Bayesian models of cultural evolution

Tom Griffiths
Department of Psychology
Cognitive Science Program
University of California, Berkeley

QuickTime™ and a TIFF (Uncompressed) decompressor are needed to see this picture.

Explaining inductive leaps

How do people…

- predict the future
- infer causal relationships
- identify the work of chance
- interpret words and sentences
- discover meaningful features of objects
- learn functions, languages, and concepts

. . . from limited data?

The importance of inductive biases

“blicket”

A theory of induction

Posterior probability

Likelihood

Prior probability

$$P(h \mid d) = \frac{P(d \mid h)P(h)}{\sum_{h'} P(d \mid h')P(h')}$$

Sum over space of hypotheses

h: hypothesis
d: data

Predicting the future

A movie has made \$90 million so far…

\$6 million

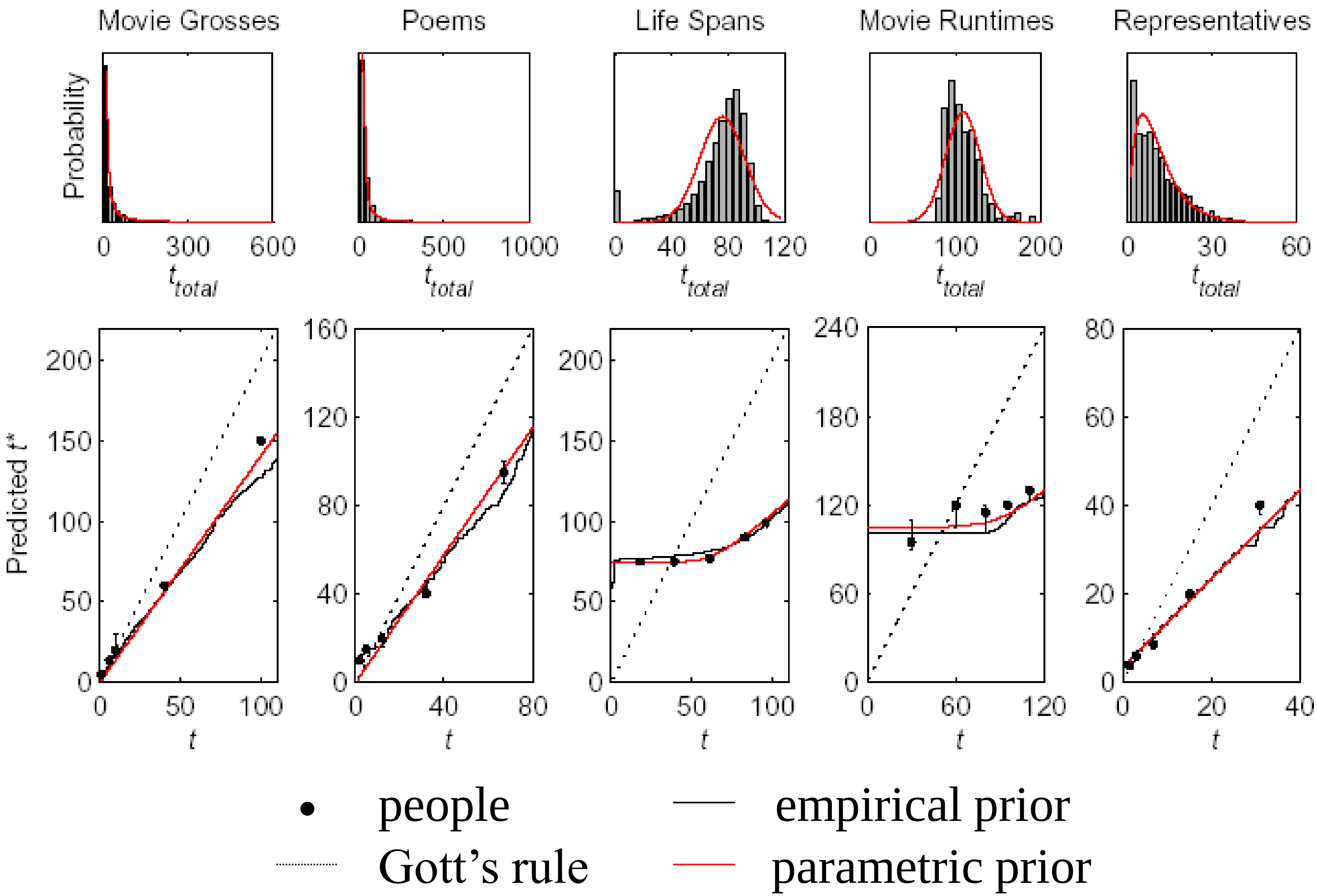
You meet a 90 year old man…

6 year old boy

t = elapsed duration or extent

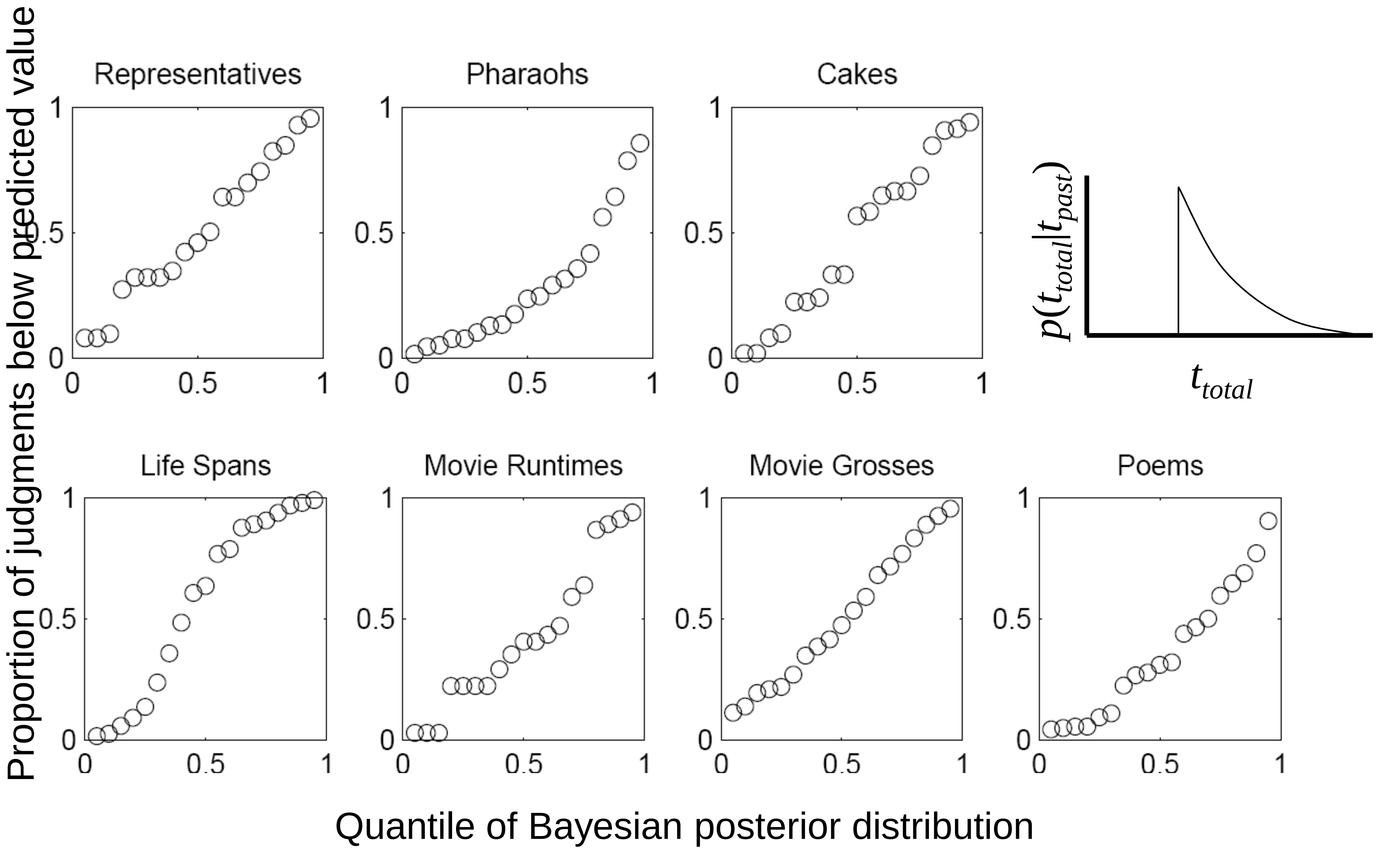
t_{total} = total duration or extent

What should we guess for t_{total} given t?

