Foundation of Base-Level Activation in the Environment

John Anderson
Carnegie Mellon

1

Foundation of Base-Level Activation in the Environment

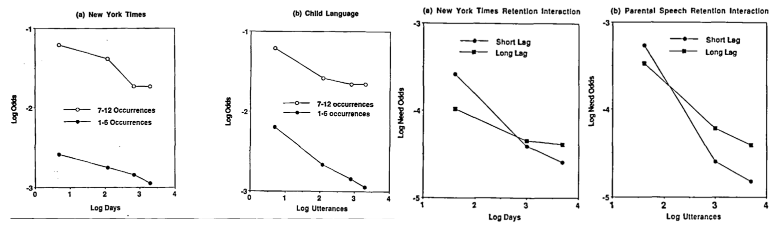
- Base-level activation in ACT-R is supposed to reflect the log odds that a memory is needed now given its past history.
- Based on studies of repetition of items in the environment (Anderson & Schooler, 1991, and much subsequent work of Schooler) the base-level action of memory i is:

  $B_i = \log\left(\sum_{j=1}^{n} t_j{}^{-d}\right) + B$

  where t_j is time since jth presentation of i.
- This model accounts for the effects of retention interval and practice and some of the interactions of the two, both in the environment and in human memory.
- However, it does not account for the spacing effect, which is found in both the environment and memory -- Pavlik's(Pavlik & Anderson, 2005) extension to ACT-R does.

2

Extending Anderson & Schooler, 1991



- Today we have access to much larger data bases and the compute power to process them, enabling us to perform a much more detailed analysis.
- **Data Base 1 (Stanley, 2014):** All tweets from the top 500 most followed English tweeters as of Jan 7, 2014, from July 11, 2007 to Jan 7, 2014.
- **Reddit Data Base:** The top 500 subreddits (by subscribers) on two days, April 23 and May 5, 2021. The visible comments in the top 25 topics.

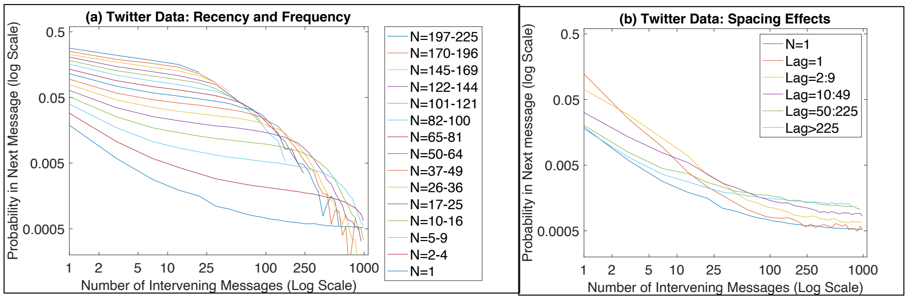
3

Analysis of the New Data Bases

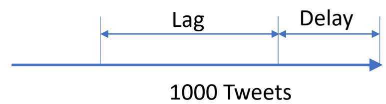
- **Twitter:** Following a tweeter what is the probability of a string in a message like *This debate is not just about numbers. It's a set of major decisions that are going to affect millions of families.*
- **Reddit:** Reading the comments on a subreddit what is the probability of a string in a message like *I received 100% attendance in 4th grade and got a free ice cream dessert at Ponderosa.*
- **The Question:** Focusing on 20,000 most frequent non-functor strings, what is the probability of a string in the next message (tweet from a tweeter or comment on a subreddit) given its pattern of occrrence over the last 1000 messages.
- **Twitter:** Over a million tweets, averaging 7.5 unique strings, offering over a billon 1000-tweet patterns.
- **Reddit:** Over a million comments, averaging 14.9 unique strings, offering over a billon 1000-comment patterns.

