Machine Learning for Analyzing Human Brain Function

Tom M. Mitchell
Center for Automated Learning and Discovery
Carnegie Mellon University

PAKDD 2005

Collaborators: Rebecca Hutchinson, Marcel Just, Francisco Pereira, Jay Pujara, Indra Rustandi, Wei Wang

KDD for the Sciences

- Empirical science has changed forever
  - Online data capture
  - Huge data sets
  - Web-based data collections
  - Machine learning
- Human genome project
  - HMMs for analyzing gene sequences,…
- Sloan sky survey
  - Unsupervised clustering of galaxies, stars,…
- Cell biology
  - Bayesian network models of gene expression

Human Brain Imaging

- fMRI Location with millimeter precision

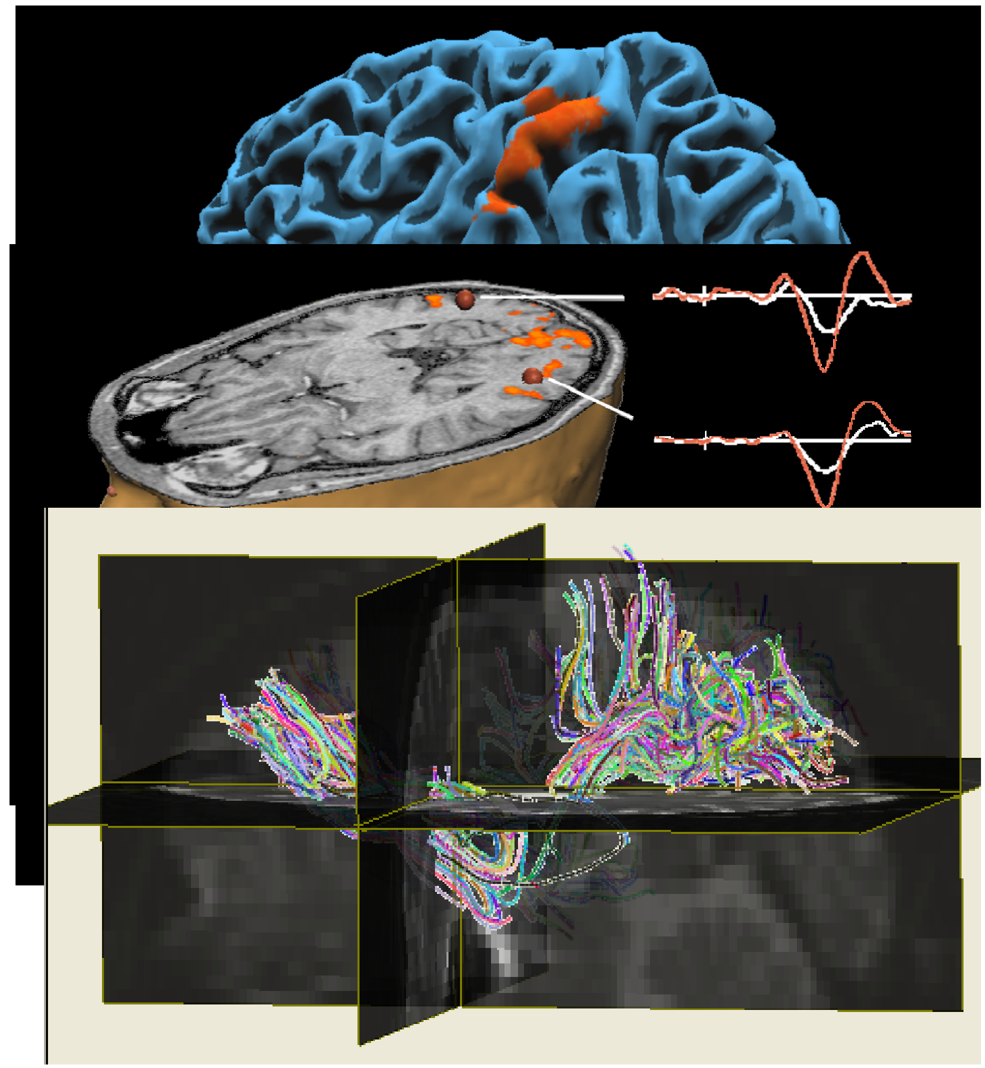
1 mm = 0.0004% of cortex

- ERP Time course with millisecond precision

10 ms = 10 % of human production cycle

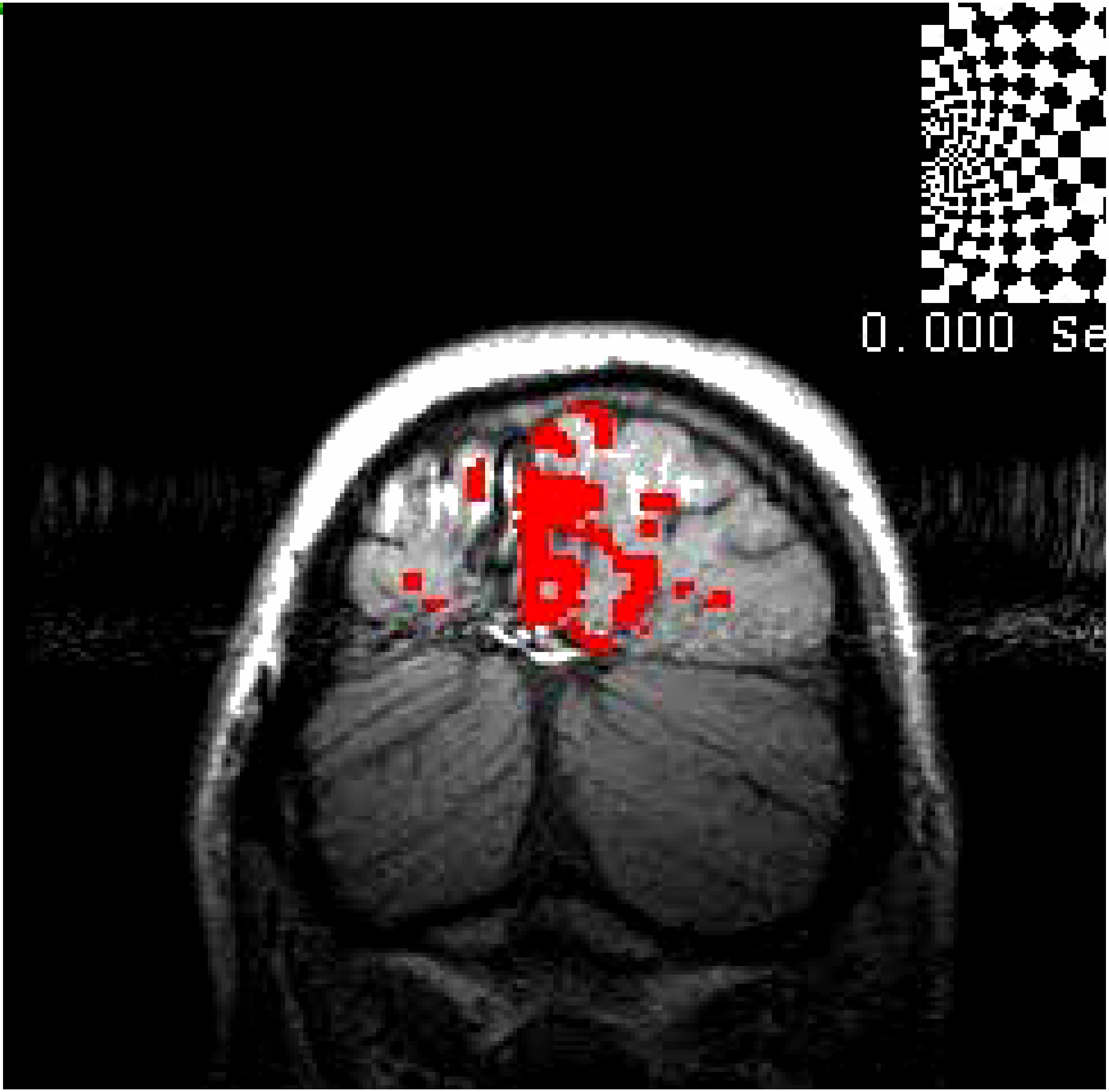
- DTI Connections tracing millimeter precision

1mm connection ~10k fibers, or 0.0001% of neurons



Functional MRI

Brain scans can track activation with precision and sensitivity



functional Magnetic Resonance Imaging (fMRI)

