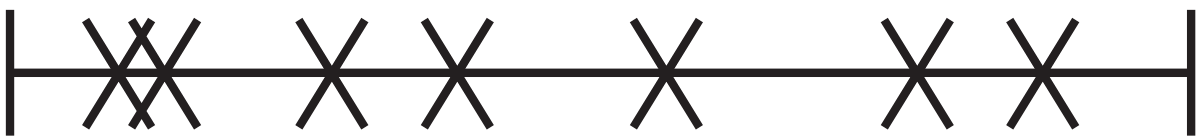
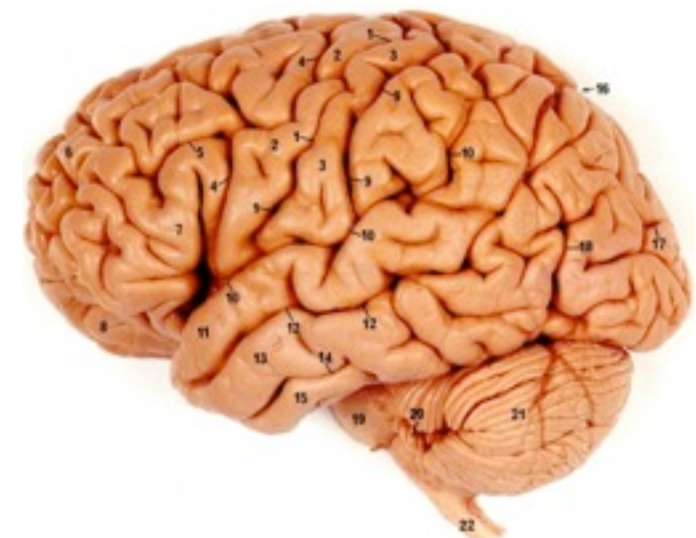
FMIE Analysis of Human Memory

Joe Austerweil





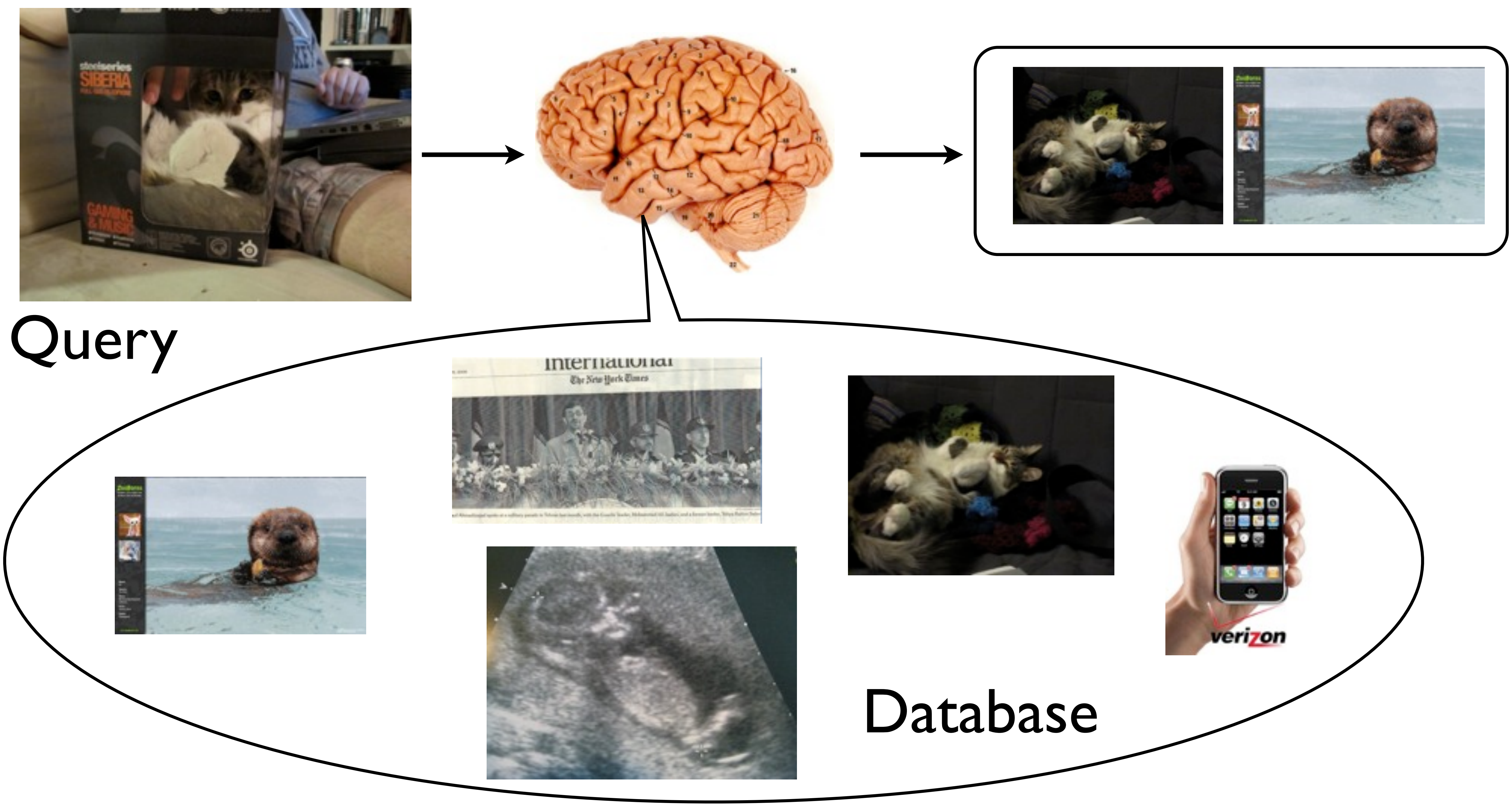
Poisson Processes & Brains

What is memory?

View memory as a information retrieval engine.

Given query of items, which items in my database are relevant?

Prob. of relevance is a function of features in common and prior use.



Adaptive memory

Let $h_1, h_2, \ldots; h_i \in \{0,1\}^D$ be items of memory defined by the state of the brain and sensory receptors.

Let $q \in \{0,1,?\}^D, m = \sum_{i=1}^{D} I(q_i =?), D - m << D,$ be a query.

It is an incomplete description of an item in memory.

Adaptive memory

Let $h_1, h_2, \ldots; h_i \in \{0,1\}^D$ be items of memory defined by the state of the brain and sensory receptors.

Let $q \in \{0,1,?\}^D, m = \sum_{i=1}^{D} I(q_i =?), D - m << D,$ be a query.

It is an incomplete description of an item in memory.

We want $P(h|d) \propto P(d|h)P(h)$

Assume the missing bits are picked uniformly at random without replacement

$$P(d|h) = \begin{cases} \frac{m!(D-m)!}{D!} & M(d,h) \\ 0 & o.w. \end{cases}$$

where $M(d,h)$ means that the query is consistent with the memory.

Adaptive memory

Let $h_1, h_2, \ldots; h_i \in \{0,1\}^D$ be items of memory defined by the state of the brain and sensory receptors.

Let $q \in \{0,1,?\}^D, m = \sum_{i=1}^{D} I(q_i = ?), D - m << D,$ be a query.

It is an incomplete description of an item in memory.

