Understanding Human Behavior from Sensor Data

Henry Kautz

Dieter Fox

Don Patterson

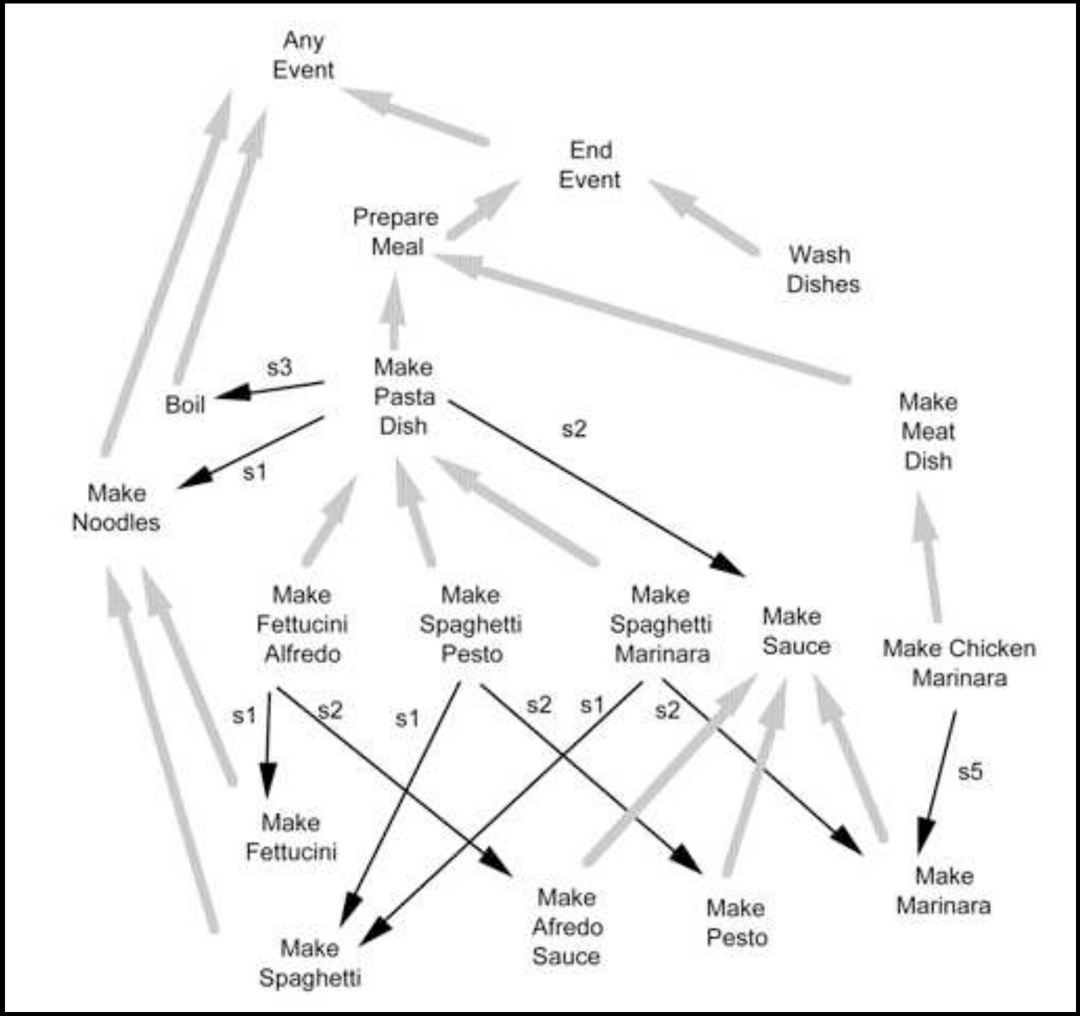
Liao Lin

University of Washington

A Dream of AI

- Systems that can understand ordinary human experience
- Work in KR, NLP, vision, IUI, planning…
  - o Plan recognition
  - o Behavior recognition
  - o Activity tracking
- Goals
  - o Intelligent user interfaces
  - o Step toward true AI