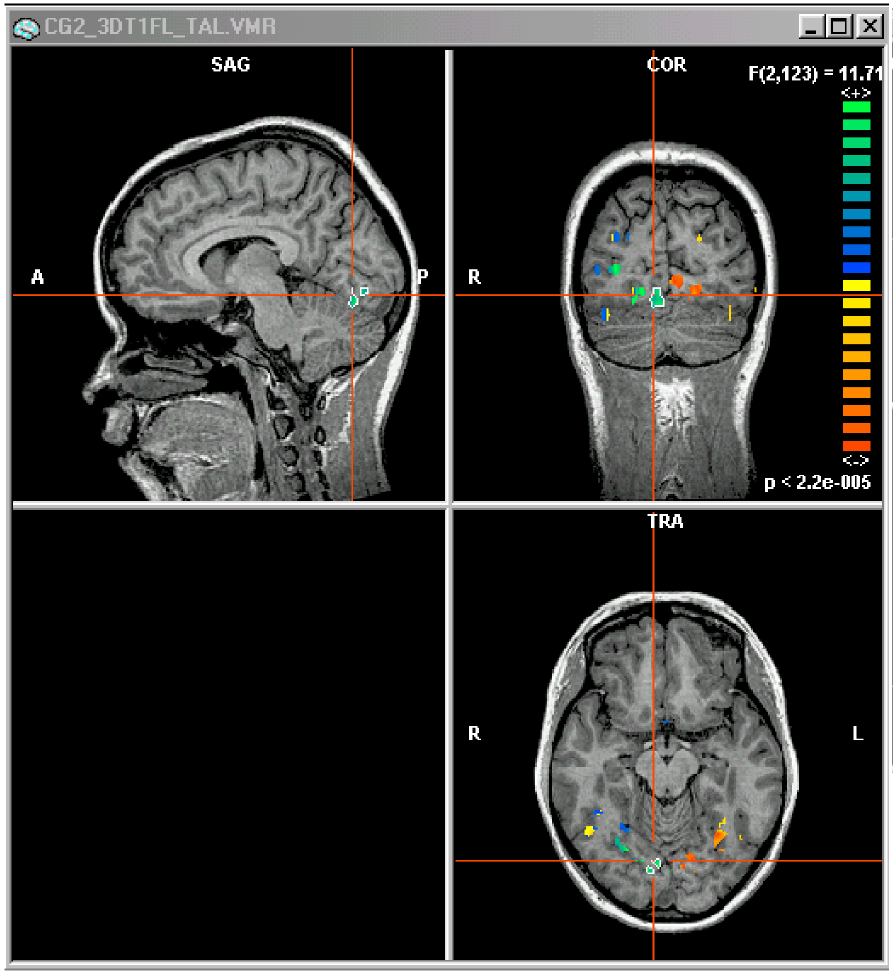
~1 mm resolution

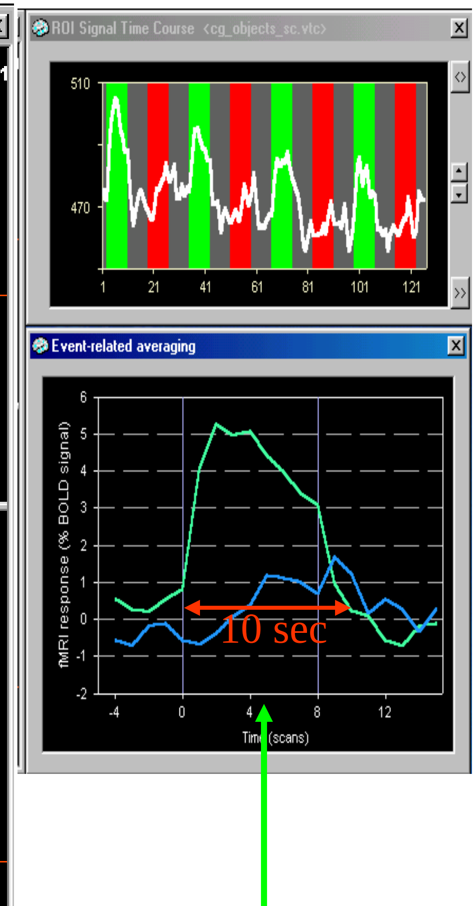
~2 images per sec.

15,000 voxels/image

non-invasive, safe

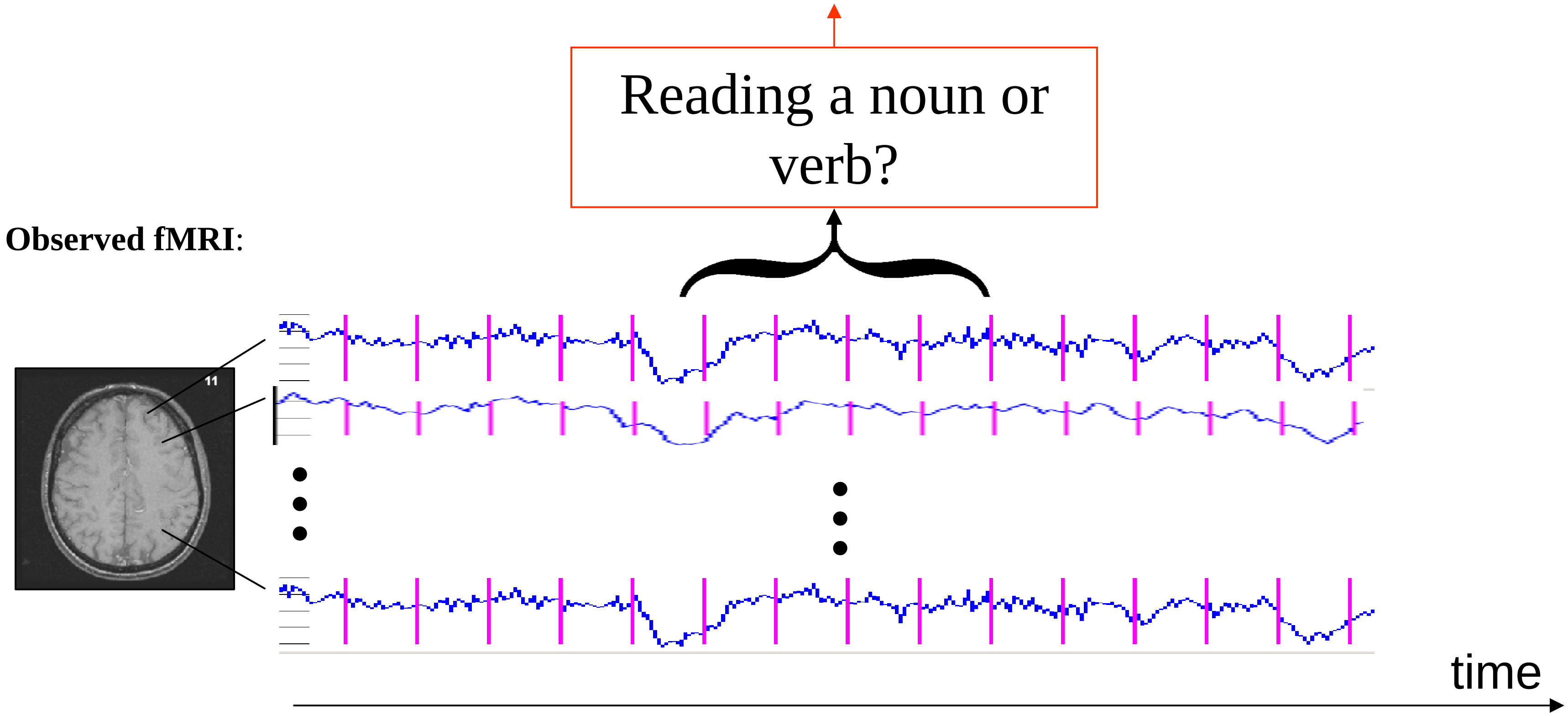
measures Blood Oxygen Level Dependent (BOLD) response





Typical impulse response

1. Can we distinguish brief cognitive processes using fMRI?



Training Classifiers of Cognitive State

Train classifiers of form:

$\text{fMRI}(t, t+\delta) \rightarrow \text{CognitiveStates}$

e.g., $\text{fMRI}(t, t+7) = \{\text{ReadNoun}, \text{ReadVerb}\}$

Initially:

- Fixed set of cognitive states
- Fixed time interval $[t, t+\delta]$

Study 1: Word Categories

[with Francisco Pereira, Marcel Just]

For each category (vegetables, tools, trees, fish, dwellings, building parts):

- Presented blocks of 20 words from single* category
- Two blocks of each category
- Blocks from different word categories interleaved

Training Classifier for Word Categories

Learn fMRI(t,t+20) $\rightarrow$ word-category(t,t+20)

– fMRI(t1,t2) = 10^4 voxels, mean activation of each during interval [t1 t2]

Training methods:

– train single-subject classifiers
– Gaussian Naïve Bayes $\rightarrow$ P(fMRI | word-category)
– Nearest nbr with spatial-correlation as distance
– SVM

Feature selection: Select *n* voxels

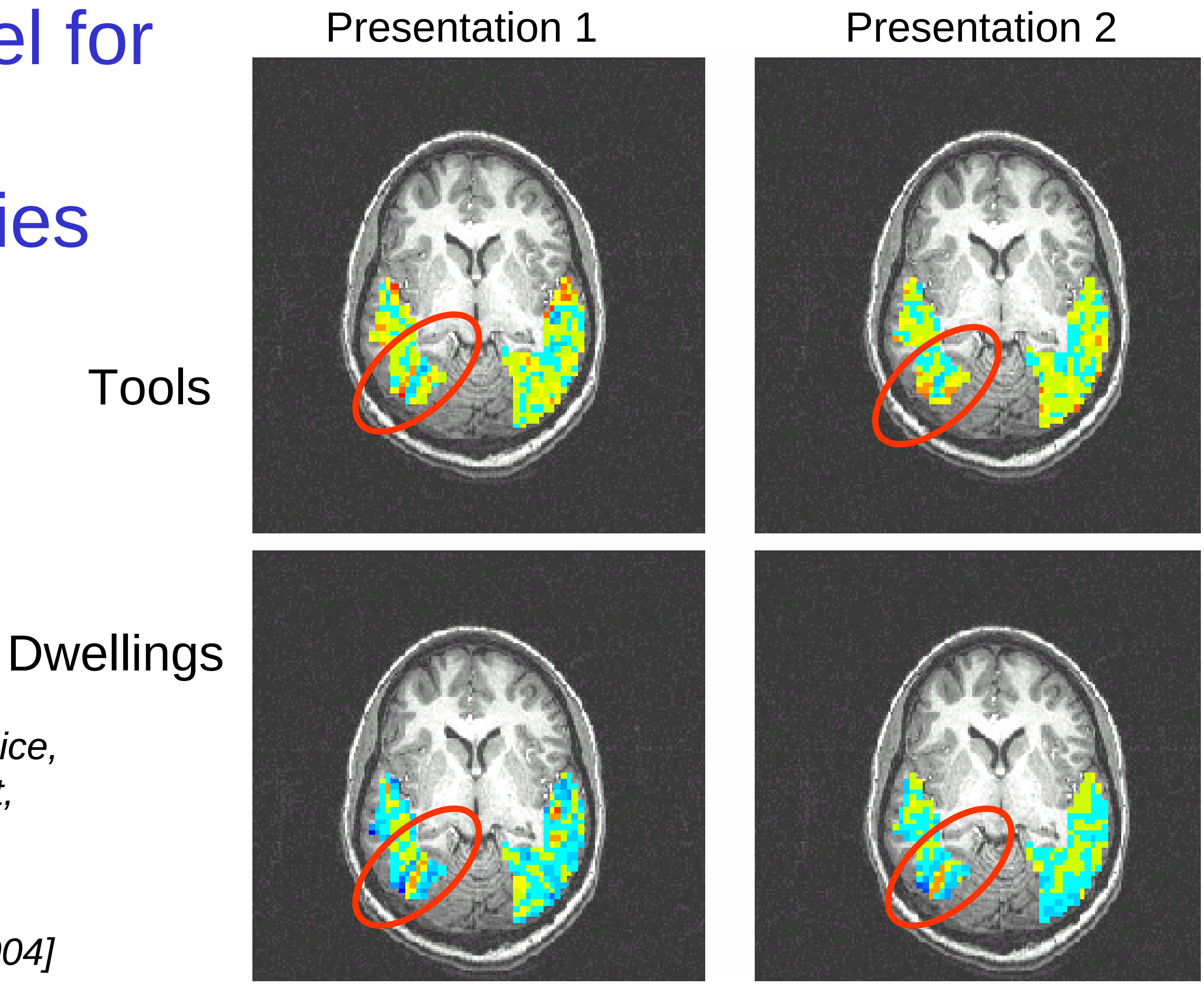
– Reduce 10^4 voxels to 10^2

Mean Activation per Voxel for Word Categories

Mean classification accuracy .90 (tools vs dwellings, 7 subjects)

one horizontal slice, from one subject, ventral temporal cortex

[Pereira, et al 2004]