We want $P(h|d) \propto P(d|h)P(h)$

Assume the missing bits are picked uniformly at random without replacement

$$P(d|h) = \begin{cases} \frac{m!(D-m)!}{D!} & M(d,h) \\ 0 & o.w. \end{cases}$$

where $M(d,h)$ means that the query is consistent with the memory.

It is less clear how to define $P(h)$ properly.

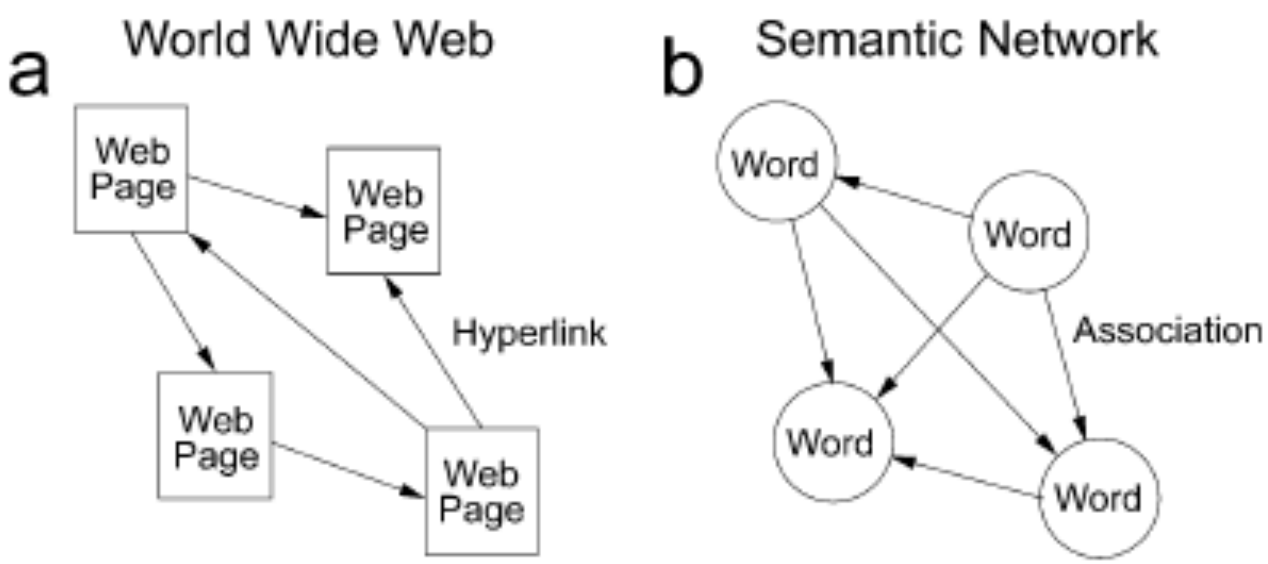
Intuition: order by $P(h)$ and search an infinite bookshelf left to right.

Adaptive memory: $P(h)$

How do we determine which memories are likely *a priori?*

Multinomial based on past history of observing items?

Connectivity to other items (Griffiths, Steyvers, & Firl, 2007; Brin & Page, 1998)



Unweighted RW on directed graph.

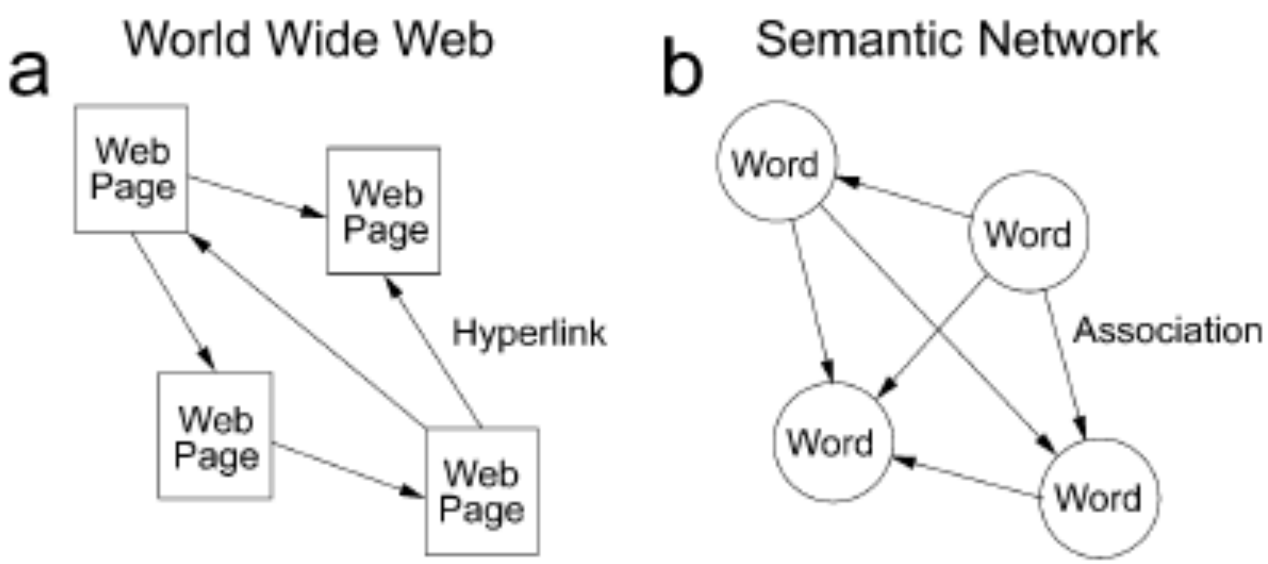
Stationary distribution gives prior.

Adaptive memory: $P(h)$

How do we determine which memories are likely *a priori?*

Multinomial based on past history of observing items?

Connectivity to other items (Griffiths, Steyvers, & Firl, 2007; Brin & Page, 1998)



Unweighted RW on directed graph.

Stationary distribution gives prior.

Similarities between web and memory search.

Subjects shown a letter (e.g., "A") and say first word that comes to mind starting with that letter (e.g., "Apple").

Based on word-association norm data, word association graph was created.

Human responses compared to PageRank, raw association and word frequencies.

PageRank ran on the word graph best explained participant responses.