4

Patterns of Repetition



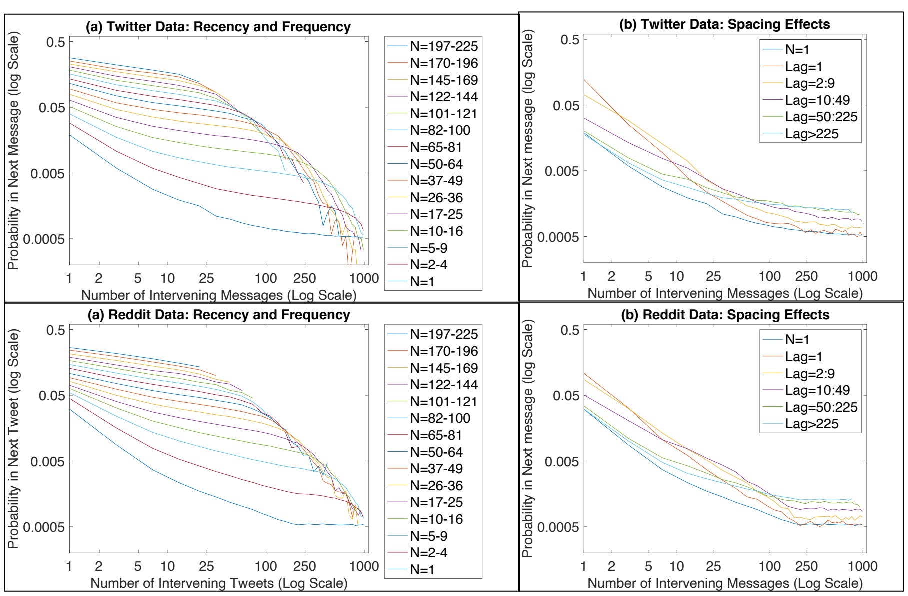
- Prob~.0255*Lag$^{-.54}$*N$^{.65,}$ R^2=.843 in log scale.
- Curves are negatively accelerated to a point after which they become positively accelerated.
- The "final crash" reflects a response to massing of presentations a long time ago.



- Short lags are best at brief delays and worst at long delays.
- Short lags at long delays not much better than a single presentation.

