Populating ACT-R's Declarative Memory with Internet Statistics

Daniela Link (Daniela.Link@unil.ch) and Julian N. Marewski (Julian.Marewski@unil.ch)
Department of Organizational Behavior, UNIL-Internef
1015 Lausanne Switzerland

Keywords: memory; accessibility; ACT-R; simple heuristics; decision making

Introduction

Memory processes drawing on declarative knowledge play an important role in many cognitive models, for example in models of decision making. Within the fast-and-frugal heuristics research program, several strategies have been proposed that describe how people infer unknown criteria using knowledge associated with these criteria as *cues*. Much of the success of fast-and-frugal heuristics lies in their *ecological rationality*, or fit to regularities in the environment. Ecological rational decision strategies exploit regularities in the structure of the environment as they are reflected in basic cognitive capacities, such as memory. However, little research has looked at how environmental structures are mapped into mental representations. The ACT-R architecture offers a quantitative theory about how patterns of occurrences and co-occurrences of information in the environment are reflected in the memory activation of corresponding chunks. In this poster, we propose an ACT-R based ecological memory model representing objects and associated knowledge contingent on environmental frequencies of information encoded in the corresponding memory chunks. Based on internet statistics, we predict retrieval probabilities and retrieval latencies for associative knowledge, which will serve, for example, as input for simulating the selection and performance of knowledge-based decision strategies. A corresponding model could provide the missing link explaining how the interplay between the environment and the cognitive system promotes ecologically rational decision making.

Modeling Associative Memory in ACT-R

The basic unit of knowledge in ACT-R's declarative memory is the *chunk*. New declarative knowledge is added to memory by encoding representations of objects that are attended in the environment. A chunk can encode discrete elements of information as well as associations between elements being attended at the same time. The type of pattern encoded in the chunk is given in a *isa* slot, whereas other slots indicate the relationship between the elements of information that is being configured together. The knowledge that Berlin has an airport, for example, can be represented in a chunk with the following structure:

BERLIN-AIRPORT

ISA	CITY_FACT
CITY	BERLIN
FACT	AIRPORT

In addition to symbolic information, each chunk encodes subsymbolic information about the likelihood that the chunk will be needed to reach one of the system's processing goals -the chunk's *activation*. The likely usefulness is a Bayesian estimate of posterior need odds derived from the past usefulness of the chunk (prior odds, or *history factor*) as well as from the current context (likelihood ratio, or *context factor*). ACT-R's theory of human associative memory offers a set of equations to calculate a chunk's activation from these two factors. Specifically, the activation, A_i, of a chunk i is determined by the base-level activation, B_i, plus the spreading activation the chunk receives from each of the j elements in the current context:

$$A_i = B_i + \sum_{j=1}^{m} W_j S_{ji} \quad \text{(Activation).}$$

Assuming approximately equal spacing of encounters of a chunk since its time of creation L, the base-level activation of a chunk can be approximated by (Anderson, 1993):

$$B_i = \ln n/(1\text{-}d) - d \ln L \quad \text{(History Factor),}$$

where d is a decay parameter and n is the number of encounters of the object or relation encoded by the chunk.

In addition to the base-level activation which reflects the prior use of the chunk itself, a chunk receives spreading activation from related chunks currently attended in the current context. The amount of spreading activation a chunk receives depends on the associative strength, S_{ji}, between elements j stored in the buffers and chunk i as well as on the weight W_j given to each source of activation. The associative strength factor S_{ji}, can be calculated from environmental frequencies of occurrences and co-occurrences of chunk i and elements j according to the following equation (Schooler & Anderson, 1997):

$$S_{ji} = \ln \frac{P(i|j)}{P(i)} \quad \text{(Context Factor),}$$

where $P(i|j)$ is an estimate of the probability of i occurring when j is present and $P(i)$ is the base rate of i occurring. Source activation is typically divided equally among the number of sources of activation, m, and sums to a constant, W, which implies that

$$W_j = W/m \quad \text{(AttentionWeighting),}$$

In ACT-R, only chunks that exceed a certain amount of activation A_i, as defined by the retrieval threshold, τ, can be retrieved. Because of the stochastic volatility in momentary activation levels, chunks exceed this threshold with a certain probability. The retrieval probability, p, for chunk i, is a logistic function of the chunk's activation:

$$p_i = \frac{1}{1+e^{\frac{-(A_i-\tau)}{s}}} \quad \text{(Retrieval Probability),}$$

where s is a scale parameter representing noise in the retrieval process. Given a chunk is successfully retrieved, the retrieval time can be expressed as an exponential function of the chunk's activation:

$$T_i = Fe^{-A_i} \quad \text{(Retrieval Time).}$$

The above equations describe how patterns of occurrences and co-occurrences of objects in the environment are reflected in subsymbolic properties related to the activations and associative strengths of chunks in ACT-R's declarative memory. Importantly, the chunk's activations, in turn, allow for behavioral predictions about retrieval probabilities and retrieval times for the corresponding memories. We use observed retrieval probabilities and retrieval time distributions for knowledge about cities to calibrate our memory model for chunks encoding associative knowledge about these cities. We then use our model to predict people's knowledge of and retrieval speed for cities and associated facts from frequency statistics obtained from the internet.