Beginning letter						
A	B	C	D	P	S	T
Human responses						
Apple (25)	Boy (11)	Cat (26)	Dog (19)	People (5)	Snake (11)	Tea (5)
Alphabet (7)	Bat (6)	Car (8)	Dad (16)	Penguin (3)	Stop (4)	Television (5)
Ant (6)	Banana (5)	Cool (3)	Door (5)	Pizza (3)	Saw (2)	Time (4)
Aardvark (3)	Balloon (4)	Card (2)	Down (4)	Play (3)	Sea (2)	Tree (4)
Ace (2)	Book (4)	Class (2)	Dark (3)	Pop (3)	Sex (2)	Table (3)
Ambulance (2)	Baby (3)	Coke (2)	Dumb (3)	Puppy (3)	Silly (2)	Tall (3)
Animal (2)	Ball (2)	Cookie (2)	Day (2)	Piano (2)	Sister (2)	Tank (3)
Absence (1)	Barn (2)	Crack (2)	Devil (2)	Pie (2)	Sit (2)	Telephone (3)
Acrobat (1)	Bear (2)	Cross (2)	Dinosaur (2)	Pig (2)	Slither (2)	Town (3)
Act (1)	Beef (2)	Cut (2)	Do (2)	Power (2)	South (2)	Train (3)
PageRank						
Animal (2)	Big (0)	Cold (0)	Dog (19)	Pretty (0)	Small (1)	Time (4)
Away (0)	Bad (1)	Car (8)	Dark (3)	People (5)	Sad (1)	Tall (3)
Air (0)	Boy (11)	Cat (26)	Drink (1)	Paper (0)	School (0)	Talk (1)
Alone (0)	Black (0)	Color (0)	Down (4)	Pain (0)	Sun (2)	Tree (4)
Apple (25)	Beautiful (0)	Clothes (0)	Death (1)	Puppy (3)	Smile (0)	Tired (0)
Arm (0)	Blue (2)	Child (1)	Door (5)	Person (1)	Stop (4)	Tiny (0)
Ache (0)	Book (4)	Cute (0)	Day (2)	Play (3)	Soft (1)	Thin (0)
Answer (1)	Body (0)	Clean (0)	Dirty (0)	Place (1)	Sex (2)	Top (1)
Apartment (0)	Bright (0)	Close (0)	Dirt (0)	Party (0)	Sky (0)	Together (0)
Alcohol (0)	Baby (3)	Cry (0)	Dead (0)	Pen (0)	Sleep (0)	Train (3)
Associate frequency						
Animal (2)	Bad (1)	Car (8)	Dog (19)	Paper (0)	School (0)	Time (4)
Air (0)	Book (4)	Clothes (0)	Death (1)	Pain (0)	Small (1)	Tree (4)
Army (0)	Black (0)	Cold (0)	Drink (1)	People (5)	Sex (2)	Talk (1)
Away (0)	Big (0)	Clean (0)	Dirty (0)	Person (1)	Sad (1)	Together (0)
Anger (0)	Baby (3)	Child (1)	Dark (3)	Play (3)	Soft (1)	Test (1)
Answer (1)	Ball (2)	Class (2)	Down (4)	Party (0)	Stop (4)	Television (5)
Art (0)	Body (0)	Church (0)	Dirt (0)	Pretty (0)	Smell (0)	Think (0)
Apple (25)	Bird (0)	Cut (2)	Dead (0)	Problem (0)	Strong (0)	Top (1)
Alcohol (0)	Break (0)	Color (0)	Dance (0)	Police (1)	Smart (0)	Teacher (0)
Arm (0)	Boring (0)	Cat (26)	Danger (1)	Place (1)	Sick (0)	Take (0)
Word frequency						
A (0)	Be (1)	Can (0)	Do (2)	People (5)	She (0)	There (0)
All (0)	Before (0)	Come (0)	Down (4)	Place (1)	Some (0)	Than (0)
After (1)	Back (0)	Course (0)	Day (2)	Part (0)	State (1)	Time (4)
Another (0)	Because (0)	City (0)	Development (0)	Public (1)	Still (0)	Two (1)
Against (0)	Between (0)	Case (0)	Done (1)	Put (2)	See (0)	Through (0)
Again (0)	Being (0)	Children (0)	Different (0)	Point (0)	Same (0)	Take (0)
American (0)	Better (0)	Church (0)	Door (5)	Program (0)	Since (0)	Three (0)
Around (0)	Business (0)	Country (0)	Death (1)	President (0)	Small (1)	Thought (0)
Always (0)	Become (0)	Certain (0)	Department (0)	Present (0)	Say (1)	Think (0)
Away (0)	Big (0)	Company (0)	Dark (3)	Possible (0)	School (0)	Thing (0)

Memory as a Cond. Poisson Process

What about time? Memories, like library books, become less used over time.

Different approach, item usages are a nonhomogeneous Poisson process* in time.

Desirability of history item: $\lambda_i \sim \mathrm{Gamma}(v, b)$

Rate of Poisson Process: $r(t)\lambda_i$

From conjugacy, posterior desirability of a history used n_i times in t_i time is:

$$\lambda_i | n_i, t_i \sim \mathrm{Gamma}\left(v + n_i, b + \int_0^{t_i} r(s)ds\right)$$

Thus, the prior probability that a memory is useful now is:

$$P(h_i | n, t_i) \propto \frac{v + n_i}{b + \int_0^{t_i} r(s)ds} r(t_i)$$

(Burrell, 1985; Anderson & Milson, 1989)

Memory as a Cond. Poisson Process

What about time? Memories, like library books, become less used over time.

Different approach, item usages are a nonhomogeneous Poisson process* in time.

Desirability of history item: $\lambda_i \sim \text{Gamma}(v, b)$

Rate of Poisson Process: $r(t)\lambda_i$

From conjugacy, posterior desirability of a history used n_i times in t_i time is:

$$\lambda_i | n_i, t_i \sim \text{Gamma}\left(v + n_i, b + \int_0^{t_i} r(s)ds\right)$$

Thus, the prior probability that a memory is useful now is:

$$P(h_i | n, t_i) \propto \frac{v + n_i}{b + \int_0^{t_i} r(s)ds} r(t_i)$$

Simple exponential decay function with parameter d_i: $r(t_i; d_i) = e^{-d_i t_i}$

whose prior is $d_i \sim \text{Exp}(\alpha)$

"Renewed interest" adds revivals of interest in item (resets time).

Happens at points given by a Poisson process with parameter β

(Burrell, 1985; Anderson & Milson, 1989)

Is this a FMIE process?

Remember that FMIE processes are defined by two levels:

1. A **meeting model** (matrix) specifying the prob. of two agents meeting.
2. How do agents update their states when they meet?

Is this a FMIE process?

Remember that FMIE processes are defined by two levels:

1. A **meeting model** (matrix) specifying the prob. of two agents meeting.
2. How do agents update their states when they meet?

The quick answer is not really because histories (agents) only "meet" with themselves, which results in them resetting their last use state to zero.

Is this a FMIE process?

Remember that FMIE processes are defined by two levels:

1. A **meeting model** (matrix) specifying the prob. of two agents meeting.
2. How do agents update their states when they meet?

The quick answer is not really because histories (agents) only "meet" with themselves, which results in them resetting their last use state to zero.

However, a simple alteration to the model could yield a FMIE:

1. Assume memories (agents) are linked in a topology (like a semantic web...).
2. Agents meet each other at a rate given by the "PageRank."
3. When two agents meet, they set their last use to the min of their states.

This would embed a "consensus" model.

