Optimal teaching for Limited-capacity Human Learners

Kaustubh R Patil[1,3] Xiaojin Zhu[2] Łukasz Kopeć[1] and Bradley C Love[1]

kaustubh.patil@gmail.com, jerryzhu@cs.wisc.edu, l.kopec.12@ucl.ac.uk, b.love@ucl.ac.uk

[1] Department of Experimental Psychology, University College London, UK [2] Department of Computer Sciences, University of Wisconsin-Madison, USA
[3] MIT Sloan Neuroeconomics laboratory, MIT, USA

EXPERIMENTAL PSYCHOLOGY

WISCONSIN UNIVERSITY OF WISCONSIN–MADISON

Neural Information Processing Systems Foundation

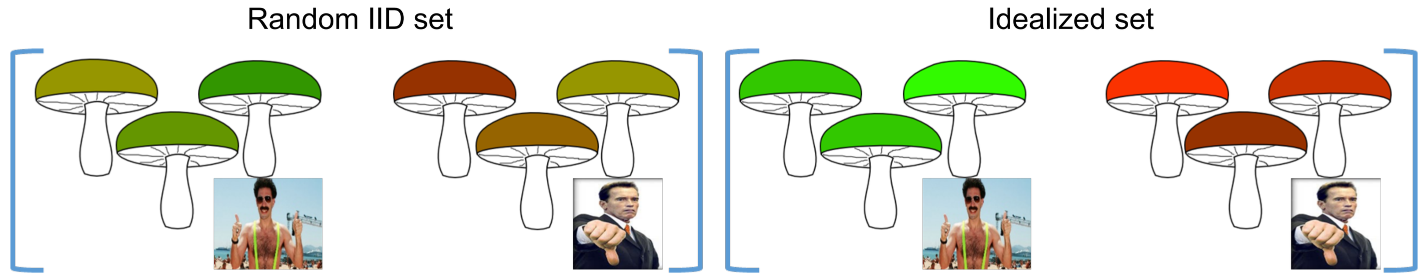
Montreal 2014

Human category learning

- Basic decisions, such as identifying a mushroom as edible or not, involve categorizing novel examples.
- How can we efficiently teach humans to make such decisions?
- Which will be the best examples if we want a small training set?

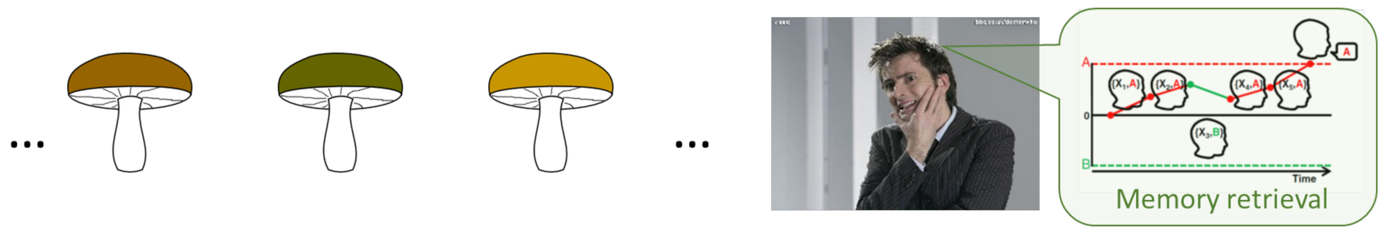
Categorization toy example

Training



While learning specific exemplars are stored in memory.

Test



At decision time, stored examples are retrieved and a decision is made based on their similarity to the novel example.

Capacity limitations [1]

- The memory retrieval process is stochastic and noisy.
- Wrongly retrieved examples will lead to wrong results.
- Idealization of the training set leads to cleaner memory retrieval and hence better decisions.

Our contribution

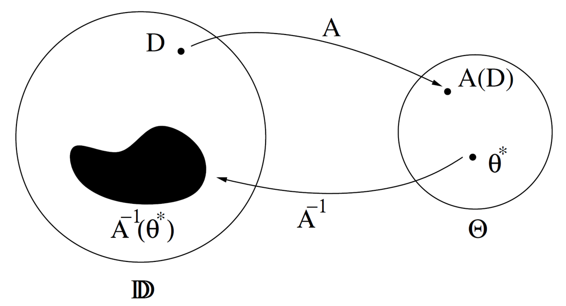
- Previous work has proposed idealization of training sets.
- But no systematic approach to find good training sets.