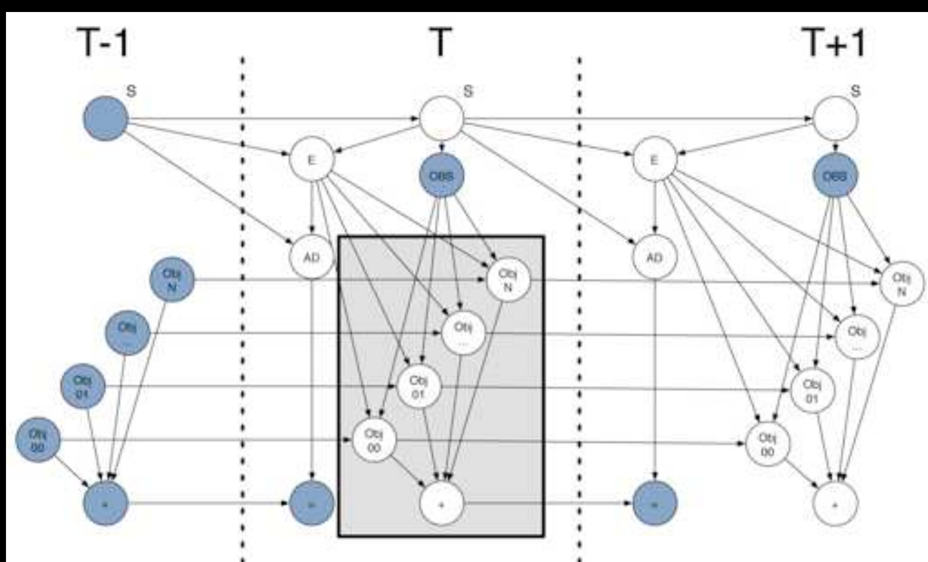
Plan Recognition, circa 1985



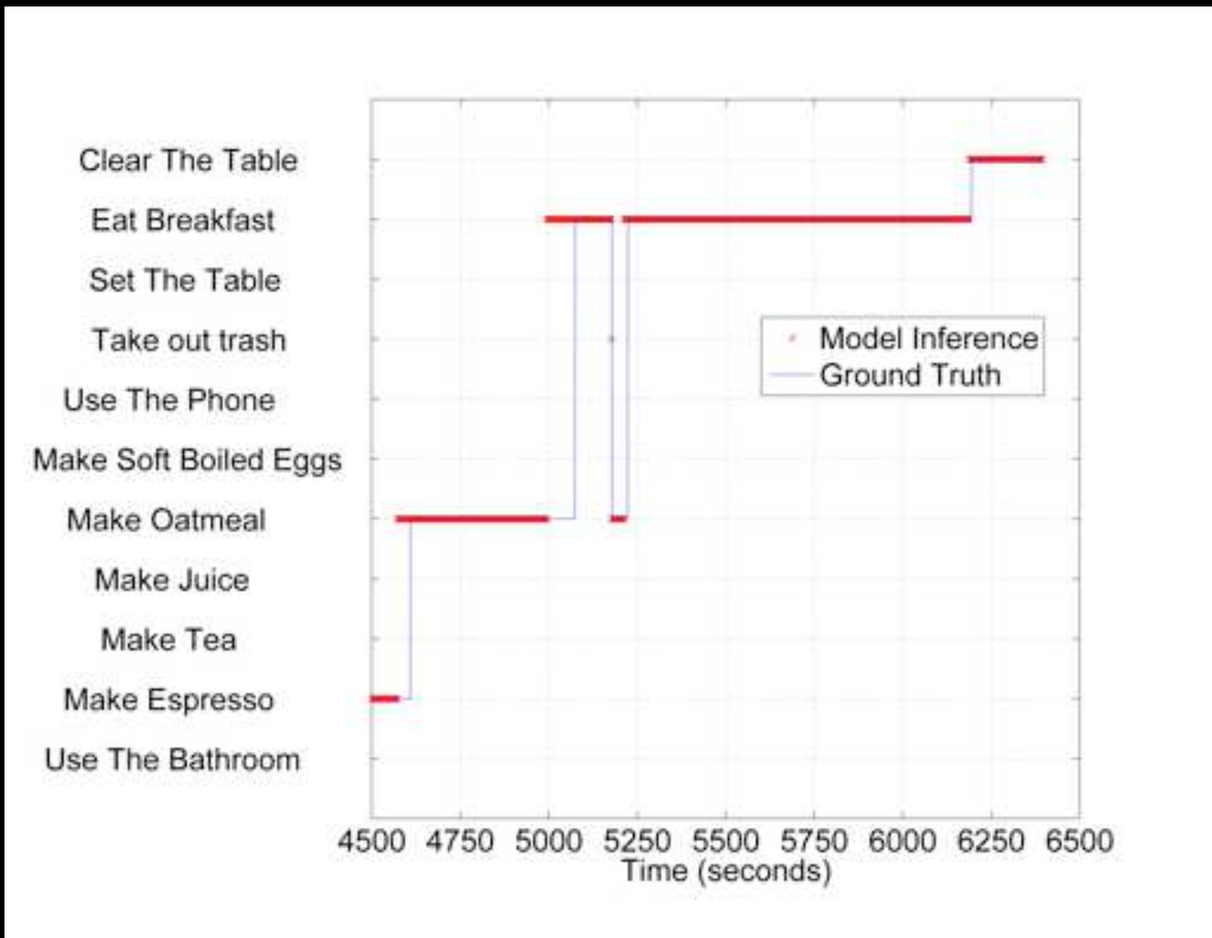
Logical abduction

+ Hierarchy
- Probabilities
- User errors
- Grounding

Activity Tracking, circa 2005



Activity Tracking, circa 2005



Probabilistic inference from sensor data

+ Probabilities

+ User errors

+ Grounding

+/- Hierarchy

Punch Line

- Resurgence of work in behavior understanding, fueled by
  - o Advances in probabilistic inference
    - Graphical models
    - Scalable inference
    - KR U Bayes
  - o Ubiquitous sensing devices
    - RFID, GPS, motes, …
    - Ground recognition in sensor data

Research Issues

- Domain modeling
  - o What is to be modeled about the domain?
  - o What is to be modeled about the user?
  - o What is the data?
  - o What prior knowledge?
  - o What features are useful for learning?
- New representations & algorithms
  - o 101 relational statistical models
- No representation without computation

This Talk

- Activity tracking from RFID tag data
  - ADL Monitoring
- Learning patterns of transportation use from GPS data
  - Activity Compass
- Learning to label activities and places
  - Life Capture

This Talk

- Activity tracking from RFID tag data
  - ADL Monitoring
- Learning patterns of transportation use from GPS data