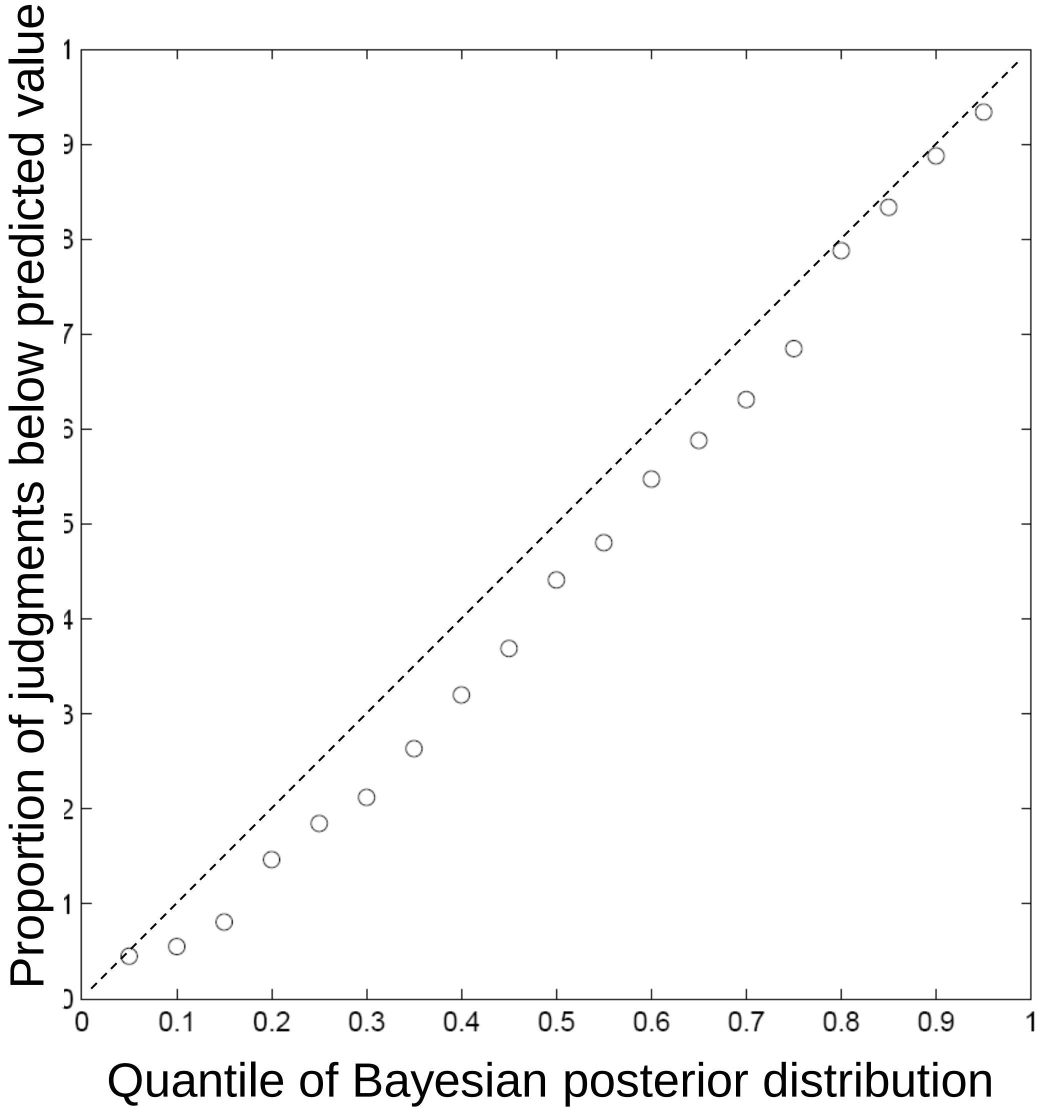


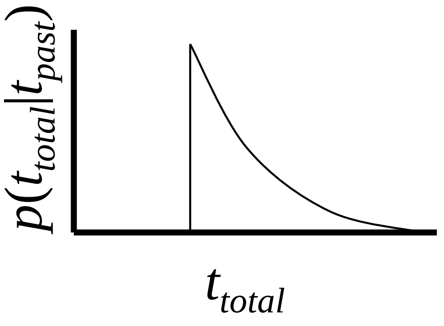
(Griffiths & Tenenbaum, 2006)

Probability matching



Probability matching





Accumulated over all prediction tasks:

- movie run times
- movie grosses
- poem lengths
- life spans
- terms in congress
- cake baking times

Outline

Cultural transmission of information

Cumulative cultural evolution

Creating communication systems

Outline

Cultural transmission of information

Cumulative cultural evolution

Creating communication systems

Iterated learning

(Kirby, 2001)

QuickTime™ and a
TIFF (LZW) decompressor
are needed to see this picture.

What are the consequences of learners learning from other learners?

Objects of iterated learning

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

How do constraints on learning (inductive biases) influence cultural universals?

Analyzing iterated learning

$P_L(h|d)$ $P_L(h|d)$

$P_P(d|h)$ $P_P(d|h)$

QuickTime™ and a
TIFF (LZW) decompressor
are needed to see this picture.

$P_L(h|d)$: probability of inferring hypothesis h from data d

$P_P(d|h)$: probability of generating data d from hypothesis h

Markov chains

$$\longrightarrow \boldsymbol{x} \longrightarrow \boldsymbol{x} \longrightarrow \boldsymbol{x} \longrightarrow \boldsymbol{x} \longrightarrow \boldsymbol{x} \longrightarrow \boldsymbol{x} \longrightarrow \boldsymbol{x} \longrightarrow \boldsymbol{x} \longrightarrow$$

Transition matrix
$P(\boldsymbol{x}^{(t+1)}|\boldsymbol{x}^{(t)})$

- Variables $\boldsymbol{x}^{(t+1)}$ independent of history given $\boldsymbol{x}^{(t)}$
- Converges to a *stationary distribution* under easily checked conditions (i.e., if it is ergodic)