We propose a systematic approach to find an optimal set for training humans on probabilistic categorization tasks.

Optimal teaching solution

Optimal teaching [2,3]

- Aims to find a training set that best teaches a hypothesis to a given learner.
- Works to the extent to which the model of the learner complies with the true learning mechanism.



$$\min_{D \in \mathbb{D}} \ \mathrm{loss}(D) + \mathrm{effort}(D).$$

$$\mathrm{loss}(D) = \mathbb{E}_{(x,y)\sim p(x,y)} \mathbb{E}_{\hat{y}\sim \hat{p}(y|x,D)} \mathbb{1}_{y \neq \hat{y}}.$$

Optimal teaching for GCM

The Generalized Context Model [4]

$$\hat{p}(y=1 \mid x, D) = \frac{\left(b + \sum_{i\in D: y_i=1} e^{-c\, d(x,x_i)}\right)^{\gamma}}{\left(b + \sum_{i\in D: y_i=1} e^{-c\, d(x,x_i)}\right)^{\gamma} + \left(b + \sum_{i\in D: y_i=-1} e^{-c\, d(x,x_i)}\right)^{\gamma}}$$

Optimal teacher for GCM

$$\min_{x_1 \ldots x_n \in [0,1]} \frac{1}{m} \sum_{j=1}^{m} \left(\frac{1 - 2p(y=1 \mid z_j)}{1 + \left(\frac{b + \sum_{i: x_i < 0.5} e^{-c\, d(z_j, x_i)}}{b + \sum_{i: x_i \geq 0.5} e^{-c\, d(z_j, x_i)}} \right)^{\gamma}} + p(y=1 \mid z_j) \right).$$

We found the capacity parameter γ by fitting GCM to previous data on a similar task from [1].

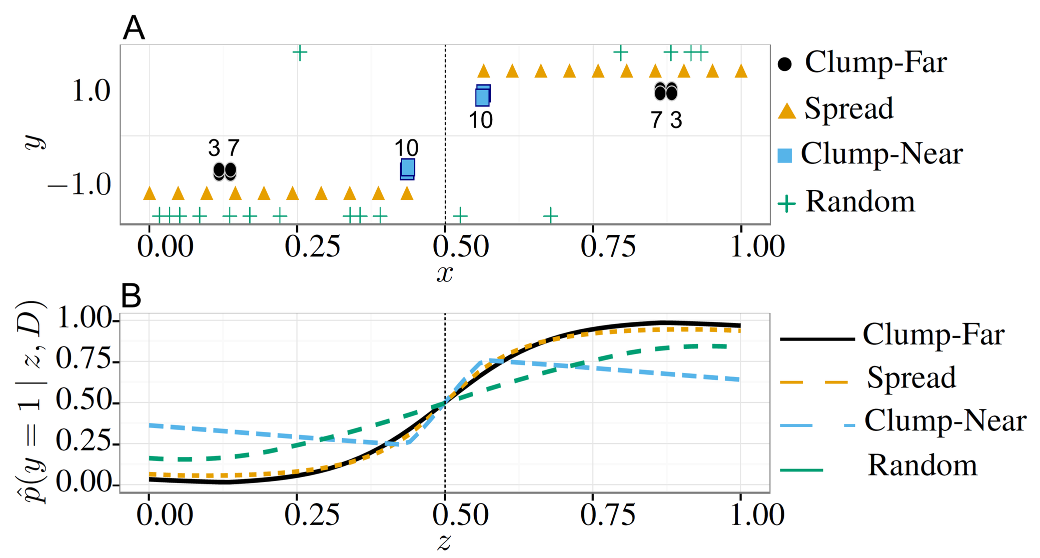
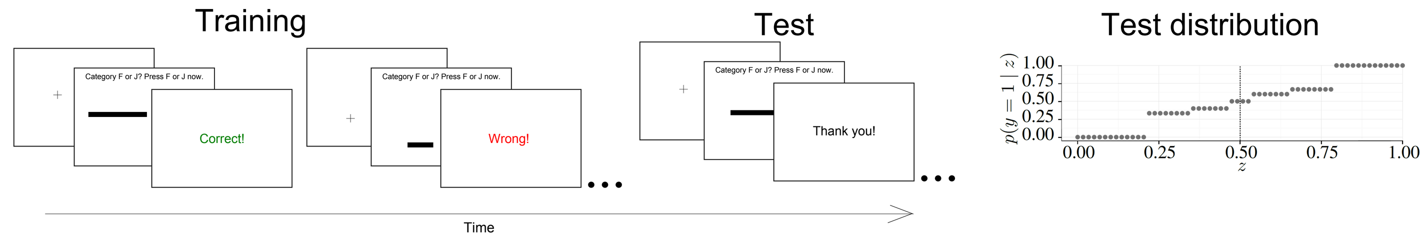