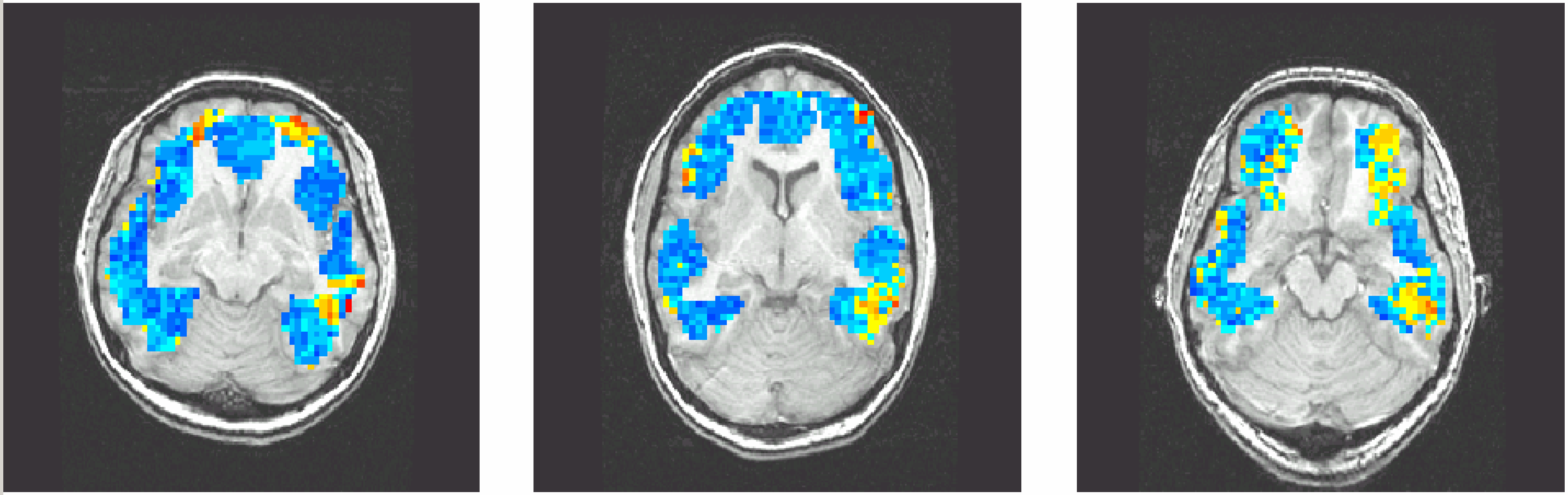
Question:

Where in the brain is the activity that discriminates word category?

Plot of single-voxel classification accuracies.

Gaussian naïve Bayes classifier

(yellow and red are most predictive).



Subject 1 Subject 2 Subject 3

Question:

Do different people's brains
'encode' semantic categories
using the same spatial patterns?

No.

But, there are cross-subject regularities in "distances" between categories, as measured by classifier error rates.

Six-Category Study: Pairwise Classification Errors (ventral temporal cortex)

* Worst * Best

	Fish	Vegetables	Tools	Dwellings	Trees	Bldg Parts
Subj1	.20	.55 *	.20	.15	.15	.05 *
Sub2	.10 *	.55 *	.35	.20	.10 *	.30
Sub3	.20	.35 *	.15 *	.20	.20	.20
Sub4	.15	.45 *	.15	.15	.25	.05 *
Sub5	.60 *	.55	.25	.20	.15 *	.15 *
Sub6	.20	.25	.00 *	.30 *	.30 *	.05
Sub7	.15	.55 *	.15	.25	.15	.05 *
Mean	**.23**	**.46**	**.18**	**.21**	**.19**	**.12**

Lessons Learned

Yes, one can train machine learning classifiers to distinguish a variety of mental states

- Nouns about “tools” vs. nouns about “building parts”
- Noun vs. Verb
- Ambiguous sentence vs. unambiguous
- Picture vs. Sentence

Failures too:

- True vs. false sentences
- Negative vs. affirmative sentences

ML methods:

- NNbr, Naïve Bayes, SVMs, NNets, …
- Case study in high dimensional, noisy classif [MLJ 2004]
- New approaches to feature selection

Which Feature Selection Strategy to Use?

Wish to learn F: $<x_1,x_2,\ldots x_n> \rightarrow \{A,B\}$

- Feature selection often reduces error 30-40%
- Conventional wisdom: pick features x_i that best distinguish between classes A and B
  - E.g., sort x_i by mutual information with target class
- Surprise:

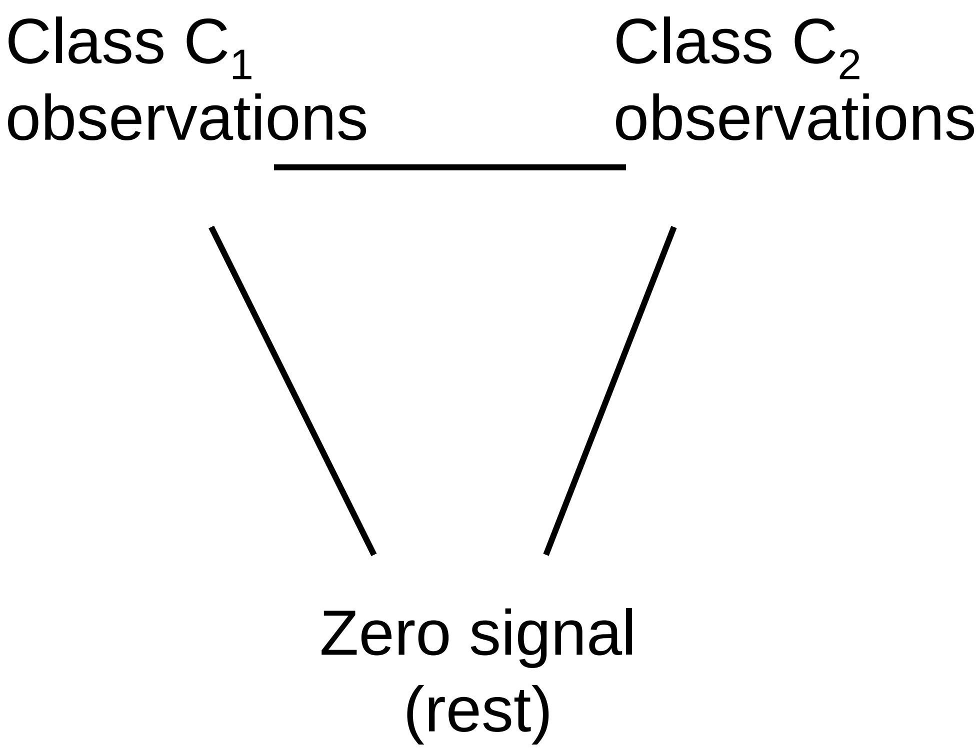
  Alternative strategy works better (23/28 subjects)

  → We have three types of data: subject is performing task A,task B, or resting

  → Pick features that distinguish whether or not subject is resting

“Zero Signal” learning setting.

Select features based on discrim(C_1,C_2) or discrim(Z,C_i)?



“Zero Signal” learning setting.

Select features based on discrim(C_1,C_2) or discrim(Z,C_i)?

Class C_1 observations

$X_1=S_1+N_1$

Class C_2 observations

$X_2=S_2+N_2$

$Z = N_0$

Zero signal (Rest)

Goal: learn f: $X \rightarrow Y$ or $P(Y|X)$

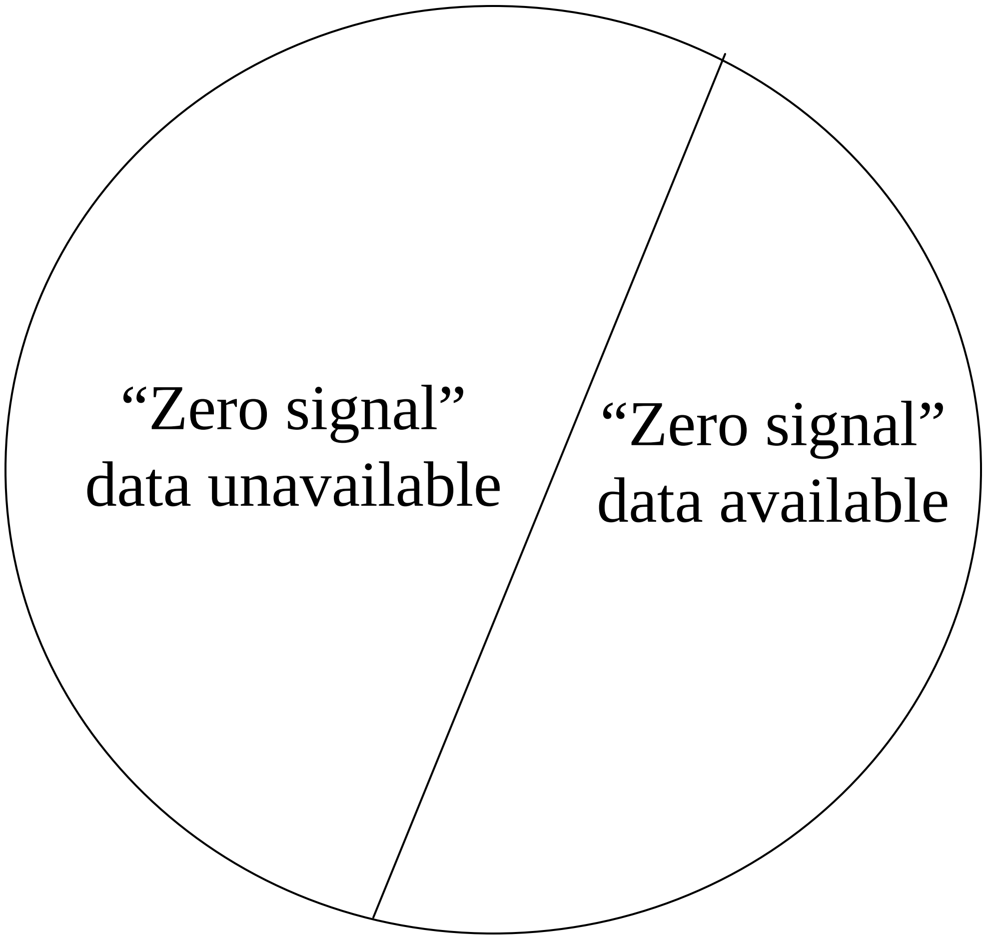
Given:

1. Training examples $<X_i, Y_i>$ where $X_i = S_i + N_i$, signal $S_i \sim P(S|Y= Y_i)$, noise $N_i \sim P_{noise}$
2. Observed noise with zero signal $N_0 \sim P_{noise}$

Discrim(C_1,C_2) or Discrim(C_i,Rest)?