  - Activity Compass
- Learning to label activities and places
  - Life Capture

Object-Based Activity Recognition

- Activities of daily living involve the manipulation of many physical objects
  - o Kitchen: stove, pans, dishes, …
  - o Bathroom: toothbrush, shampoo, towel, …
  - o Bedroom: linen, dresser, clock, clothing, …
- We can recognize activities from a time-sequence of object touches

Application

- ADL (Activity of Daily Living) monitoring for the disabled and/or elderly
  - Changes in routine often precursor to illness, accidents
  - Human monitoring intrusive & inaccurate

Activities of Daily Living/ Instrumental Activities of Daily Living	Estimated Times
Laundry/Housekeeping	4 hours total
☐ Laundry	2 hours per week
☐ Minor Sewing and Mending	10-15 minutes
☐ Other	
☐ Clean Toilet, Sink, Tub/Shower	30 minutes
☐ Clean Floors, Carpets, Rugs	30 minutes
☐ Clean Kitchen Appliances, Countertops	15 minutes
☐ Dust	10-15 minutes
☐ Make Bed &/Or Change Linens	5-10 minutes
☐ Wash Dishes	10-15 minutes
☐ Wash Cupboards, Walls, Throw Rugs, Curtains, Inside Windows	Once yearly
☐ Remove Trash	5 minutes
☐ Other	

Sensing Object Manipulation

- RFID: Radio-frequency identification tags
  - Small
  - No batteries
  - Durable
  - Cheap
- Easy to tag objects
  - Near future: use products' own tags