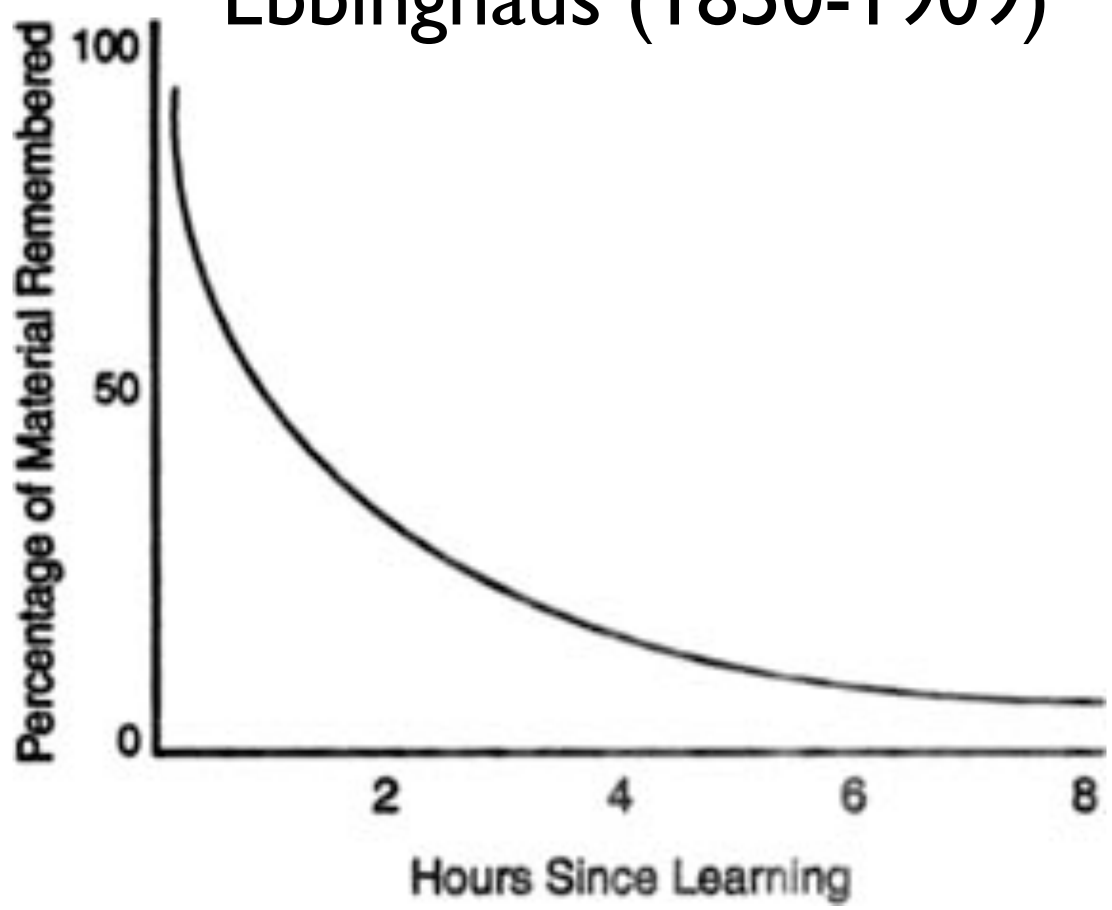
We will discuss this later, but first look at properties of the basic model and its relation to human memory phenomena.

Basic Memory Experiment Paradigm

Task:
Learn a list of words.

Ebbinghaus (1850-1909)

Measure:
% of words recalled later.

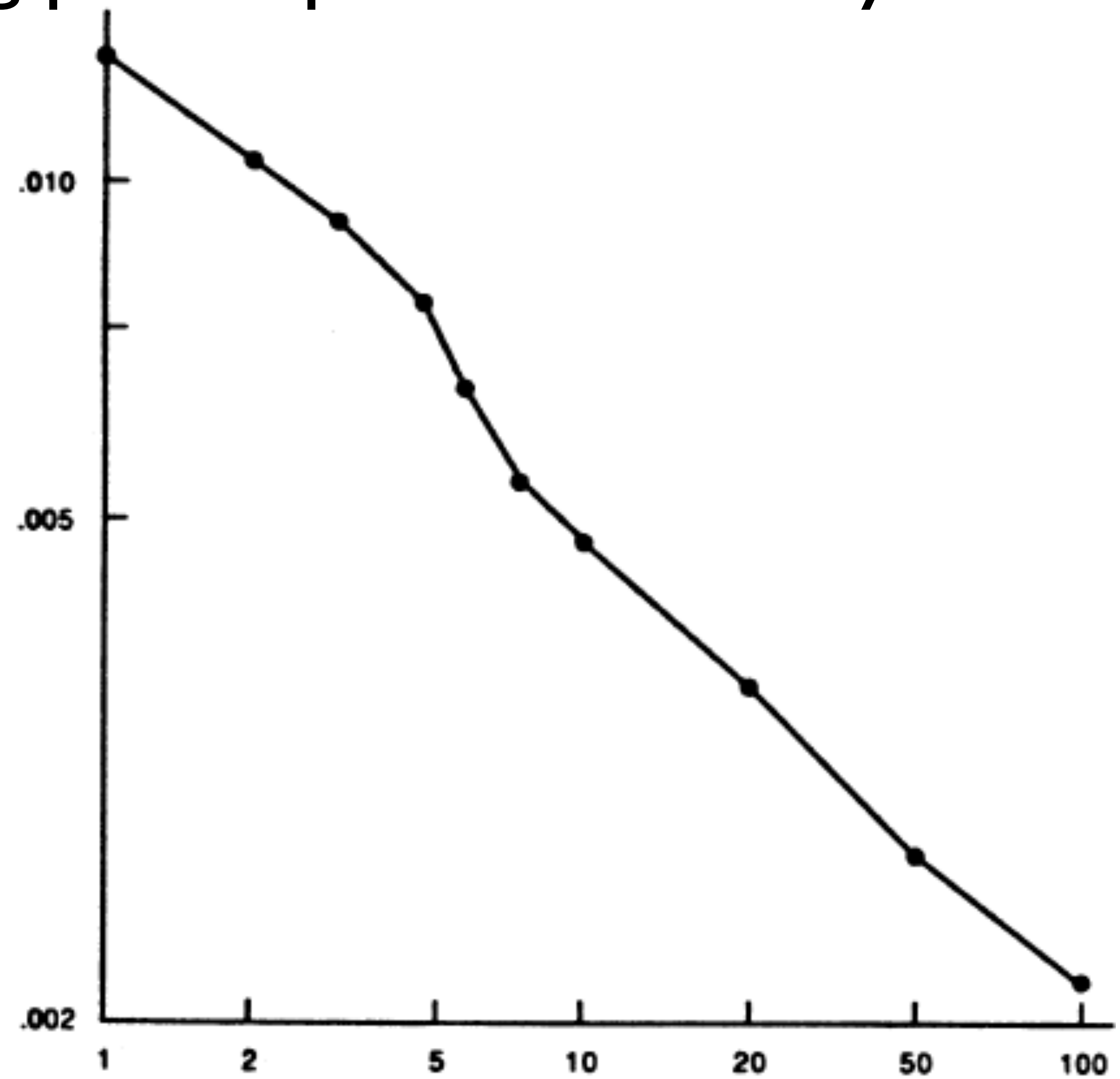


Results:
Exp. decay over time.