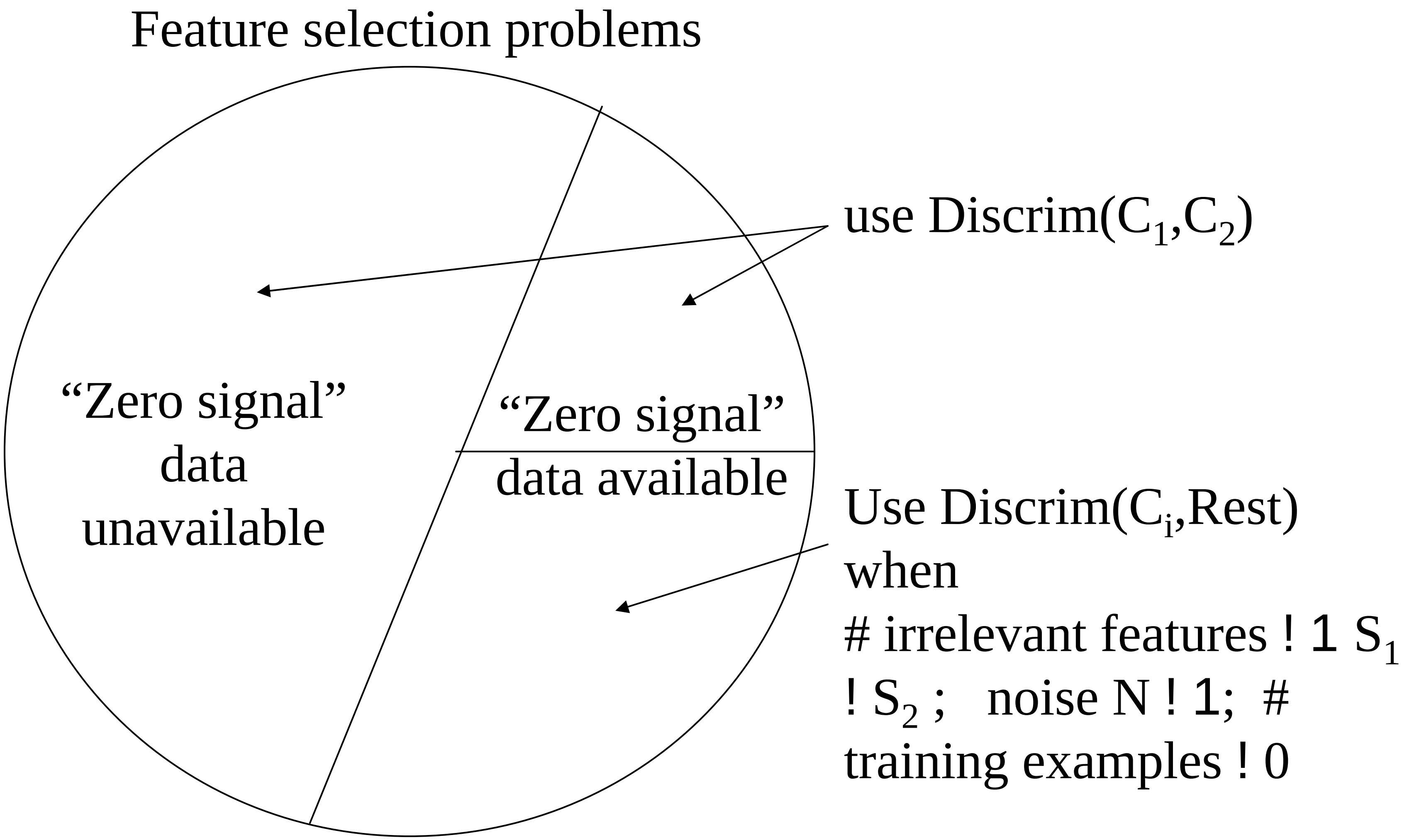
[Jay Pujara, 2005]

Feature selection problems

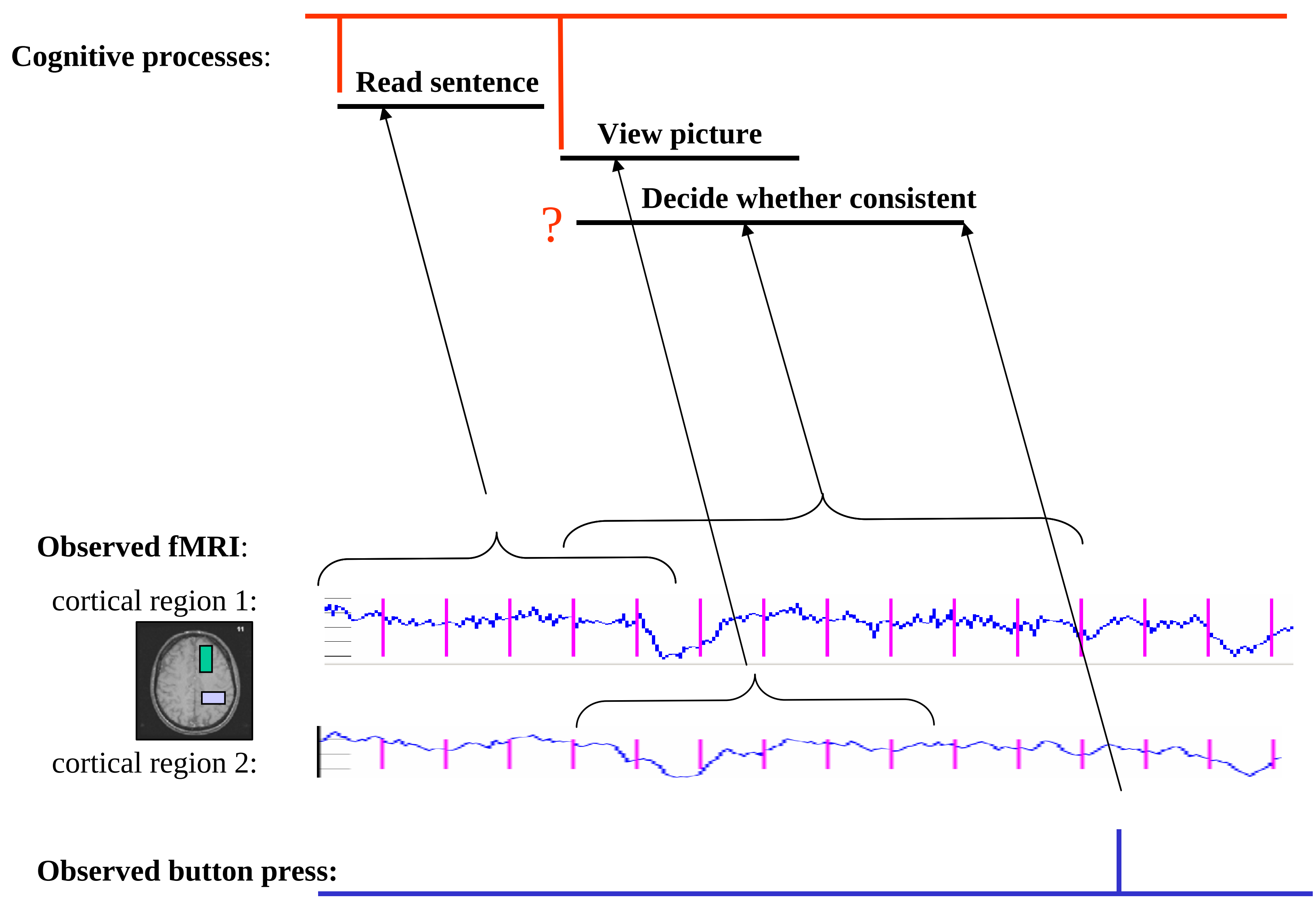


Discrim(C_1,C_2) or Discrim(C_i,Rest)?

[Jay Pujara, 2005]



2. Can we classify/track *multiple overlapping* processes?

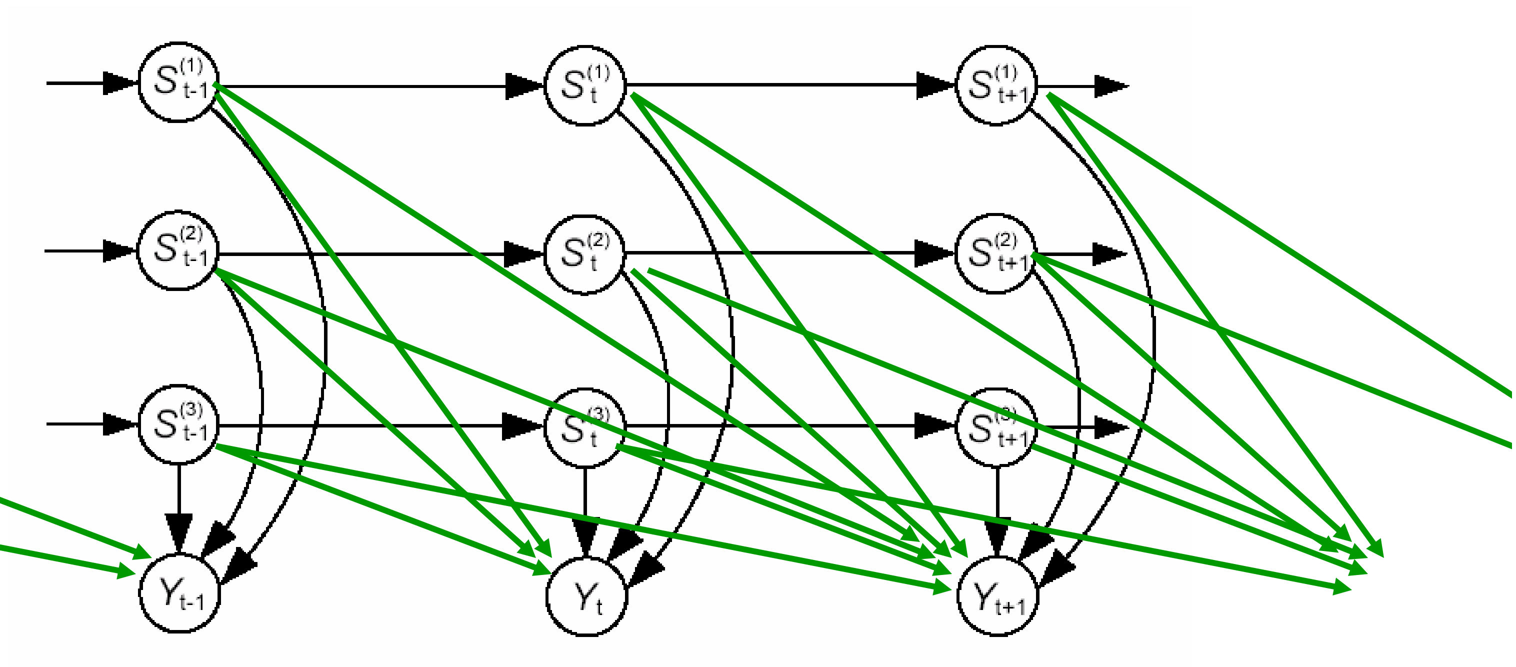


Bayes Net related State-Space Models

HMM's, DBNs, etc. e.g., [Ghahramani, 2001]