Analyzing iterated learning

$$d_0 \xrightarrow[P_L(h|d)]{} h_1 \xrightarrow[P_P(d|h)]{} d_1 \xrightarrow[P_L(h|d)]{} h_2 \xrightarrow[P_P(d|h)]{} d_2 \xrightarrow[P_L(h|d)]{} h_3 \longrightarrow$$

A Markov chain on hypotheses

$$h_1 \xrightarrow[\Sigma_d P_P(d|h) P_L(h|d)]{} h_2 \xrightarrow[\Sigma_d P_P(d|h) P_L(h|d)]{} h_3 \longrightarrow$$

A Markov chain on data

$$d_0 \xrightarrow[\Sigma_h P_L(h|d)\, P_P(d|h)]{} d_1 \xrightarrow[\Sigma_h P_L(h|d)\, P_P(d|h)]{} d_2 \xrightarrow[\Sigma_h P_L(h|d)\, P_P]{}$$

Iterated Bayesian learning

$P_L(h|d)$ $P_L(h|d)$

$P_P(d|h)$ $P_P(d|h)$

QuickTime™ and a
TIFF (LZW) decompressor
are needed to see this pict

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

Assume learners *sample* from their posterior distribution:

$$P_L(h \mid d) = \frac{P_P(d \mid h) P(h)}{\sum_{h' \in H} P_P(d \mid h') P(h')}$$

Stationary distributions

- Markov chain on h converges to the prior, $P(h)$

- Markov chain on d converges to the "prior predictive distribution"

$$P(d) = \sum_h P(d \mid h) P(h)$$

(Griffiths & Kalish, 2005)

Explaining convergence to the prior

$P_L(h|d)$ $P_L(h|d)$

$P_P(d|h)$ $P_P(d|h)$

QuickTime™ and a TIFF (LZW) decompressor are needed to see this pict

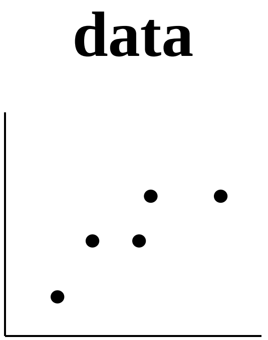
QuickTime™ and a TIFF (Uncompressed) decompressor are needed to see this picture.

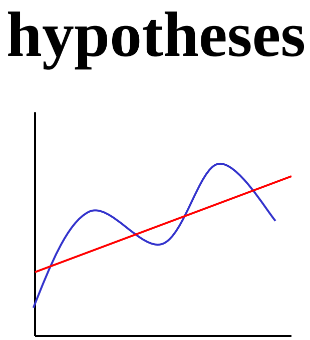
QuickTime™ and a TIFF (Uncompressed) decompressor are needed to see this picture.

- Intuitively: data acts once, prior many times
- Formally: iterated learning with Bayesian agents is a *Gibbs sampler* on $P(d,h)$

(Griffiths & Kalish, 2007)

Iterated function learning

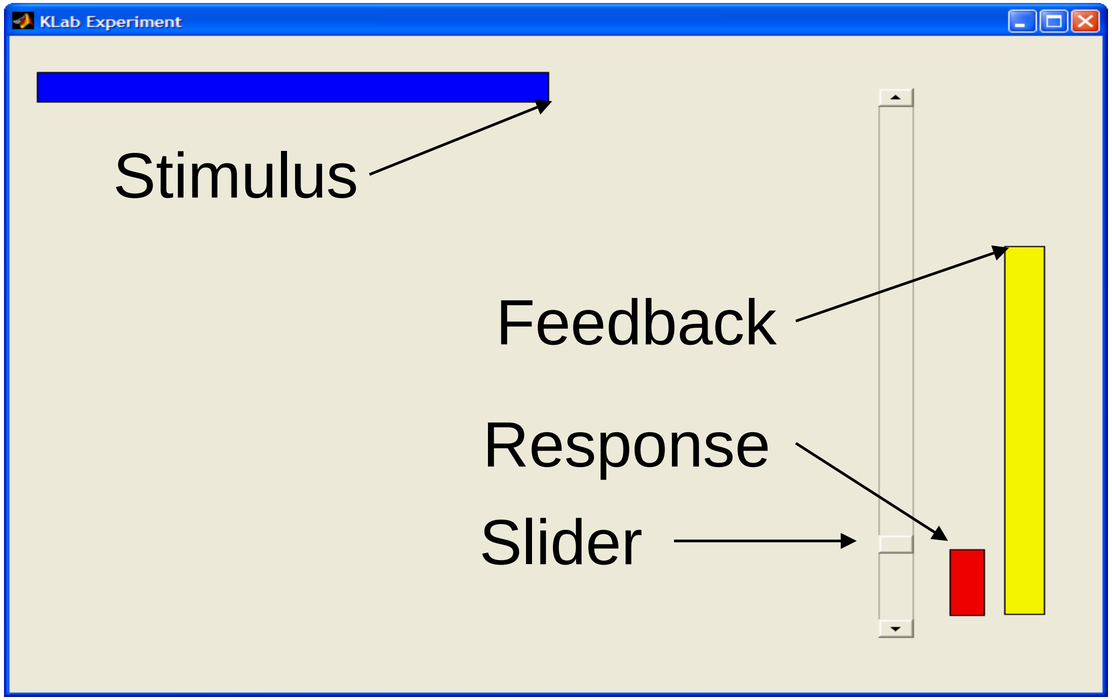




- Each learner sees a set of (*x*,*y*) pairs
- Makes predictions of *y* for new *x* values
- Predictions are data for the next learner