(Ebbinghaus, 1885)

Delay

Log-log plot of prob. of memory seen last at t_i



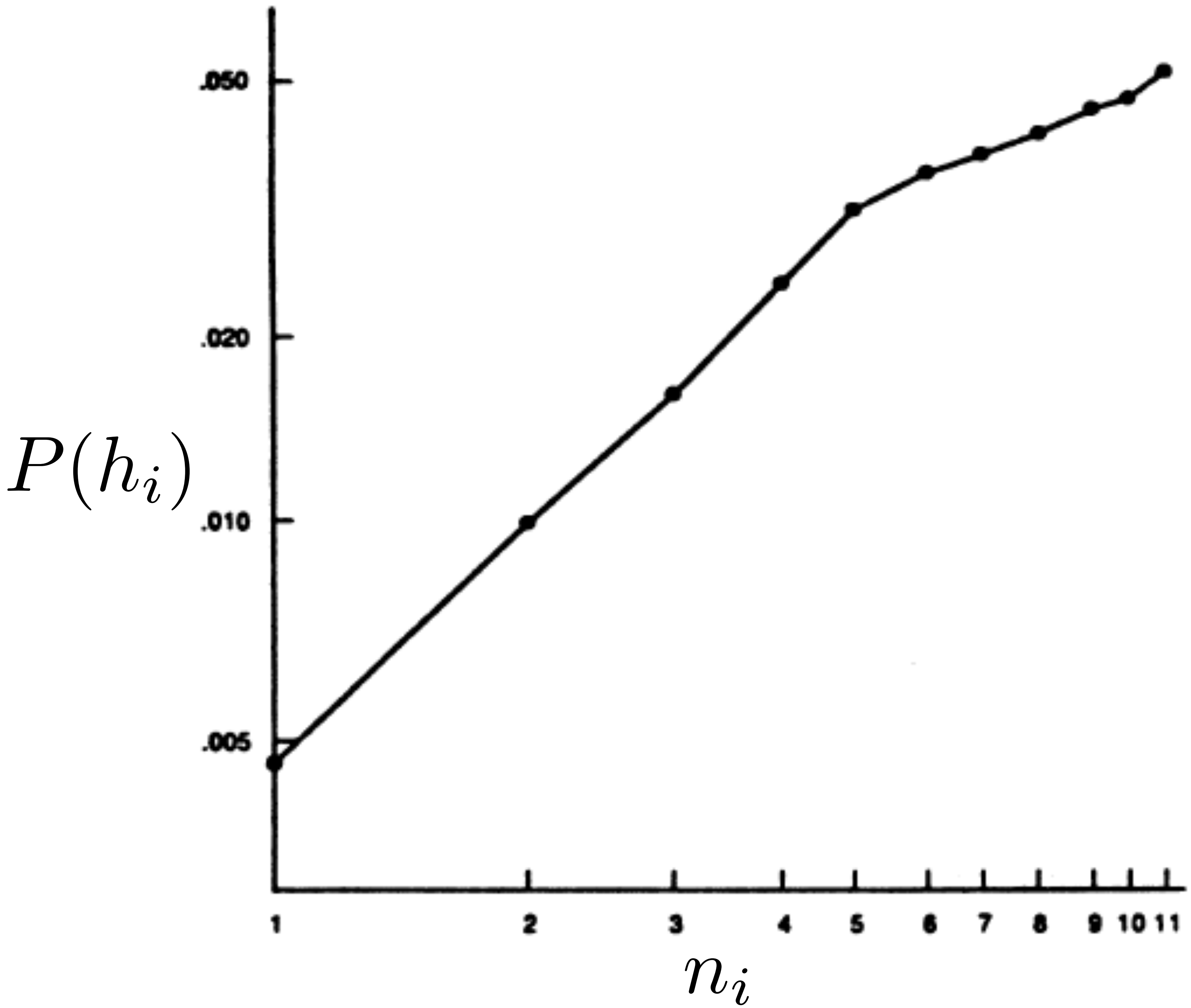
$P(h_i)$

Time t_i

(Anderson & Milson, 1989)

“Practice”

Log-log plot of prob. of memory seen n_i times



(Anderson & Milson, 1989)

Spacing Effects

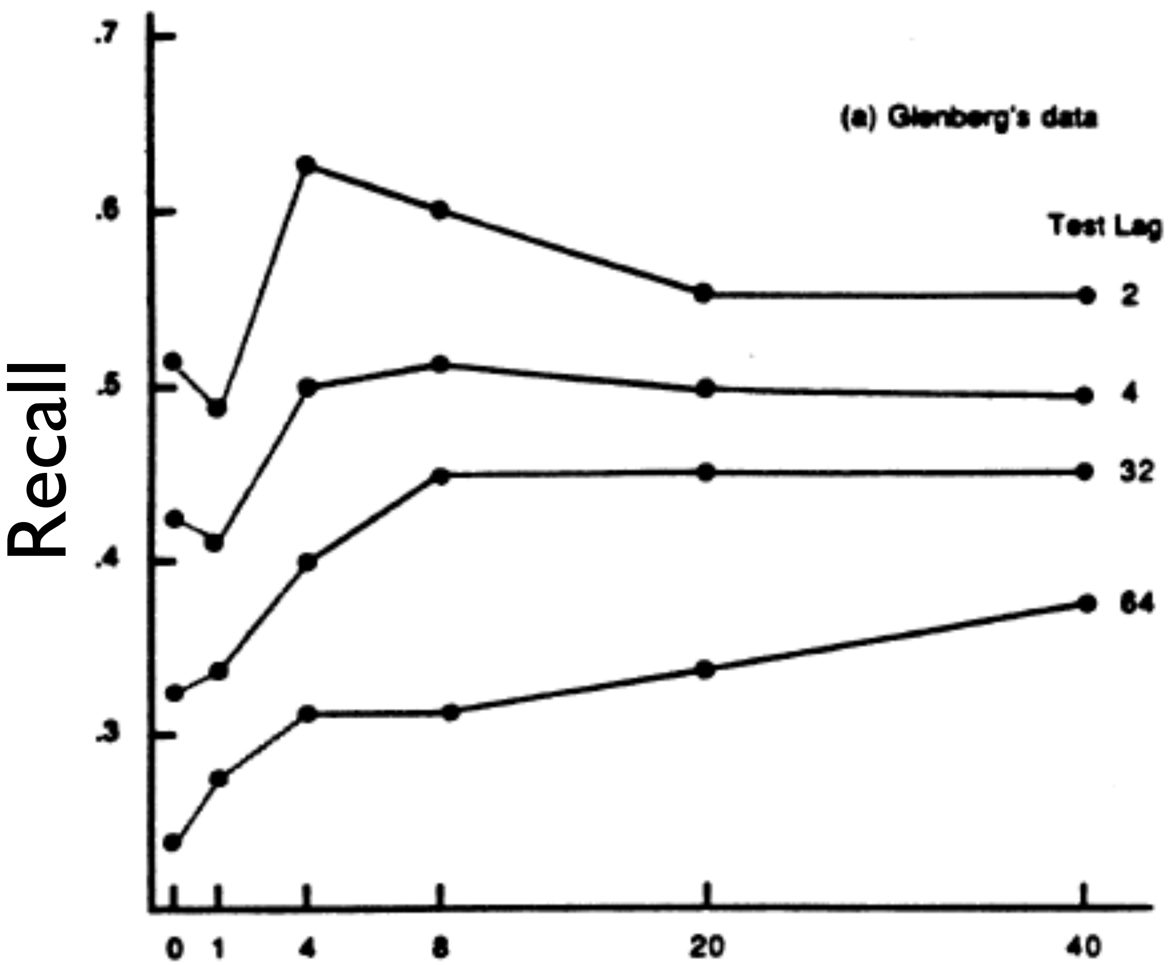
Remember a list of items, where there are two study sessions and one test.

Increased lag between study times *helps* when there's a large test lag.

Increased lag between study times *hurts* when there's a small test lag.