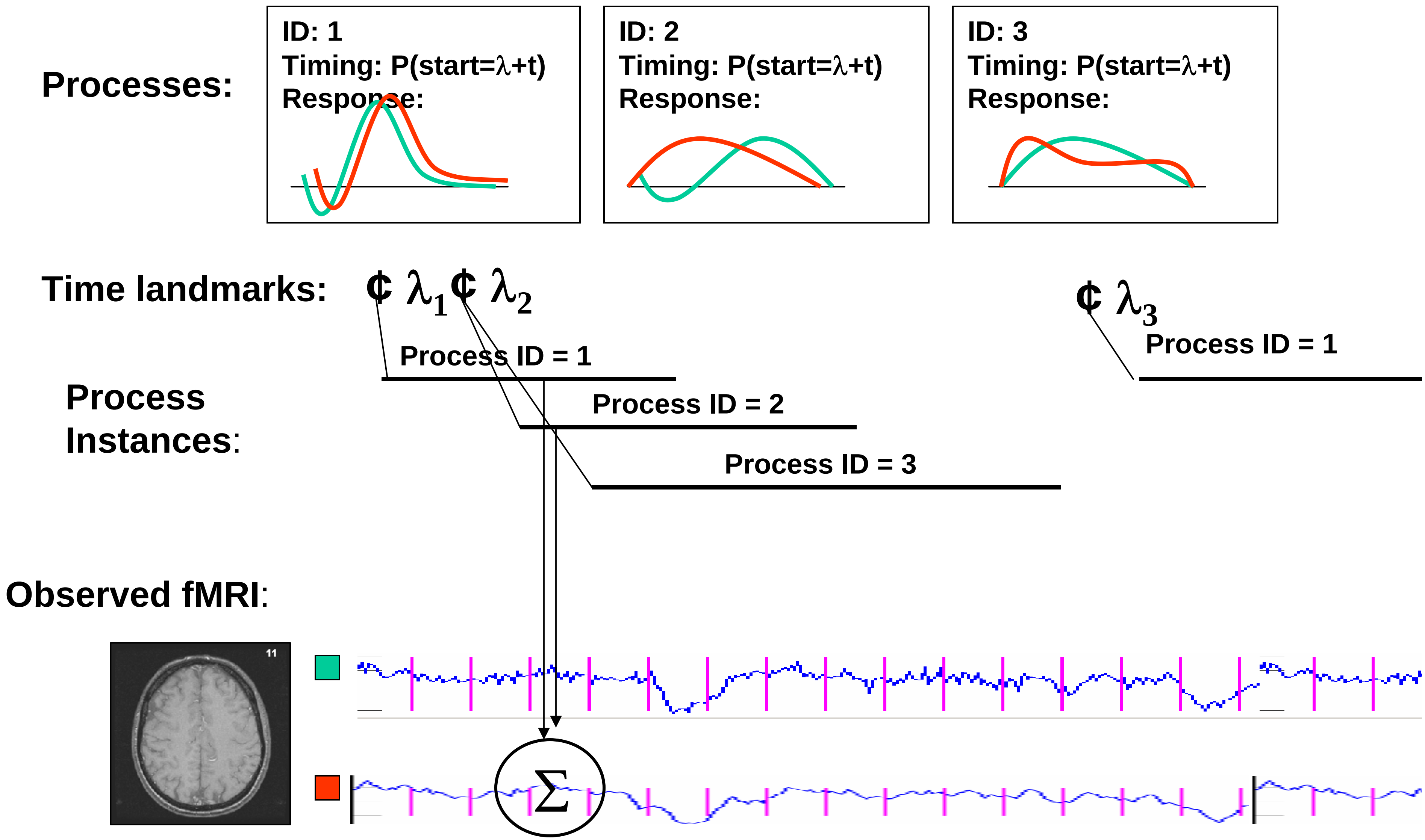
Cognitive subprocesses / state variables:

fMRI:

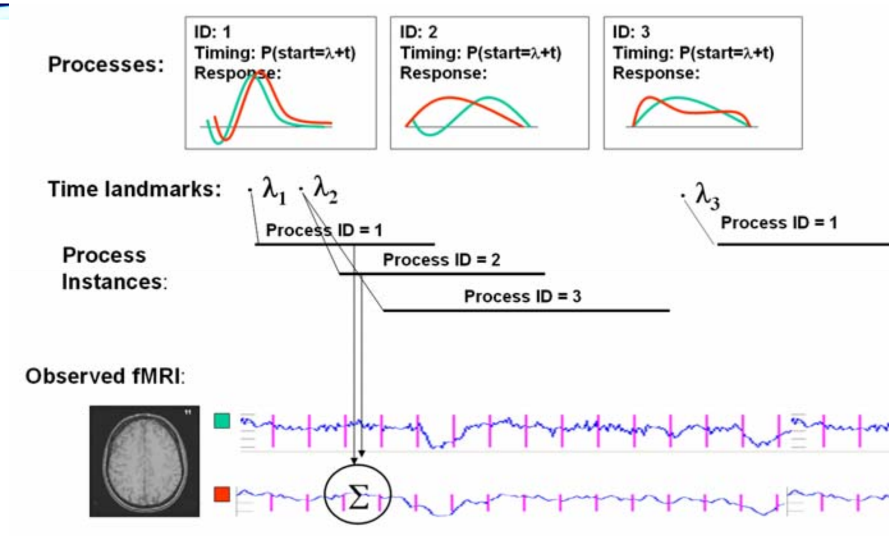


see [Hojen-Sorensen et al, NIPS99]

Hidden Process Models



Hidden Process Models



- Probability of fMRI observation $y_{v,t}$

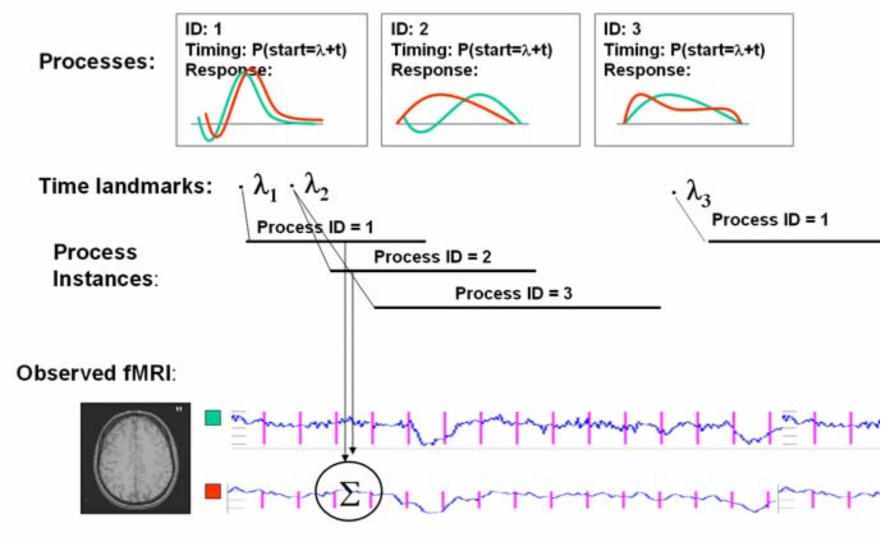
$$P(y_{v,t}|HPM, INT) = N(\mu_{v,t}, \sigma_v)$$

learned

where

$$\mu_{v,t} = \sum_{i \in \text{active process instances}} W_i(v, t - start(i))$$

Hidden Process Models

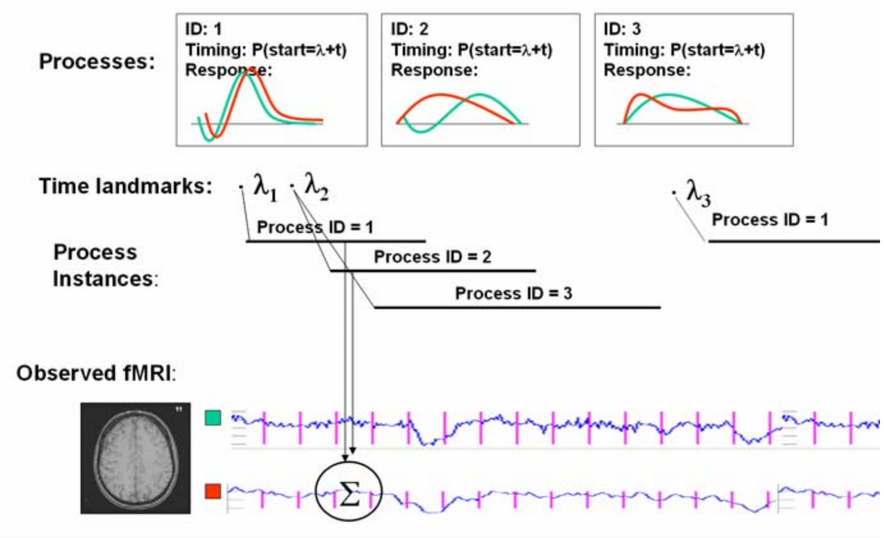


- Probability of process instance k beginning at time t

$$P(start_k = t) = \theta_{t-\lambda}$$

learned

Learning HPMs



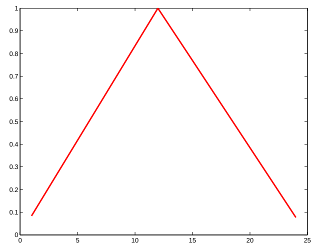
- When process IDs, start times *known*:
  - Least squares regression, eg. Dale[HBM,1999]
  - Ordinary least sq if assume noise indep over time
  - Generalized least sq if assume autocorrelated noise
- When start times *unknown*:

  EM algorithm
  - Repeat:
    - E: estimate probability distribution over start times
    - M: choose parameters to maximize expected data likelihood

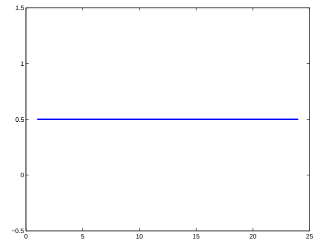
Synthetic Noise-Free Data Example

Process responses:

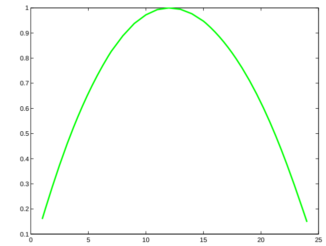
Process 1:



Process 2:



Process 3:



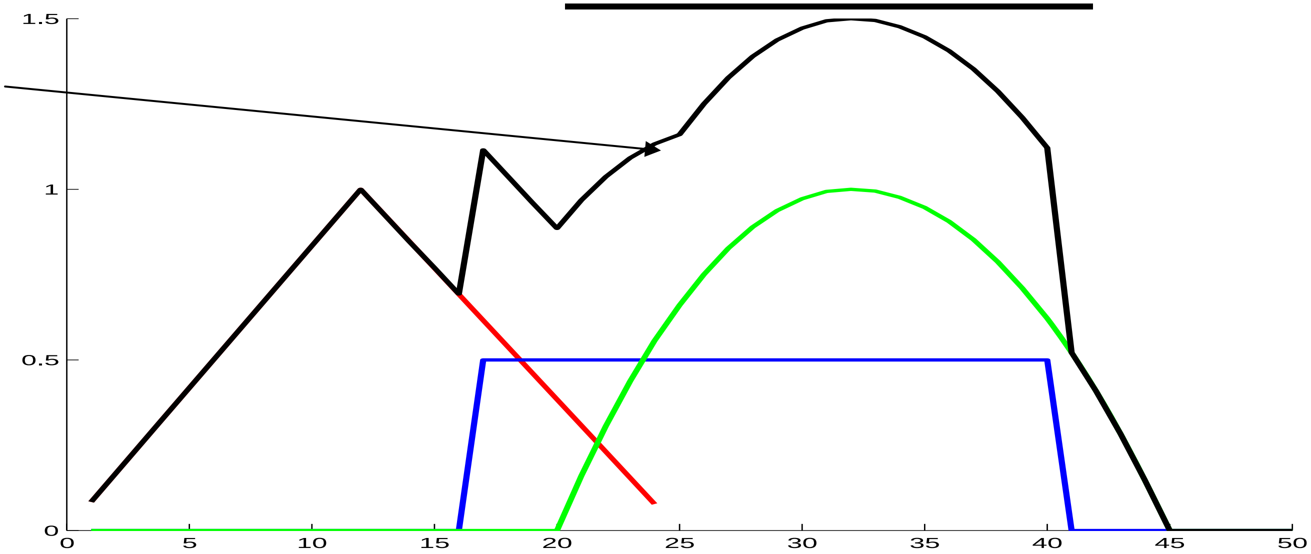
Process instances:

ProcessID=1, S=1

ProcessID=2, S=17

ProcessID=3, S=21

Observed data



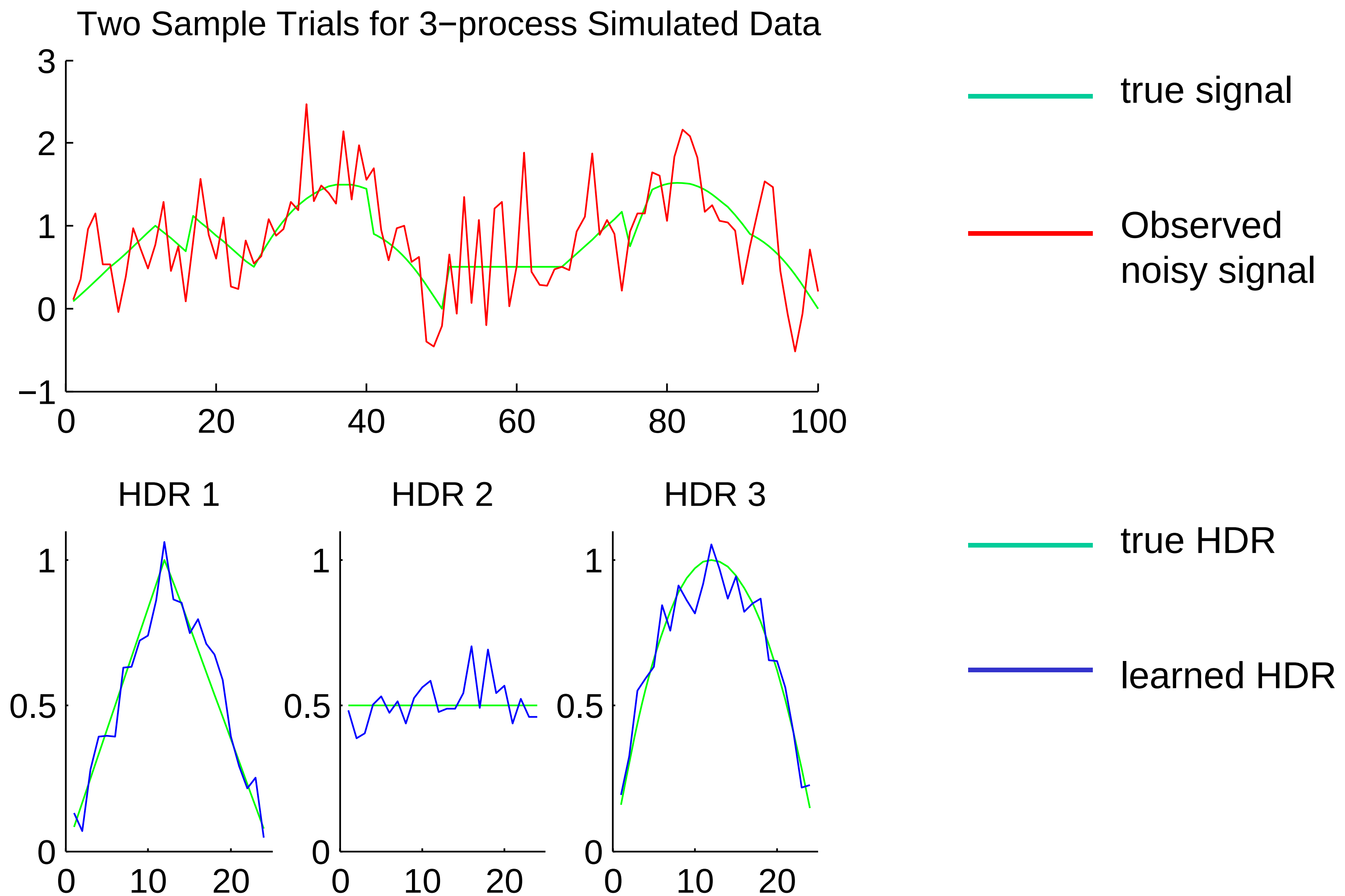
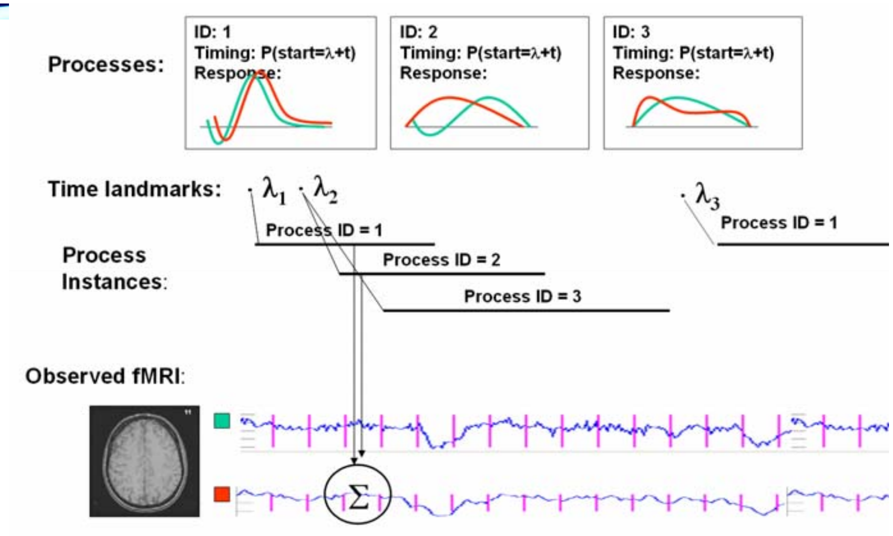


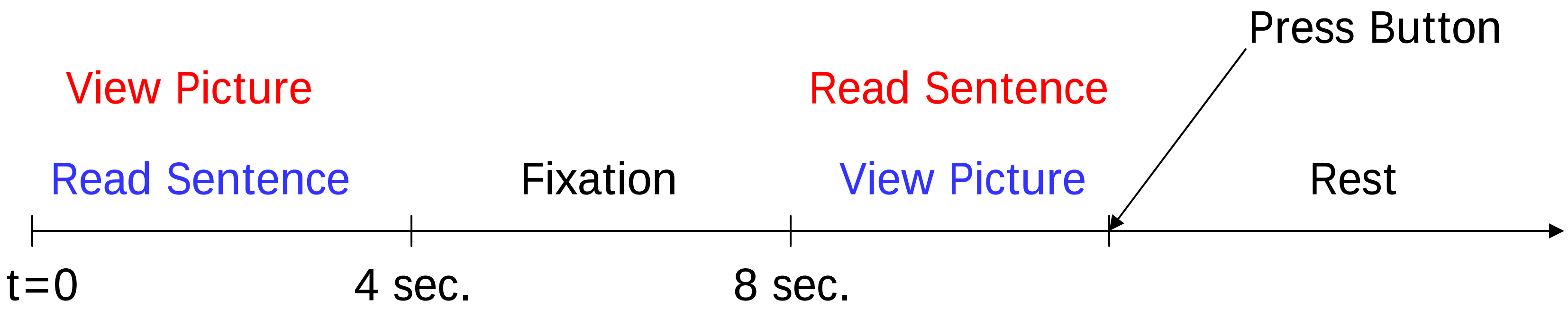
Figure 1. The learner was given 80 training examples with known start times for only the first two processes. It chooses the correct start time (26) for the third process, in addition to learning the HDRs for all three processes.