(Kalish, Griffiths, & Lewandowsky, 2007)

Function learning experiments

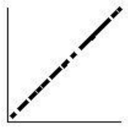


Examine iterated learning with different initial data

Initial data

Iteration

1 2 3 4 5 6 7 8 9



Iterated predicting the future

data	**hypotheses**
A movie has made $30 million so far	$60 million total

- Each learner sees values of t
- Makes predictions of t_{total}
- The next value of t is chosen from $(0, t_{total})$

(Lewandowsky, Griffiths & Kalish, submitted)

Movie grosses

Poems

Chains of predictions

t_{total}

QuickTime
decompr
are needed to see ture.

t_{total}

Iteration

Iteration

(Lewandowsky, Griffiths, & Kalish, submitted)

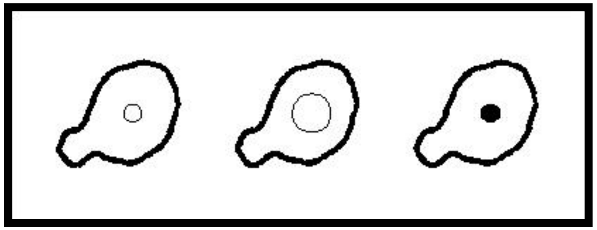
Iterated predicting the future

QuickTime™ and a
TIFF (LZW) decompressor
are needed to see this picture.

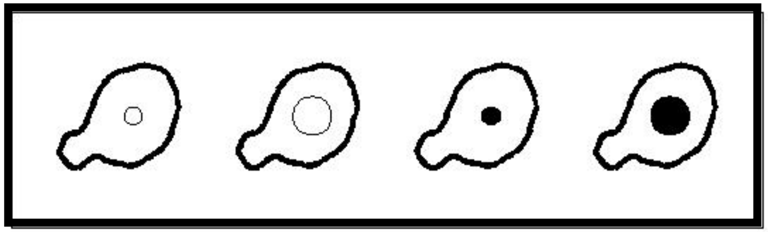
(Lewandowsky, Griffiths, & Kalish, submitted)

Iterated concept learning

data



hypotheses

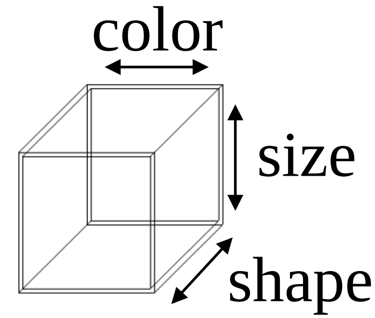


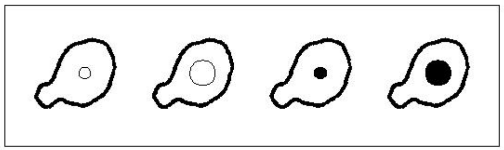
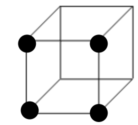
(stimuli from Feldman, 2000)

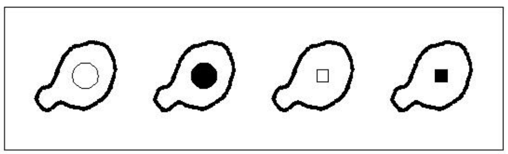
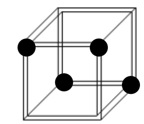
- Each learner sees examples from a species
- Identifies species of four amoebae
- Species correspond to boolean concepts

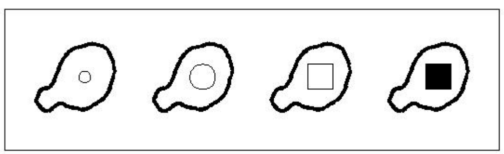
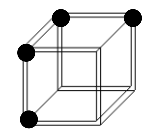
(Griffiths, Christian, & Kalish, 2006)

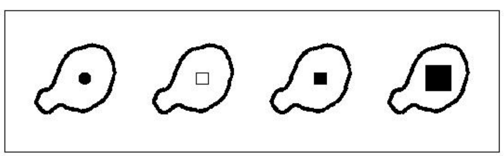
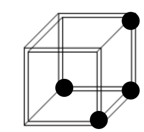
Types of concepts
(Shepard, Hovland, & Jenkins, 1961)