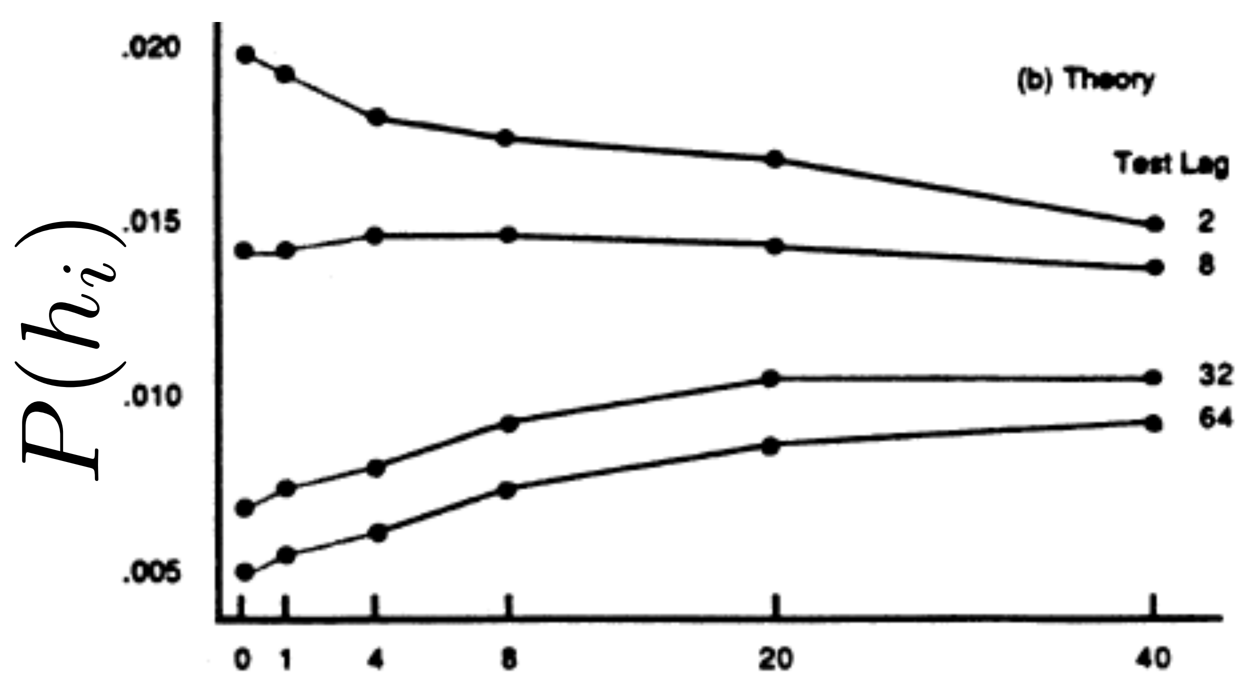
Plots are different test lags (increasing in lag from top to bottom).

Humans



Log lag between study times

Model



Log lag between study times

(Anderson & Milson, 1989)

Priming

Task: Look at letter sequences and say if it's a word or not.

Measure: Look at reaction time (RT) to respond (when correct).

Question: How is RT effected when the previous word is related?

$$P(h_i|\text{relevant}) > P(h_i)$$

Also, negative priming...

dish

Word or not word?

book

car

Word or not word?

dop

cat

Word or not word?

dog

Extending to FMIE Framework

Remember that FMIE processes are defined by two levels:

1. A **meeting model** (matrix) specifying the prob. of two agents meeting.
2. How do agents update their states when they meet?

However, a simple alteration to the model could yield a FMIE:

1. Assume memories (agents) are linked in a topology (like a semantic web...).
2. Agents meet each other at a rate given by the "PageRank" and desirability.
3. When two agents meet, they set their last use to the min of their states.

This would embed a "consensus" model.

Extending to FMIE Framework

Remember that FMIE processes are defined by two levels:

1. A **meeting model** (matrix) specifying the prob. of two agents meeting.
2. How do agents update their states when they meet?

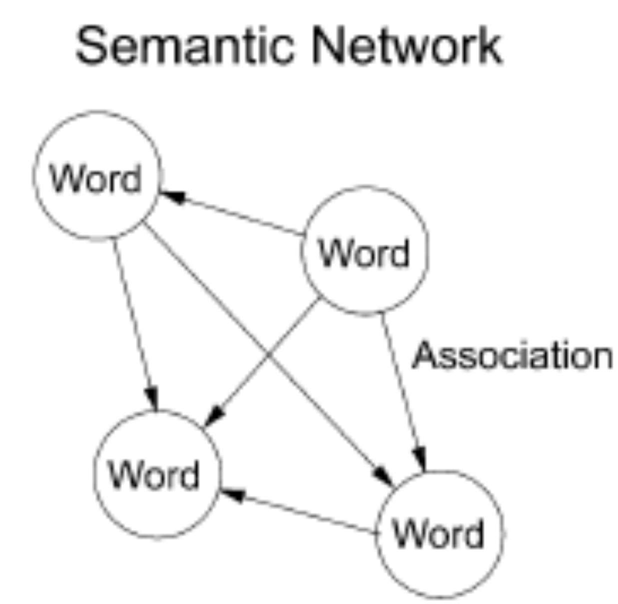
However, a simple alteration to the model could yield a FMIE:

1. Assume memories (agents) are linked in a topology (like a semantic web...).
2. Agents meet each other at a rate given by the "PageRank" and desirability.
3. When two agents meet, they set their last use to the min of their states.

This would embed a "consensus" model.

$$\mathcal{N} = \begin{pmatrix} d_1 & \dots & d_n \\ \vdots & \dots & \vdots \\ d_1 & \dots & d_n \end{pmatrix} \otimes X \otimes (\mathbf{1} - I)$$

where X is the connectivity matrix.