Figure: (A) Different training sets and (B) their estimated performance.

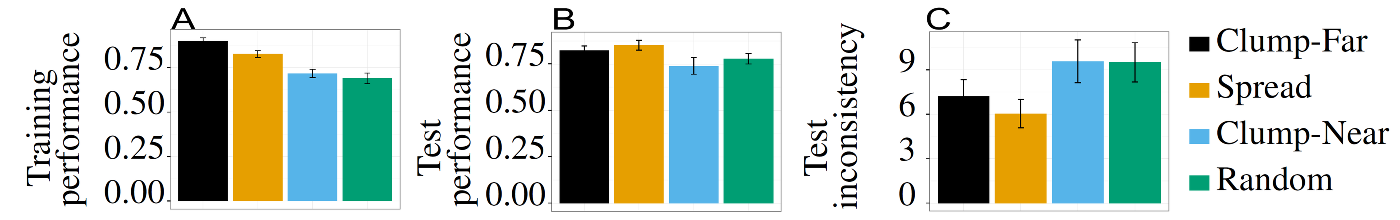
The low-capacity optimal training set 'Clump-Far' is idealized!

Human experiments

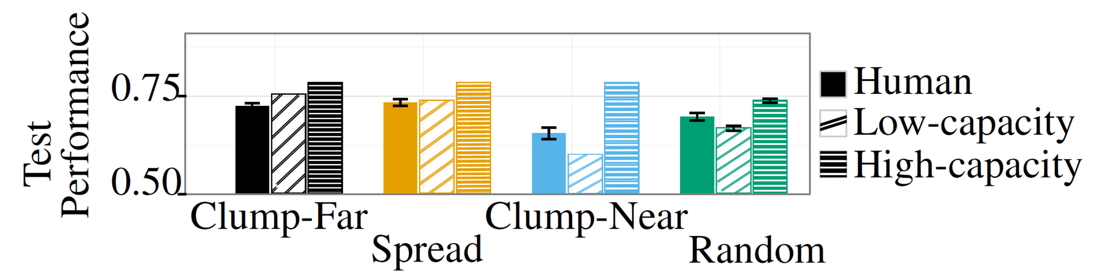
The task



Results and conclusions



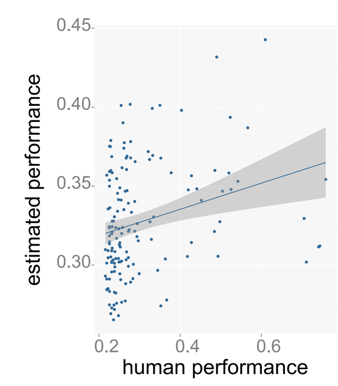
Humans perform better on the optimal Clump-Far set both in terms of lower loss and lower inconsistency.



The human performance pattern closely follows the low-capacity model.

This adds further evidence to the low-capacity account of human categorization.

Estimated and actual human performance on the IID Random sets are positively correlated Pearson's $r=0.273$, $p<0.05$.



Humans benefit from lower variability in the training set, i.e. idealization.

References

[1] G Giguere and B C Love. Limits in decision making arise from limits in memory retrieval, PNAS, 110(19):7613–8, 2013.
[2] S A Goldman and M J Kearns. On the Complexity of Teaching. J Comput Syst Sci, 50(1):20–31, 1995.
[3] X Zhu. Machine Teaching: an Inverse Problem to Machine Learning and an Approach Toward Optimal Education. In AAAI-15, 2015.
[4] R M Nosofsky. Attention, similarity, and the identification-categorization relationship, J Exp Psychol. General, 115(1):39–61, 1986.

People perform better when the optimal teacher uses a model of the learner that incorporates capacity limitations!