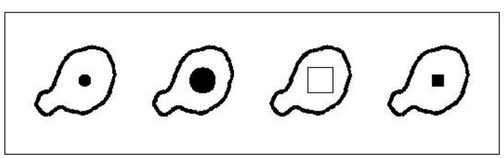
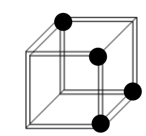


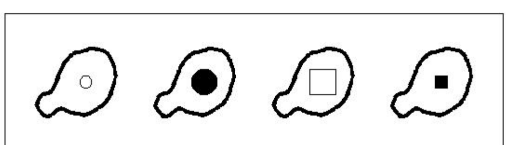
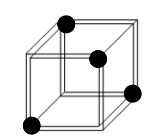

Type I



Type II


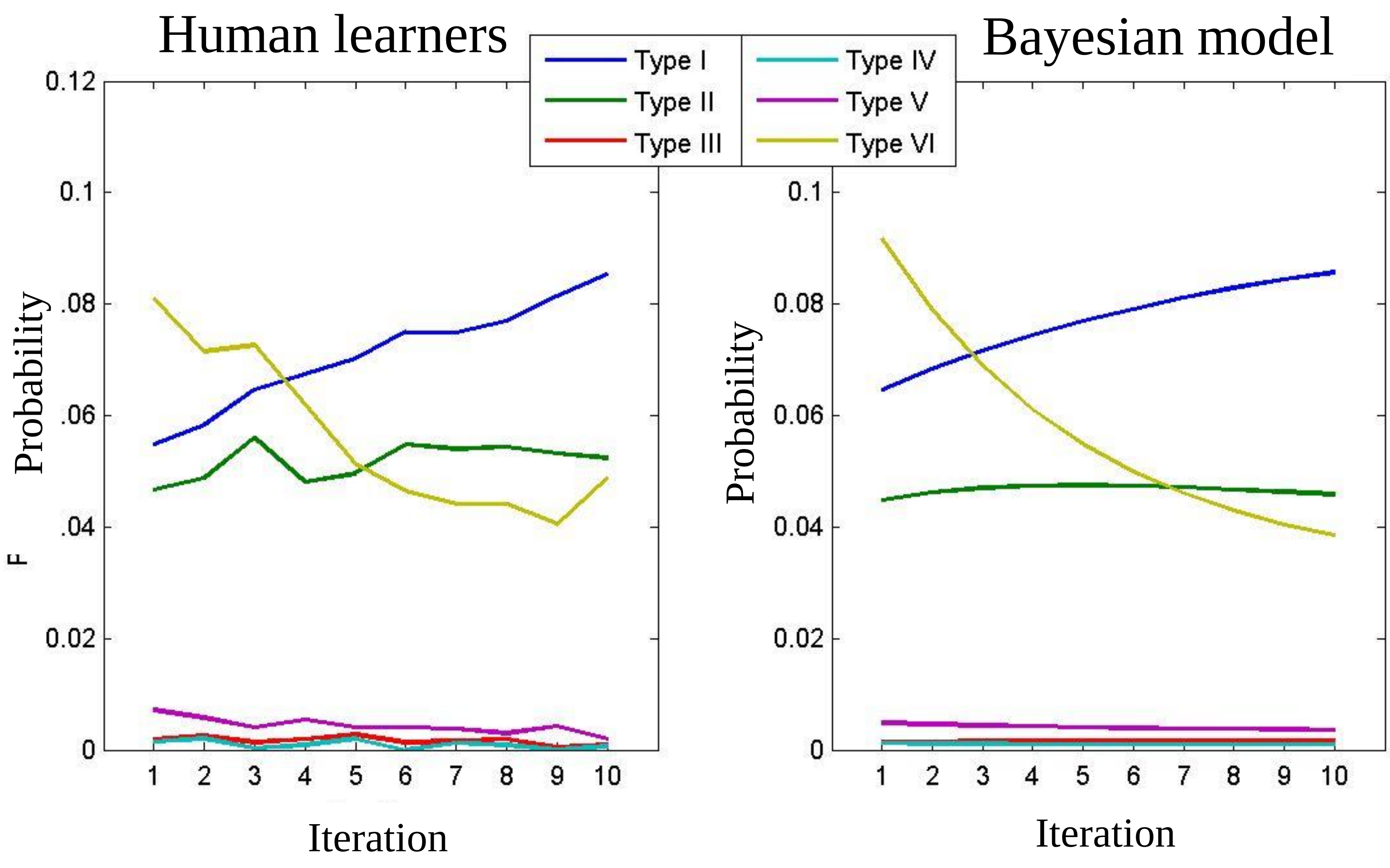

Type III



Type IV



Type V



Type VI


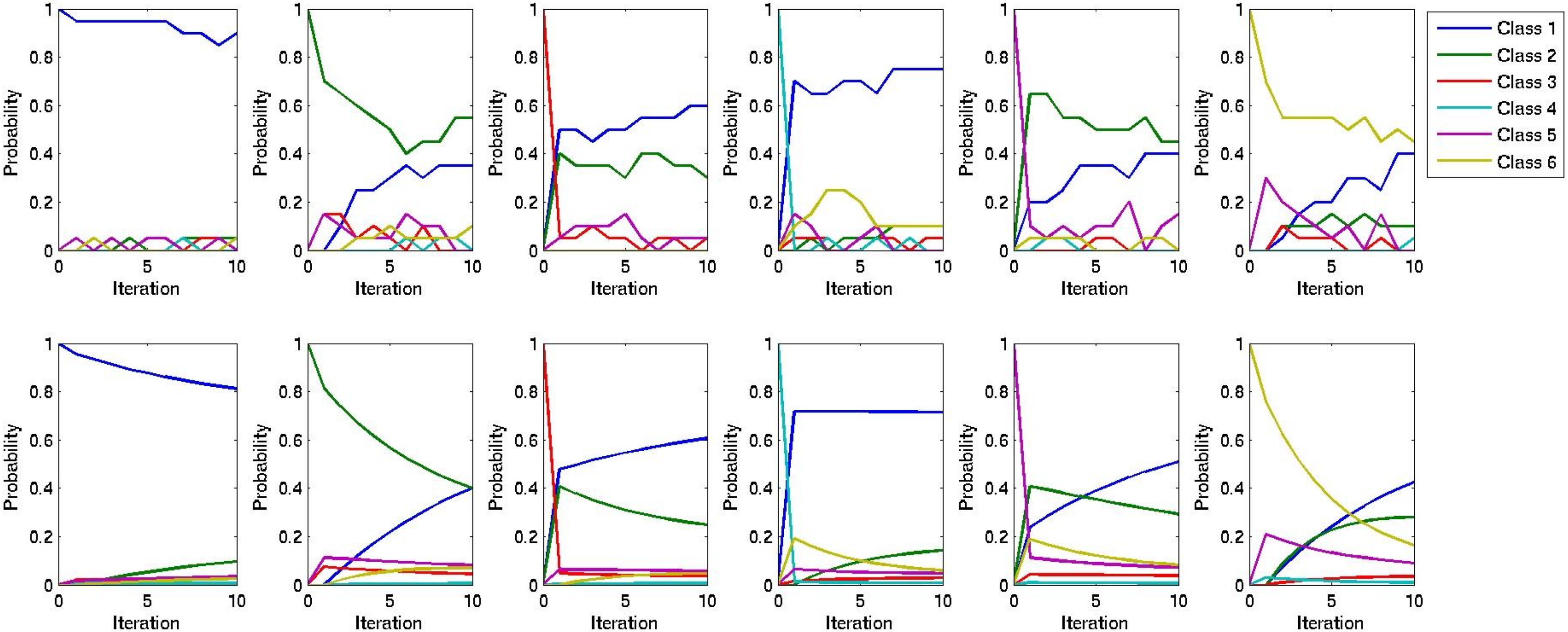
Results



Human learners
Bayesian model
Type I
Type II
Type III
Type IV
Type V
Type VI
Probability
Iteration
0.12
0.1
.08
.06
.04
0.02
0
0.08
0.06
0.04
1
2
3
4
5
6
7
8
9
10

Three positive examples

Human learners



Bayesian model

Outline

Cultural transmission of information

Cumulative cultural evolution

Creating communication systems

Making progress

? ? ?

QuickTime™ and a
decompressor
are needed to see this picture.

What needs to be transmitted between generations to support *cumulative* cultural evolution?

Formalizing the problem

- A sequence of Bayesian agents
- Each receives
  - a message from the previous agent
  - data d^* from the world (generated from $P(d^*)$)
- Selects a hypothesis h by applying Bayes' rule

- What kinds of messages result in the ultimate selection of hypotheses that best match $P(d^*)$?

(Beppu & Griffiths, in prep)

Observational learning…

QuickTime™ and a
decompressor
are needed to see this picture.

…is insufficient

- Markov chain on h converges to the average of the posterior over d^*, $\Sigma_{d*} P(h|d^*)P(d^*)$
- Asymptotic distribution over hypotheses is equivalent to a single learner observing d^*
  - no cumulative advantage of cultural transmission

(Beppu & Griffiths, in prep)

Transmitting theories…

QuickTime™ and a
decompressor
are needed to see this picture.