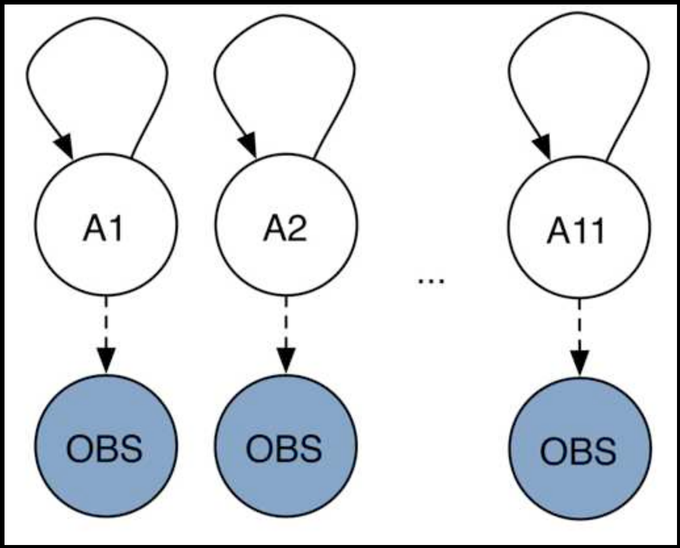
Wearable RFID Readers

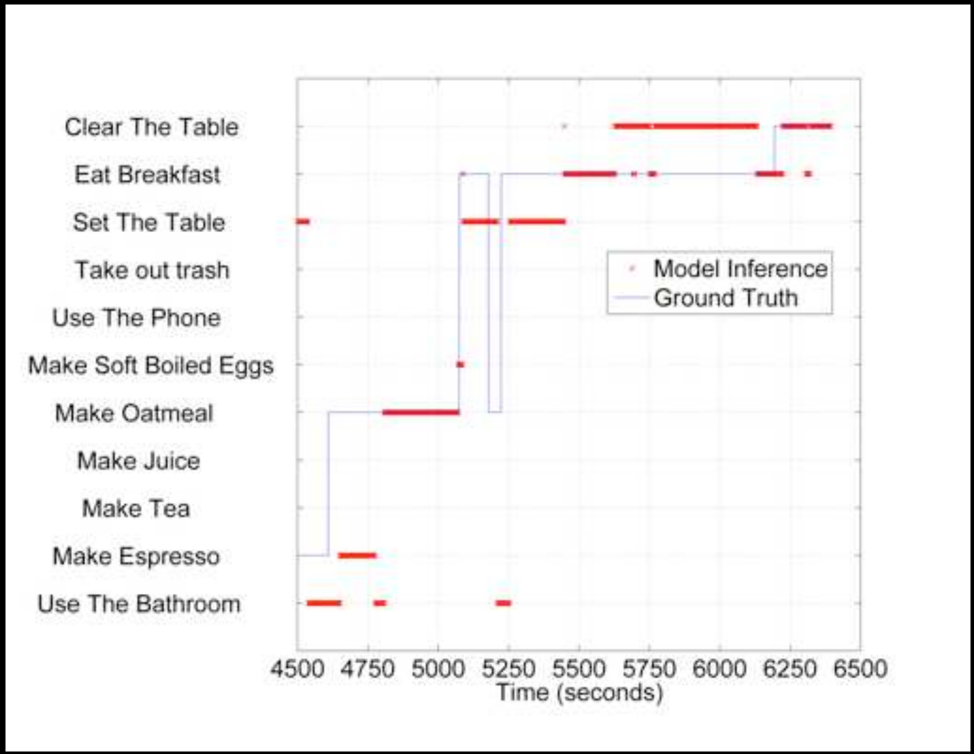
Experiment: Morning Activities

- 10 days of data from the morning routine in an experimenter's home
  - 61 tagged objects
- 11 activities
  - Often interleaved and interrupted
  - Many shared objects

Use bathroom	Make coffee	Set table
Make oatmeal	Make tea	Eat breakfast
Make eggs	Use telephone	Clear table
Prepare OJ	Take out trash	