Behavioral Data

One-hundred twenty-eight students (54 female; mean age 20 years, SD = 2.23) took part in the experiment. Participants received a fixed payment of 5 CHF (5.39 US$) supplemented by a performance bonus of up to 33 CHF (35.56 US$) depending on the coherence of their responses in the main task with responses given in a later control task where similar knowledge was tested. The stimuli for the cue-knowledge task (see Figure 1) consisted of 95 European cities and eight cues. Cue-knowledge tested was whether the city had an *airport*, a *university*, a premier league *soccer* team, the headquarters of a *company* listed on the stock market, a *cathedral*, a *subway*, a *harbor*, and whether it was served by a high-speed *train* line. Each city was paired with each of the cues, so that the items consisted of a total of 760 city-cue pairs.

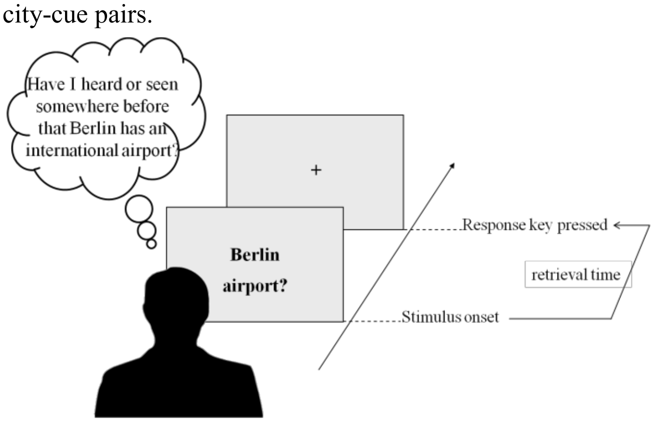

Figure 1: Illustration of the cue-knowledge task.

Participants were presented with city-attribute pairs one at a time in and were asked to respond with either "yes" if they could remember having seen or heard of the city possessing such an attribute or "no" if they could not remember having heard of this before. Responses were made by pressing keys on the right and left side of the keyboard. The order of presentation of items was randomized. All trials were preceded by a small fixation cross for 1,000 msec and participants were instructed to fixate the cross until it disappeared and to respond as quickly and accurately as possible upon stimulus onset.

Predicting Accessibility from Internet Statistics

We approximate memory activation A_i resulting from encounters with certain information in a person's environment by the activation $A_{i,web}$ estimated from *web counts*, the number of entries for this information in the knowledge base Wikipedia.

$$A_i = c + b\, A_{i,web}.$$

The parameters c and b serve as scaling parameters describing the unknown relation between how often we encounter an object in our environment and the web frequency of the corresponding search term.

We calibrated the memory model to the log odds of retrieval of cue-knowledge. Subsequently, we calibrated the model to the observed retrieval times for retrieved cue-knowledge. Activations for chunks encoding cue-knowledge estimated from web counts were then used to predict observed retrieval probabilities and retrieval time distributions.

Conclusion

Comparisons between observed and predicted retrieval probabilities, and observed and predicted retrieval time distributions show that our memory model is able to capture how the probability of retrieval and the accessibility of cue-knowledge depends on the distribution of relevant information in the environment.

Our work extends the ecological approach for populating the contents of declarative memory in ACT-R (e.g., Marewski & Schooler, 2011). Possible applications include the simulation of performance of and selection between knowledge-based inference strategies and could be used for any model interested in mirroring the statistical structure of the environment outside the laboratory.

Acknowledgments

This work was supported by a grant from the Swiss National Science Foundation [Grant number 100014_144413] to Julian Marewski.

References

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Marewski, J. N., & Schooler, L. J. (2011). Cognitive niches: an ecological model of strategy selection. *Psychological Review*, *118*, 393.

Schooler, L. J., & Anderson, J. R. (1997). The role of process in the rational analysis of memory. *Cognitive Psychology, 32,* 219–250.