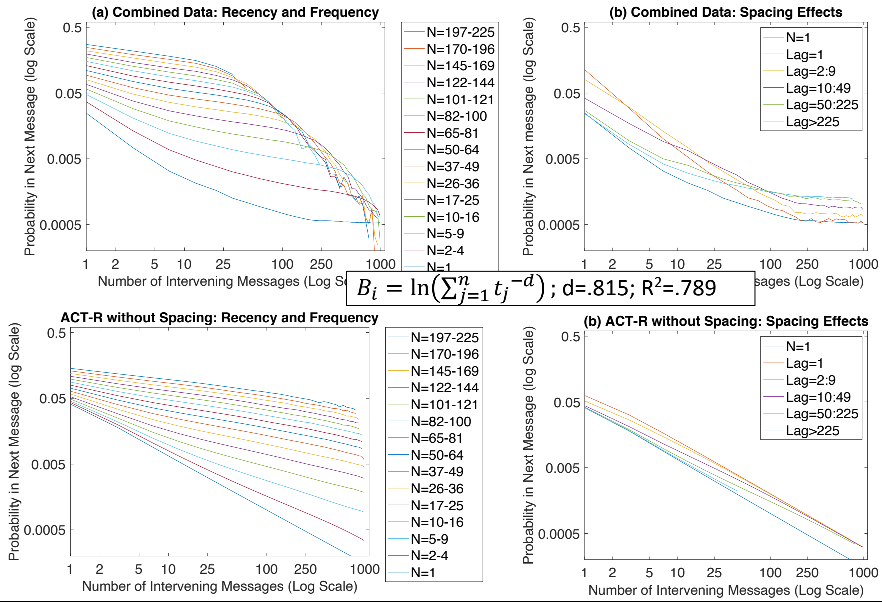
5

Patterns of Repetition



6

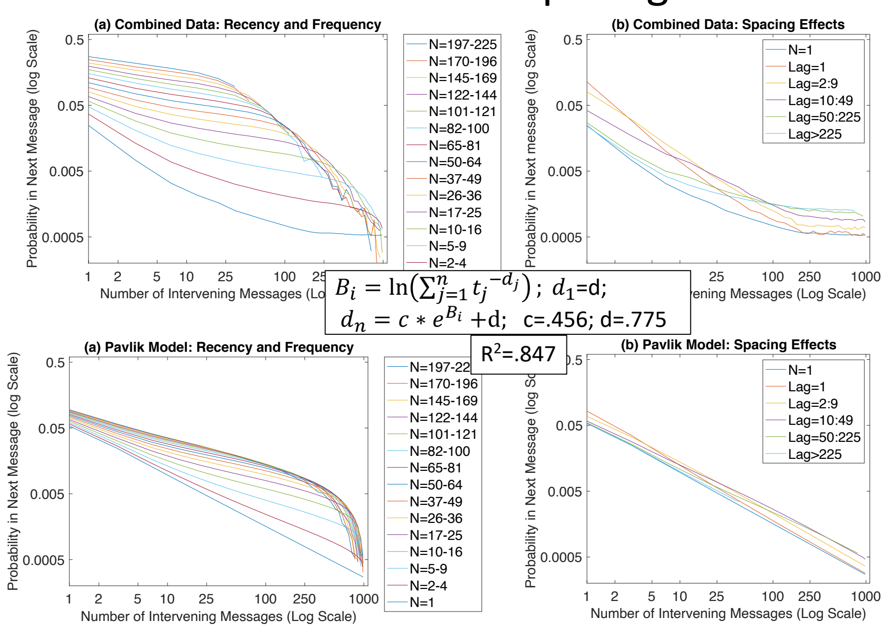
ACT-R without Spacing



$B_i = \ln\left(\sum_{j=1}^{n} t_j^{-d}\right)$; d=.815; R^2=.789

7

Pavlik Model of Spacing



$B_i = \ln\left(\sum_{j=1}^{n} t_j^{-d_j}\right)$; d_1=d;
$d_n = c * e^{B_i}$ +d; c=.456; d=.775

R^2=.847

8

Anderson & Milson (1989) Model

1. The desirabilities, λ's, of items are distributed according to a gamma function: $\pi(\lambda) = \frac{\lambda^{v-1}e^{-\lambda/b}}{\Gamma(v)b^{v}}$; $\Gamma(v)$=(v-1)! for integer v.
2. Items decay in desirability over time according to an exponential function: $r(t) = e^{-dt}$. (I will also explore the power function $r(t) = t^{-d}$.)

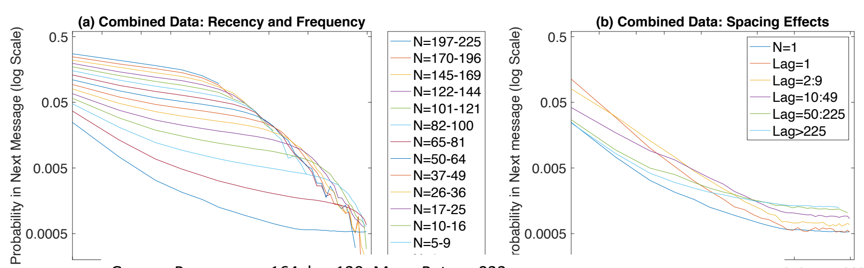
The above 2 assumptions are derived from Burrell's model of library borrowing and imply that the expected probability of an item occurring that has had n occurrences in time t is $\frac{v+n}{M(t)+1/b}$r(t) where $M(t) = \int_0^t r(s)ds$.

3. The rate of decay, d, varies for items: f$(d) = e^{-d/\alpha}/\alpha$.
4. With each unit of time there is a probability p that the item will have a revival of desirability to original level.

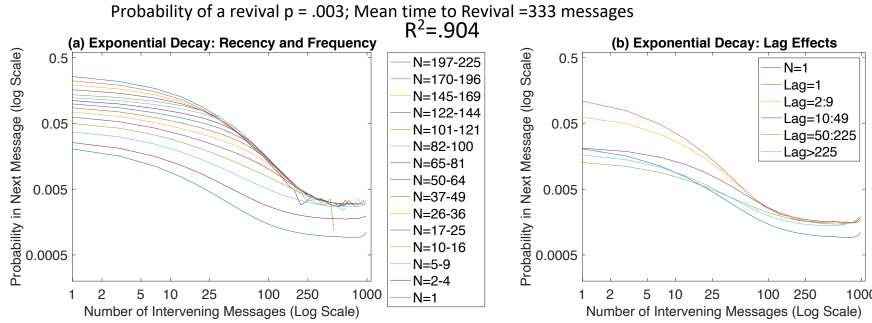
Assumptions 3 & 4 eliminate the possibility of a closed form calculation and require simulation.

9

Original Anderson & Milson Model

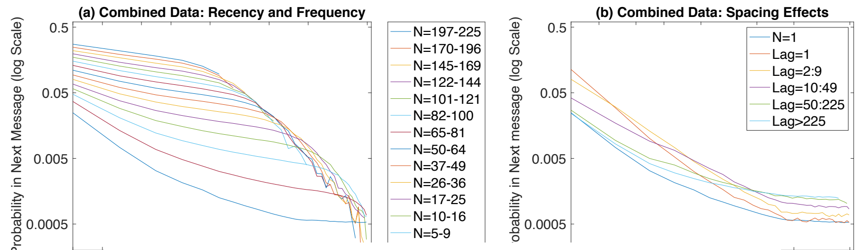


Gamma Params: v = .164, b= .139: Mean Rate = .023
Exponential Distribution of d in e^{-dt}, α = .035; average of 22% left after 100 messages
Probability of a revival p = .003; Mean time to Revival =333 messages
R^2=.904