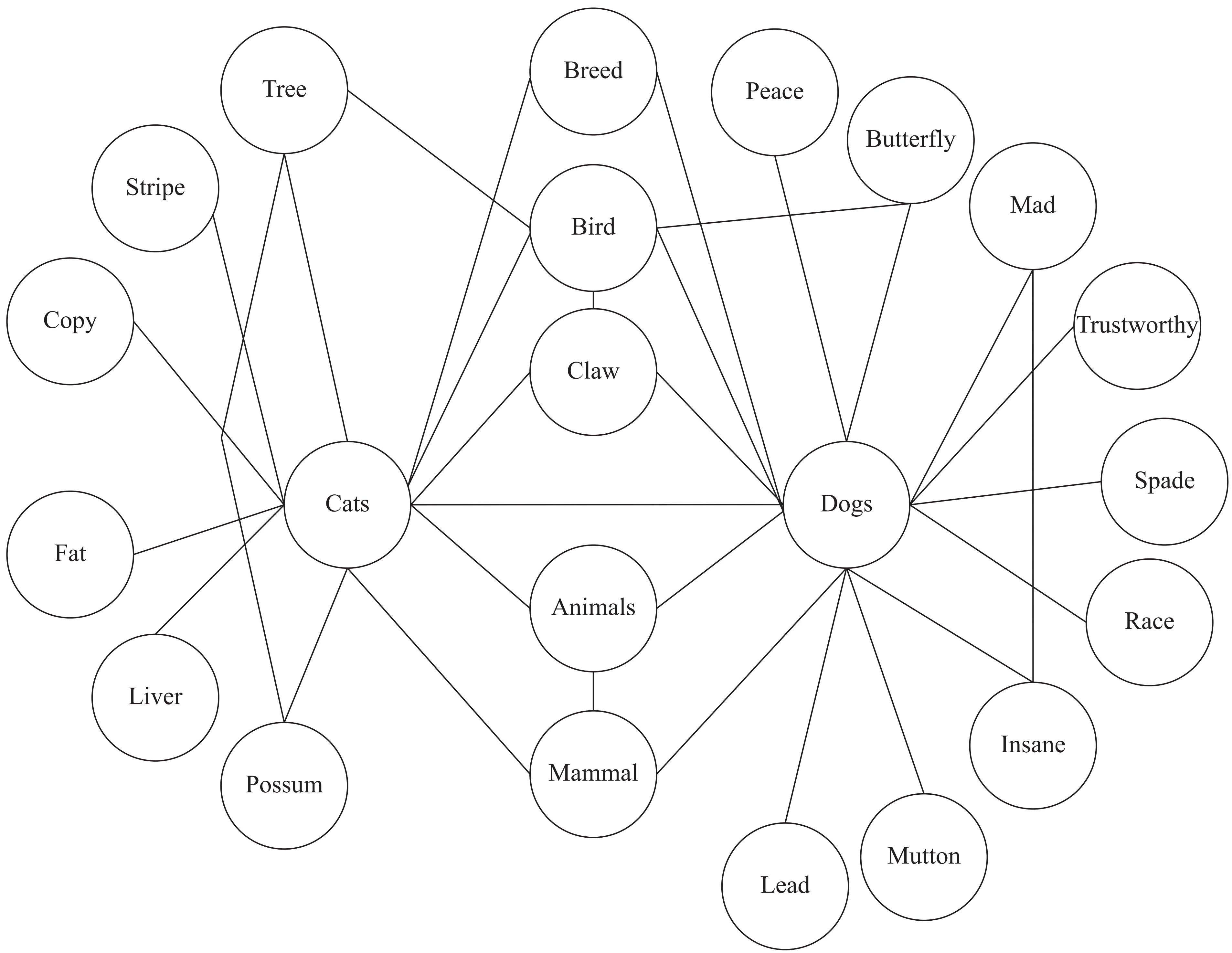


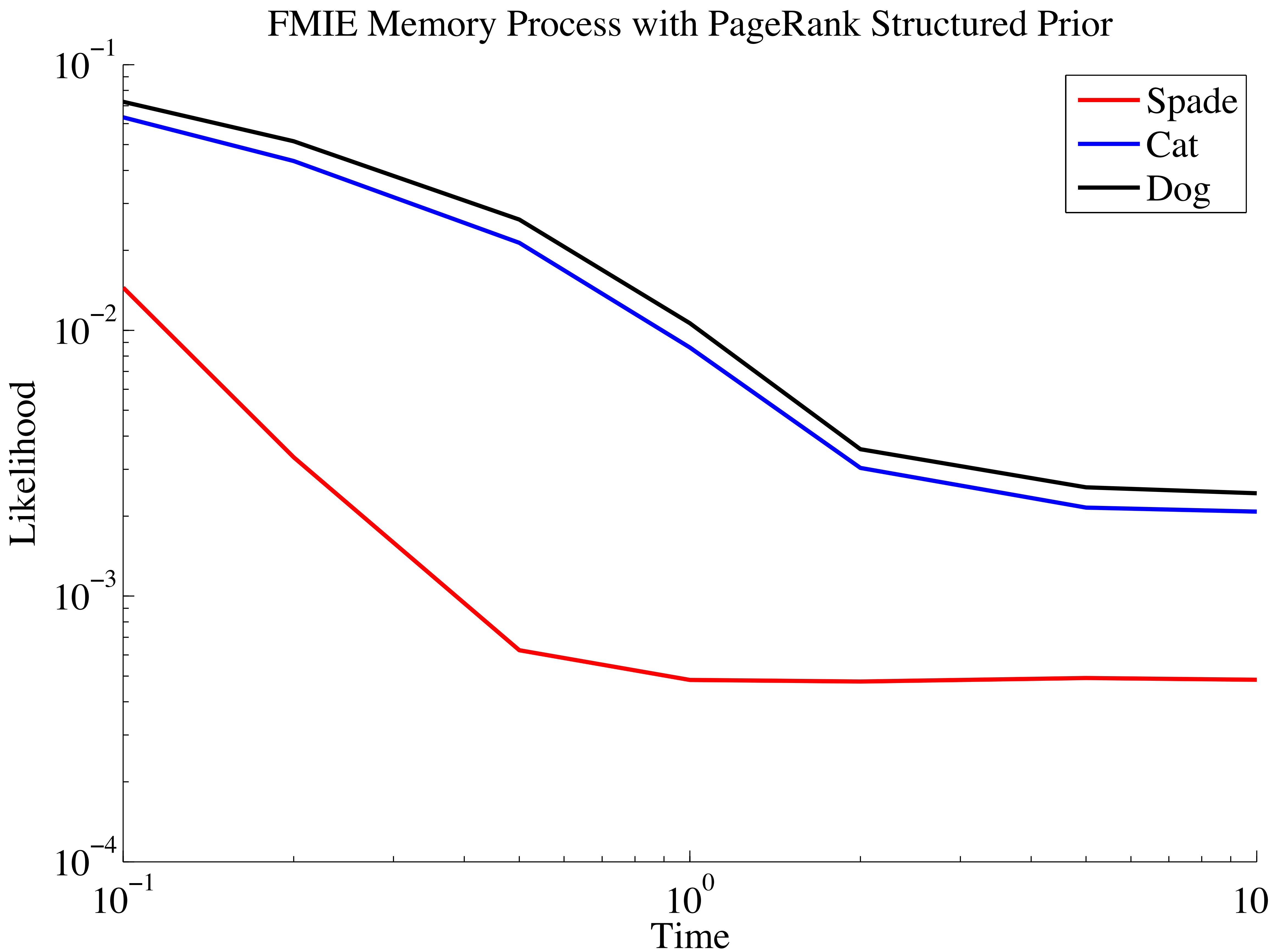
The connectivity matrix is made symmetric if it has asymmetric relations.