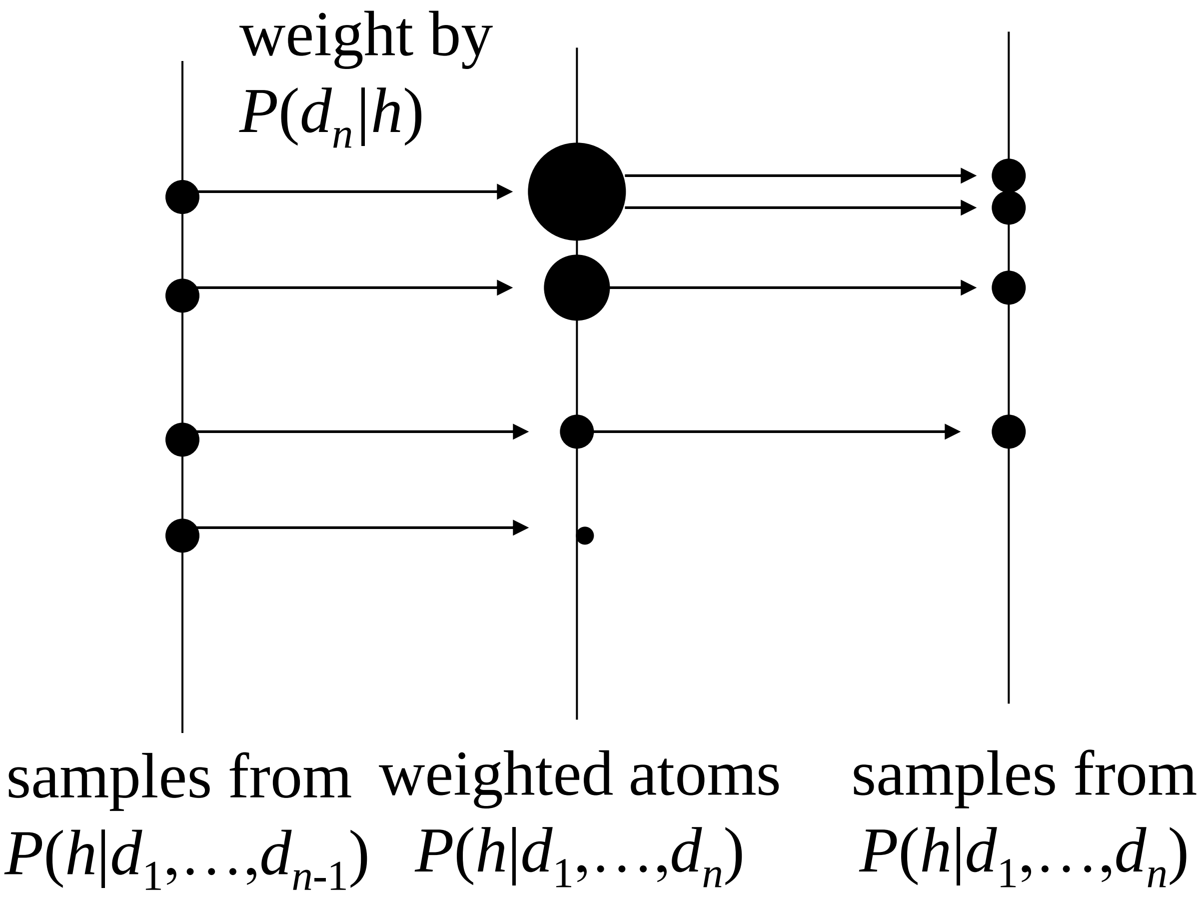
…is sufficient

- Convergence to the hypothesis h with $P(d|h)$ closest to $P(d^*)$ (in KL divergence) for
  - transmission of full posterior distribution
  - transmission of n samples from posterior distribution (via convergence of particle filters)

(Beppu & Griffiths, in prep)

Particle filters

$$P(h \mid d_1, \ldots, d_n) \propto P(d_n \mid h) P(h \mid d_1, \ldots, d_{n-1})$$



…is sufficient

- Convergence to the hypothesis h with $P(d|h)$ closest to $P(d^*)$ (in KL divergence) for
  - transmission of full posterior distribution
  - transmission of n samples from posterior distribution (via convergence of particle filters)
- These possibilities correspond to reasonable cultural practices…

(Beppu & Griffiths, in prep)

Cumulative cultural evolution in the lab

- Iterated function learning experiment, varying data from the world and type of message
- Three conditions:
  - *no data*, just iterated learning (2 chains)
  - *mixed data*, observational learning (10 chains)
  - *theory*, message typed in a box (11 chains)
- Eight participants per chain
- True function was quadratic

(Beppu & Griffiths, in prep)

Representative chains

training data
responses

no
data

(Beppu & Griffiths, in prep)

All chains

no data

mixed data

theory

Distance from true function

QuickTime
decompr
are needed to see
e.

Cumulative mean distance

Generation

Generation

(Beppu & Griffiths, in prep)

Conclusions

- Simple Bayesian models can provide insight into complex processes related to cultural evolution
- When cognition affects culture, studying groups can give us better insight into individuals

data

QuickTime™ and a
TIFF (LZW) decompressor
are needed to see this picture.

Credits

Joint work with…

Iterated learning

Mike Kalish

Steve Lewandowsky

Cumulative evolution

Aaron Beppu

Creation

Linsey Smith

Middy Pineda

QuickTime™ and a
TIFF (LZW) decompressor
are needed to see this picture.

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

Computational Cognitive Science Lab

http://cocosci.berkeley.edu/