Using HPMs



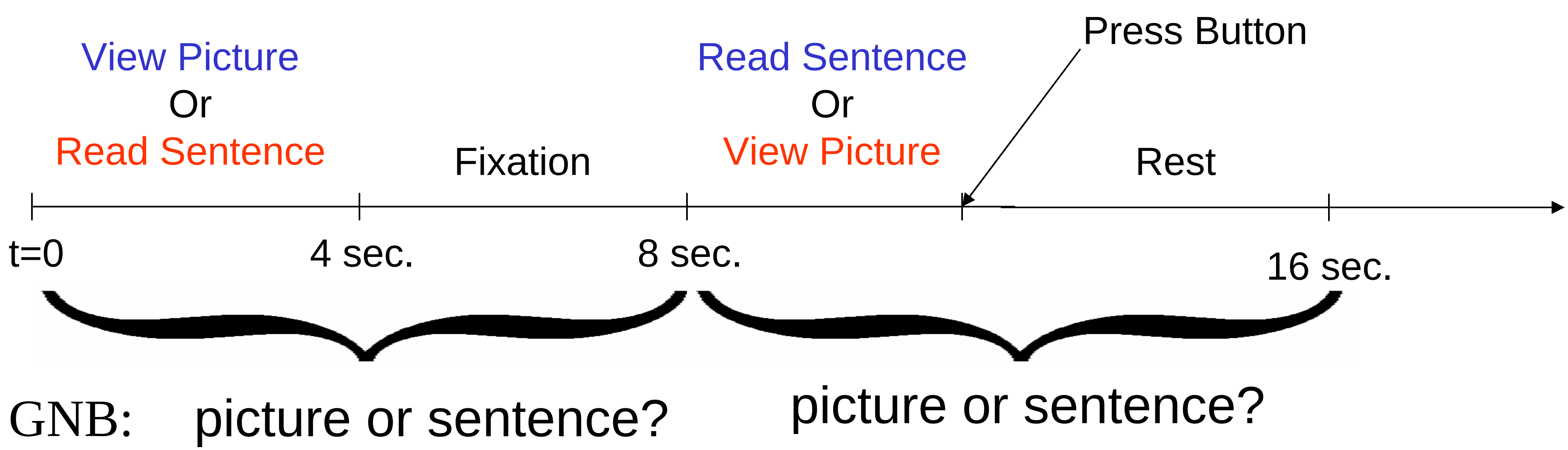
- Given an HPM and data set
  - Assign the ProcessIDs and StartTimes that maximize
    P(ObservedData | ProcessIDs,StartTimes)
  - Subject to any known processIDs, and prior probabilities on their StartTimes
- Classification = assigning processIDs

Study: Pictures and Sentences



- 13 normal subjects.
- 40 trials per subject.
- Sentences and pictures describe 3 symbols: *, +, and $, using ‘above’, ‘below’, ‘not above’, ‘not below’.
- Images are acquired every 0.5 seconds.

Standard classifier formulation

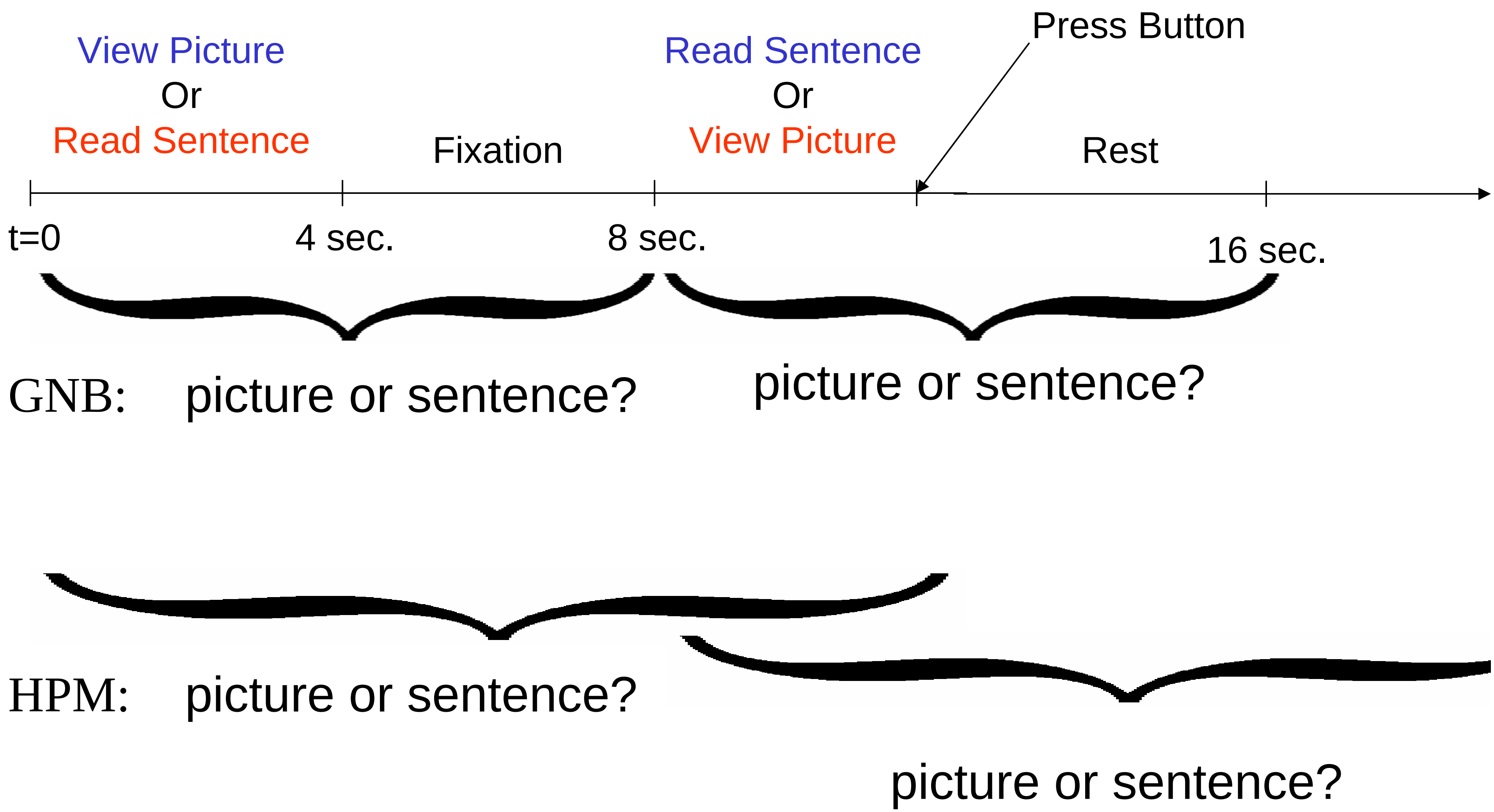


Standard formulation of classification problem.