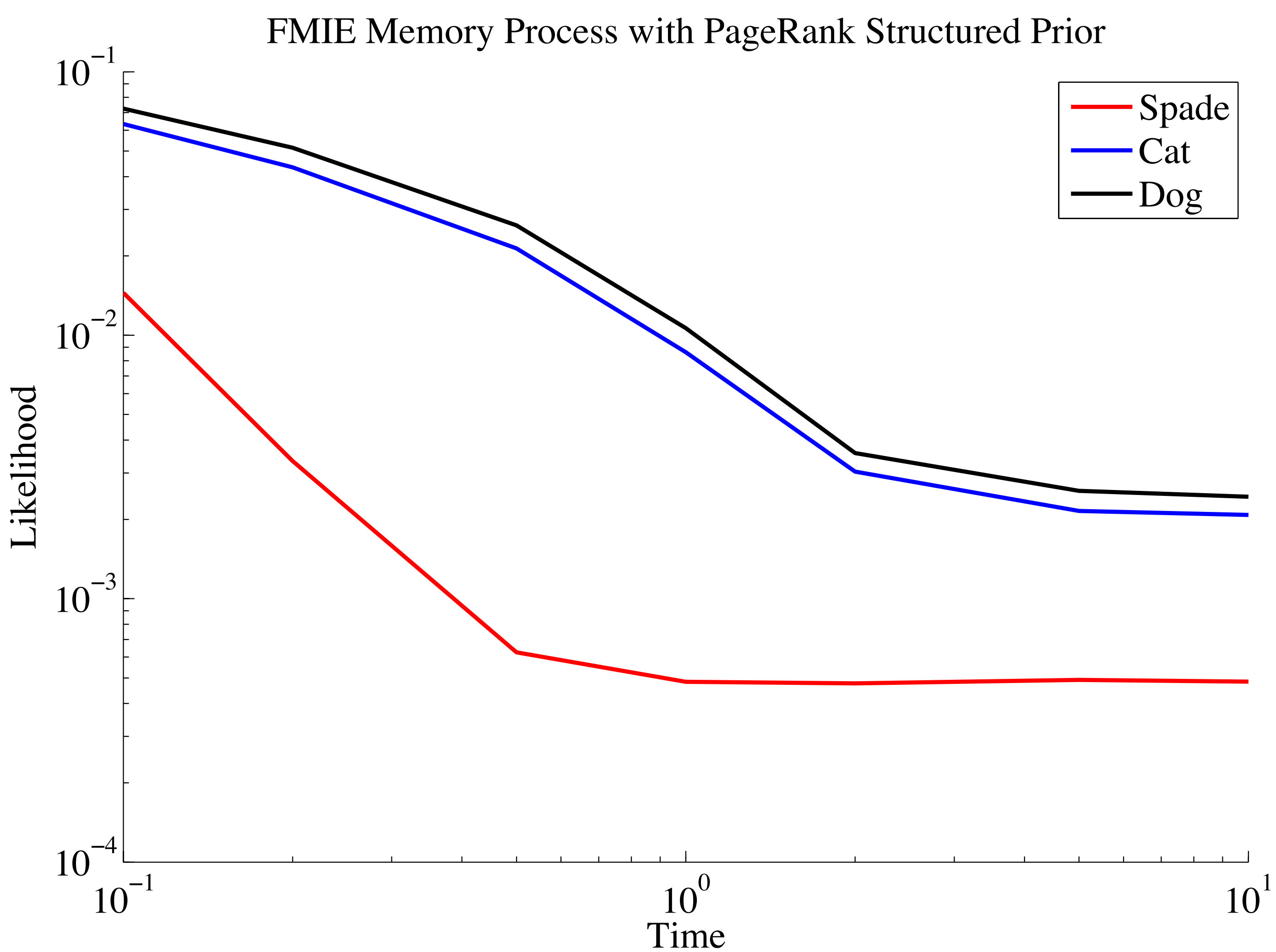
Communication Topology

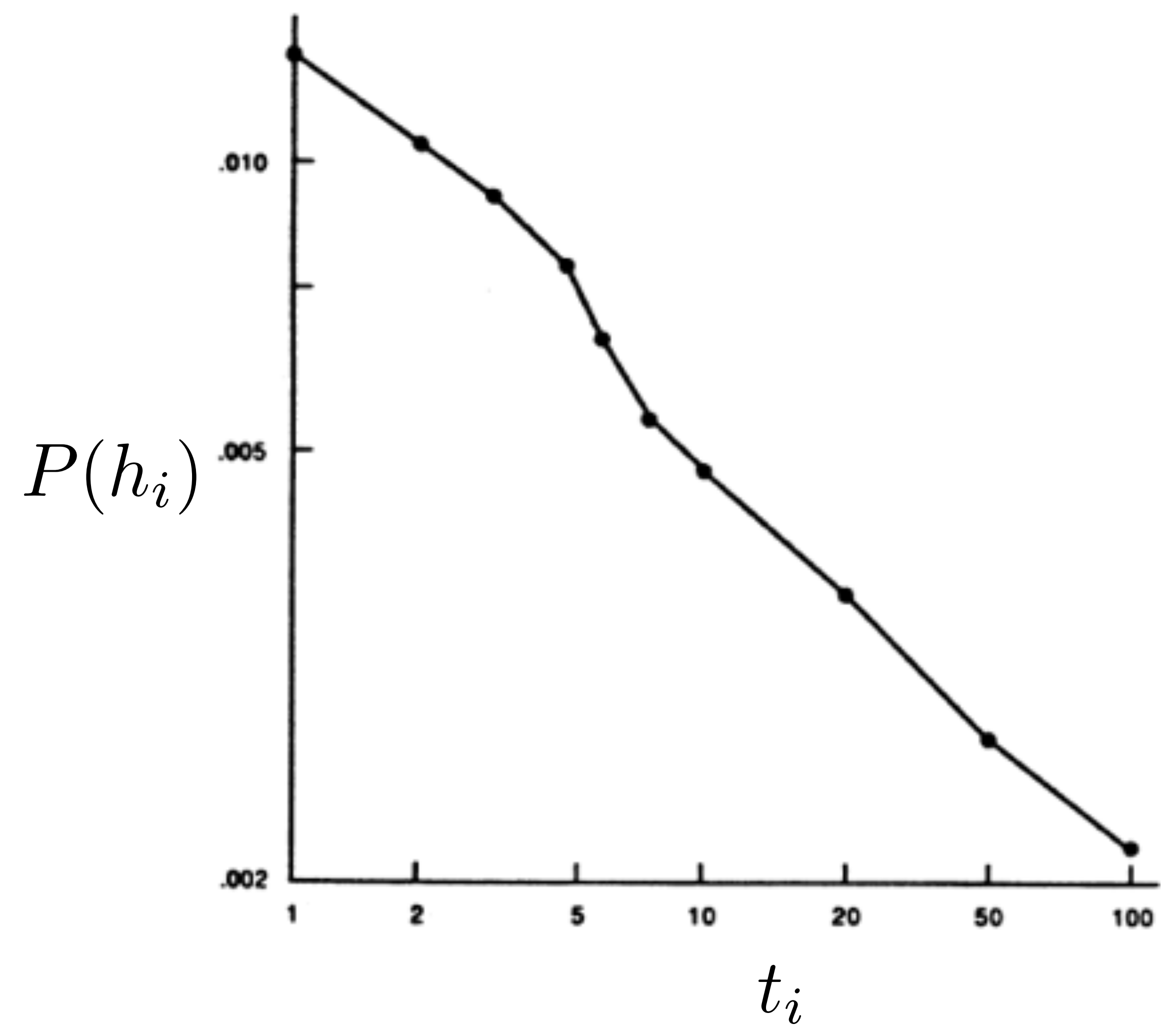
Uses a subset of a larger semantic memory data set:

Includes 10 random cat features and 10 random dog features (and cat and dog).




FMIE Memory Process with PageRank Structured Prior
Spade
Cat
Dog
Likelihood
Time
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-1}
10^{0}
10^{1}


FMIE Memory Process with PageRank Structured Prior
Spade
Cat
Dog
Likelihood
Time
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{0}
10^{1}


$P(h_i)$
.010
.005
.002
1
2
5
10
20
50
100
t_i