicating a bleeding edge vs. settled result distinction
- But, it can and has been done elsewhere . . .

Working memory benchmarks

- 15 working memory researchers
- A few meetings
- Comprehensive set of benchmarks

Oberauer et al. (2018, Psychological Bulletin)

How to make progress

- Paradigm coherence in working memory?
- Subdivide by assumed broad process?
  - Attentional reallocation
  - Averaging abstractions
- Consideration of broader goals?

Le Pelley et al. (2016, Psychological Bulletin)

Goals: Example

- Adequate formal model of X (in my case, X = learning object categories)
- But what broader goals does that serve?
  - Intrinsic / curiosity motivation ("crossword" motivation)
  - In my case: Understand learning object categories well enough to build a machine to do it.

A solved problem?

- GoogLeNet
- If it's a solved problem, then (as per Olivia's talk) we should be working out how they solve it, because either ...
  - Human-like $\rightarrow$ repurpose as model of cognition
  - Non-human-like $\rightarrow$ implications for their interaction with humans

I'm skeptical it's solved...

1. Kettle
2. Spaceship
3. Orange
4. Screwdriver
5. Rabbit
6. Elephant
7. Mountain
8. Telephone

I'm skeptical it's solved...

- Naive tests of human performance
- Generalization
  - Success on location, orientiation
  - What about occlusion?

- Learning speed?
- Catastrophic interference?

One way forward . . .

- Develop "learning object category" benchmarks
  - Richer / more naturalistic sitmuli
  - More direct comparison of accuracy, and speed of learning between human behavior and the machines.