Train on labeled data, assuming known IDs, StartTimes

Fails to account for overlapping influences of processes

HPM classifier accounts for overlap



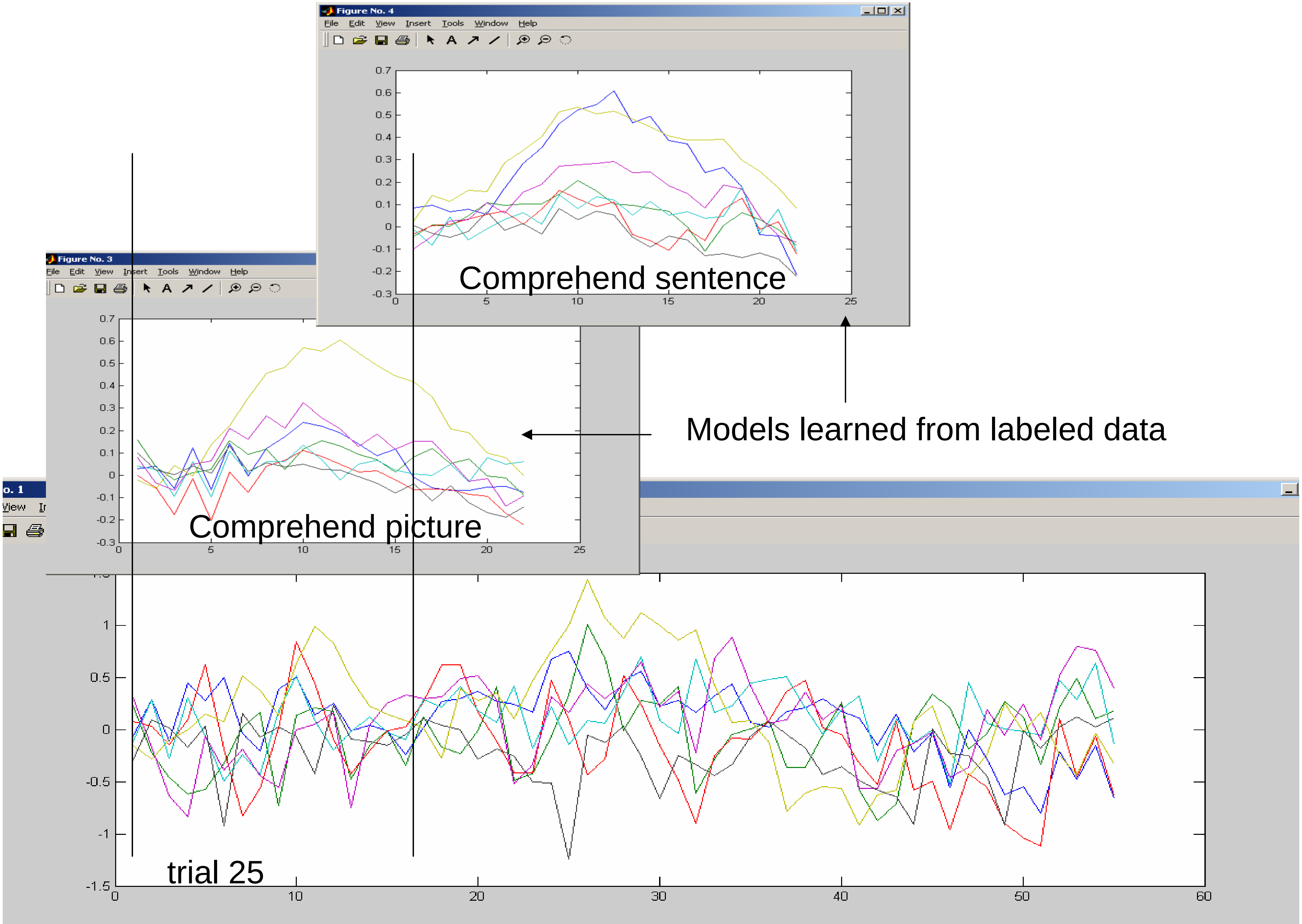


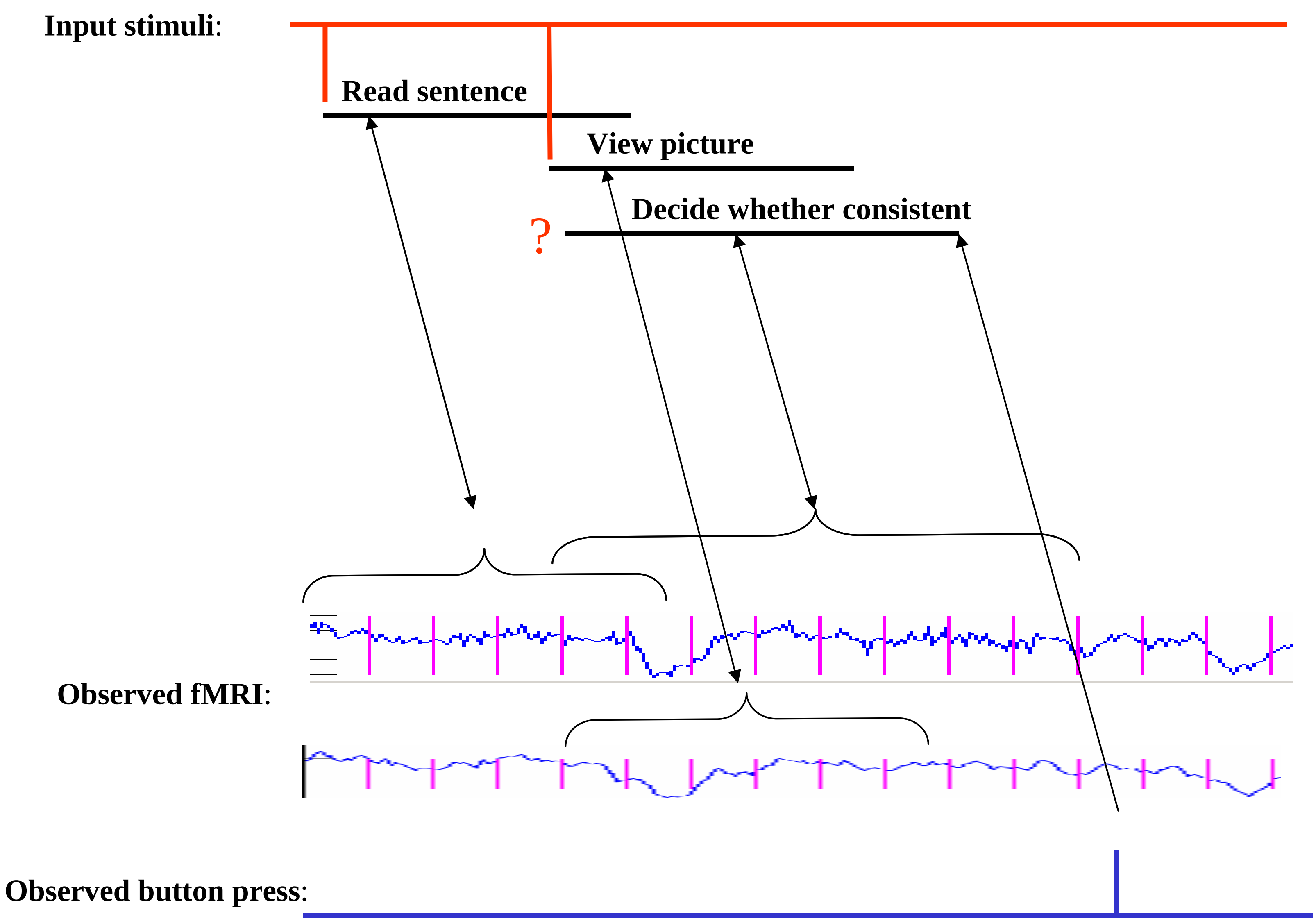
Figure No. 4
Comprehend sentence
Figure No. 3
Comprehend picture
Models learned from labeled data
trial 25

GNB vs. HPM Classification

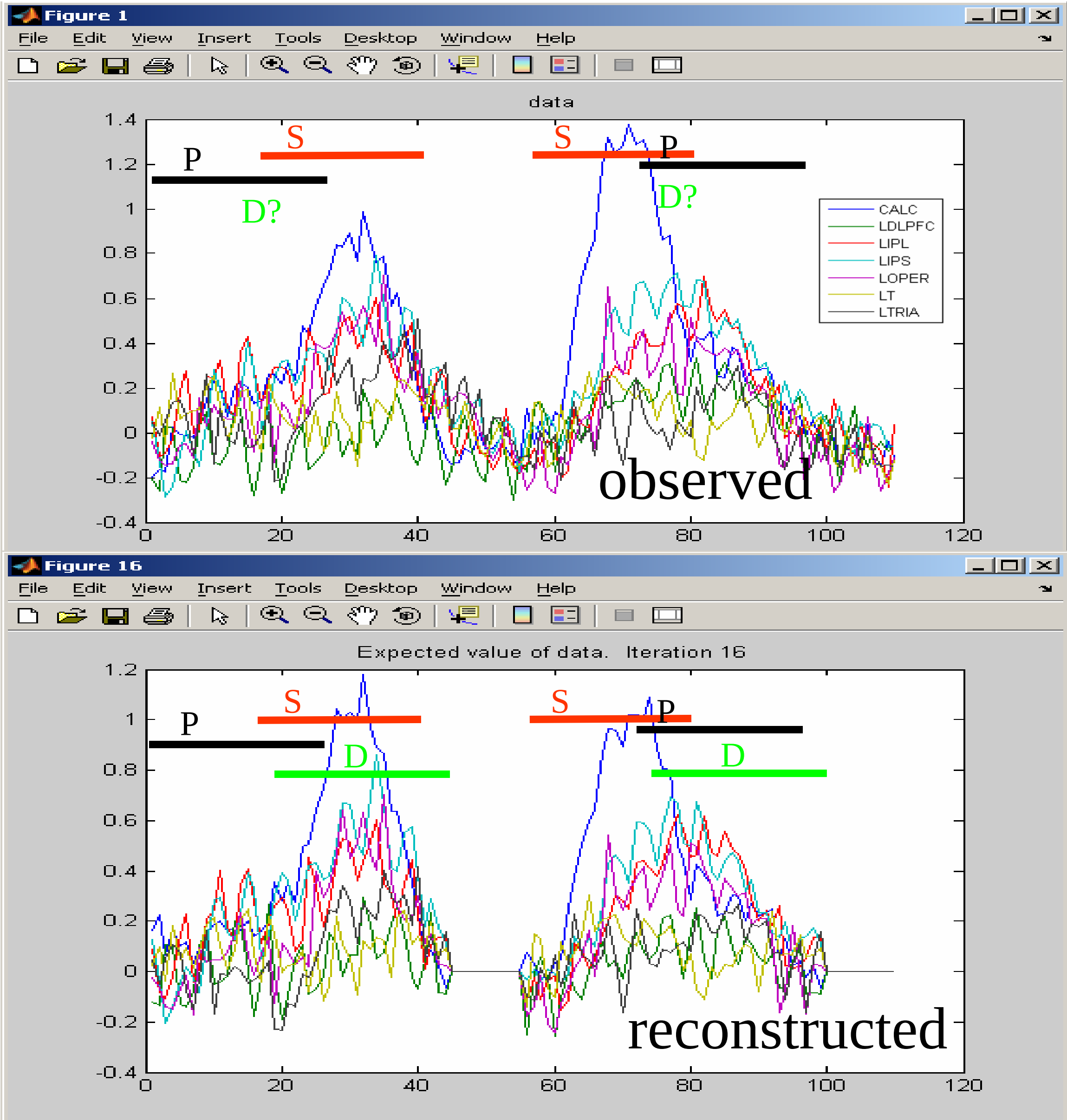
- GNB: assumes non-overlapping processes
- HPM: simultaneous classification of multiple overlapping processes

- Average improvement of 15% in classification error using HPM vs GNB
- E.g., subject 04847
  - GNB classification error: 0.14
  - HPM classification error: 0.09

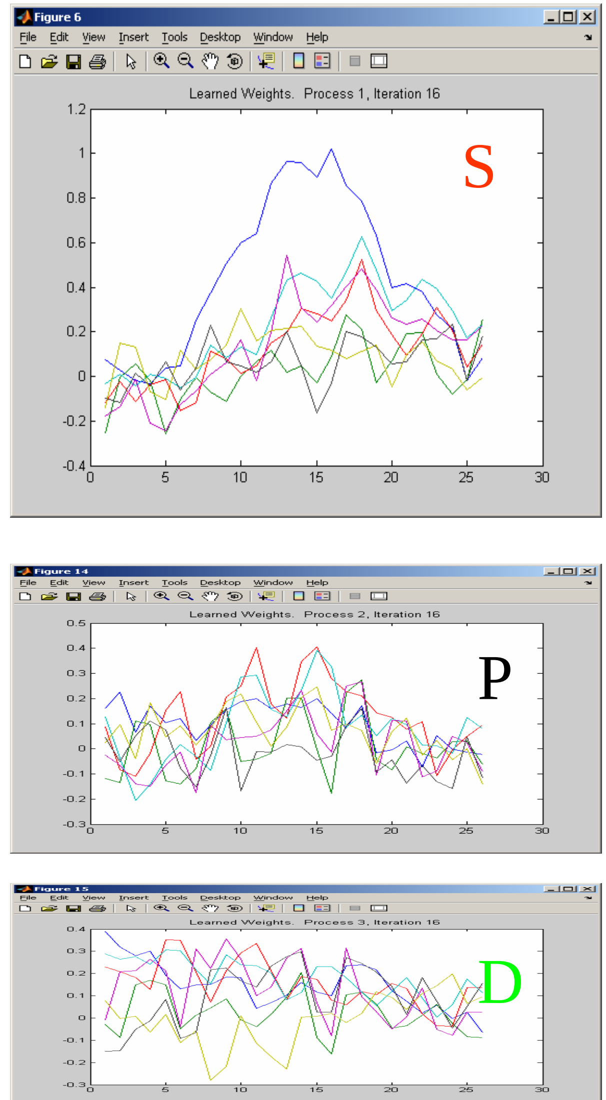
Learning hidden processes with *unknown* start times



Learned HPM with 3 processes (S,P,D), and R=13sec (TR=500msec).



Learned models



D start time chosen by program as t+18

HPM's on Picture-Sentence task

- HPM classification accuracy for Picture/Sentence better than Gaussian Naïve Bayes (GNB)
- HPMs are a strict generalization of GNB
- Model with 2 or 3 cognitive processes?
  - How would we know ground truth?
  - Cross validated data likelihood P(testData | HPM)
    - Better with 3 processes than 2
  - Cross validated classification accuracy
    - Better with 3 processes than 2

Summary

- Classification of cognitive processes from fMRI brain image data
  - Works!
  - Feature selection with “zero signal” data
- Learning models of overlapping, hidden cognitive processes
  - Hidden Process Model formalism
  - Superiority over standard classification
  - Base for studying hidden human processes