10

Anderson & Milson: Power Decay

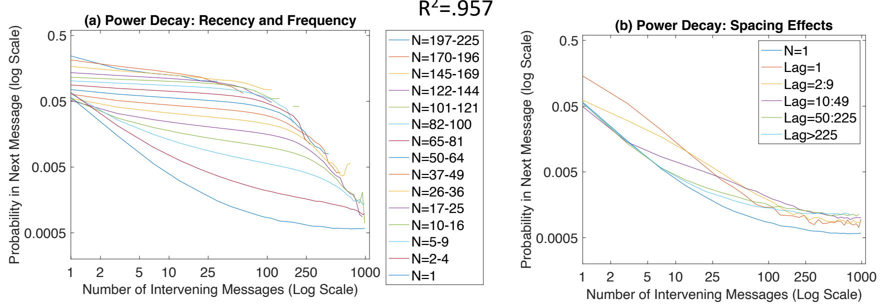


Gamma Params: v = .199, b= .482: Mean Rate = .096

Exponential Distribution of d in t^{-d}, α = 4.08; average of 5% left after 100 messages

Probability of a revival p = .001; Mean time to Revival =1000 messages

R^2=.957



11

A Mechanism for Anderson & Milson

- A&M's Monte Carlo simulation is implausible as model of human memory and not something to run as part of ACT-R
- However, it suggests something like to PPE (Walsh, Gluck, Gunzelmann, Jastrzembski , Krusmark (2018):

$$B_i = \pi_i * T_i^{-d_i}$$

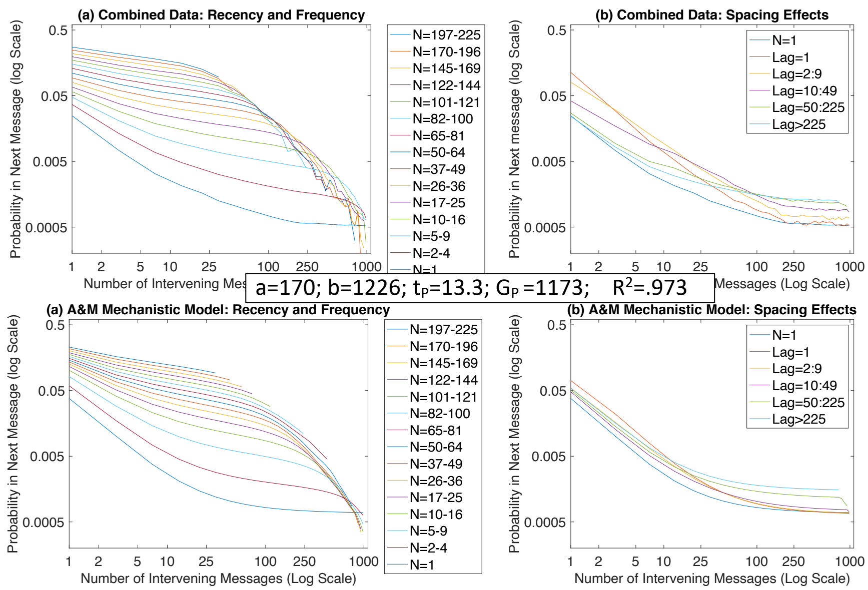
- T_i is an estimate of the time since the last revival estimated as as the harmonic of prior t_p plus observed times $t_1,t_2,\ldots,t_n$.
- The initial desirability and the decay are functions of the effective interval $M=(G_p +G)/2$ where $G =(t_n-t_1)+1$:

$$\pi_i = \frac{a*n}{M} \quad d_i = \frac{b}{M}$$

- Note opposing effects of M and hence gap G.
- Note that with respect to decay the critical feature is not the spacing of interval between the items but the period G over which they appear.

12

Anderson & Milson: Mechanistic Model



a=170; b=1226; t_P=13.3; G_P =1173; R^2=.973

13

Conclusions

- Assuming power-law decay Anderson & Milson is a surprisingly good model of the memory demands of the environment.
- A simple PPE-like equation can be used which is easy to compute and captures spacing effects: $B_i = \pi_i * T_i^{-d_i}$.
- This equation suggests that the critical variable might not be the spacing between individual items but rather the period of time spanned by all presentations.
- While high-density presentation has not been a focus of memory research, it happens in the real world and it is implausible that all those close presentations (e.g. of my wife's name) are causing rapid forgetting.
- Preliminary efforts to extend $B_i = \pi_i * T_i^{-d_i}$ to human memory experiments appear to do a good job using the ACT-R

$$prob = \frac{1}{1 + e^{\frac{\tau - B_i}{s}}}$$

14