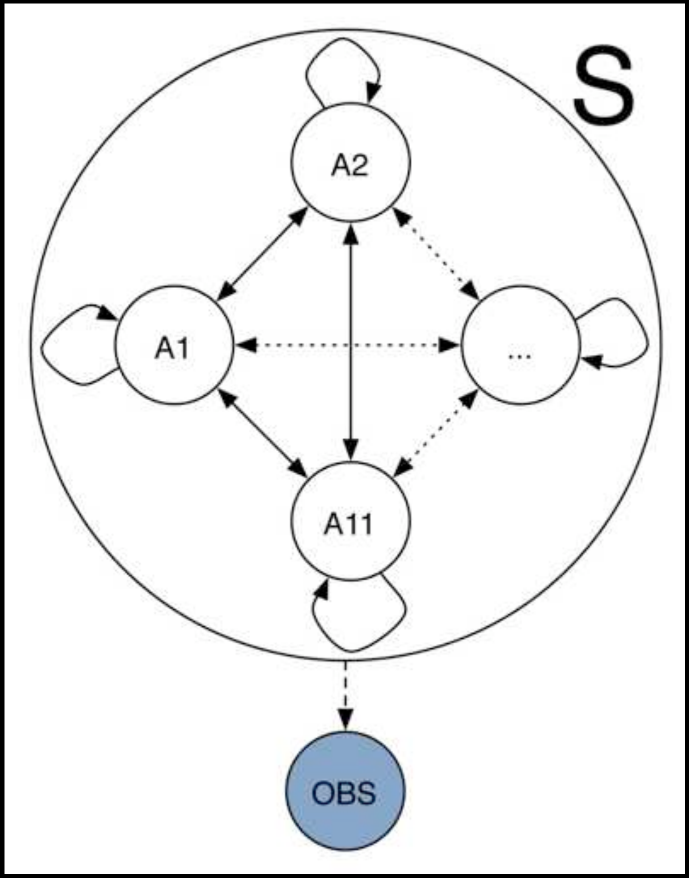
Baseline: Individual Hidden Markov Models

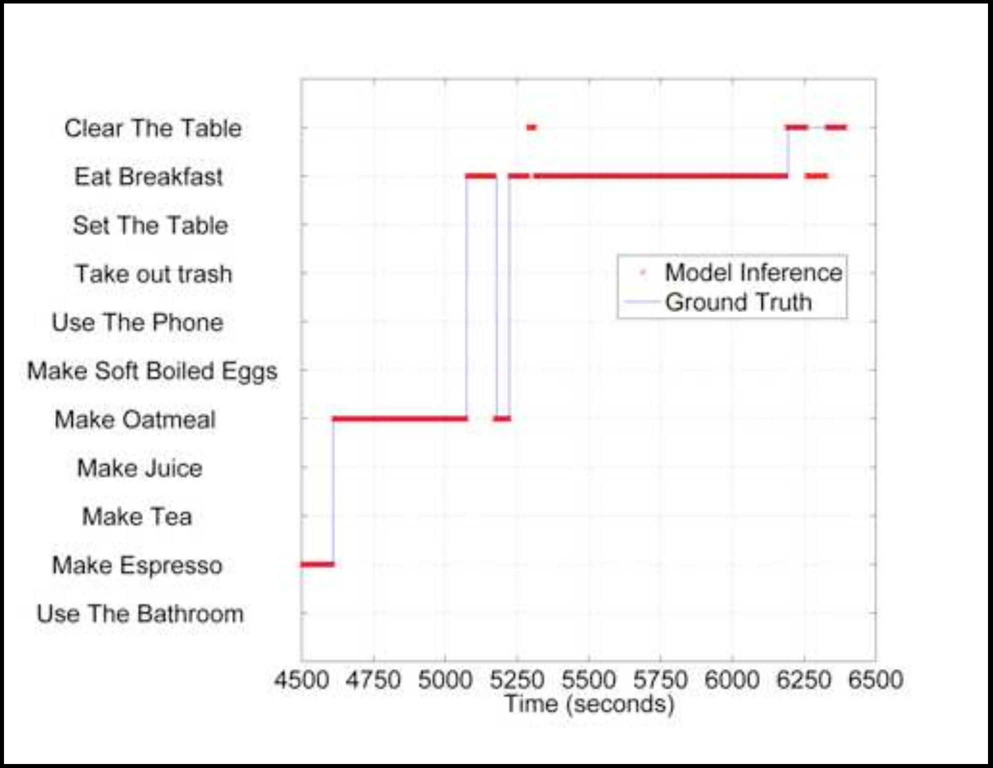




68% accuracy
11.8 errors per episode

Baseline: Single Hidden Markov Model





83% accuracy
9.4 errors per episode

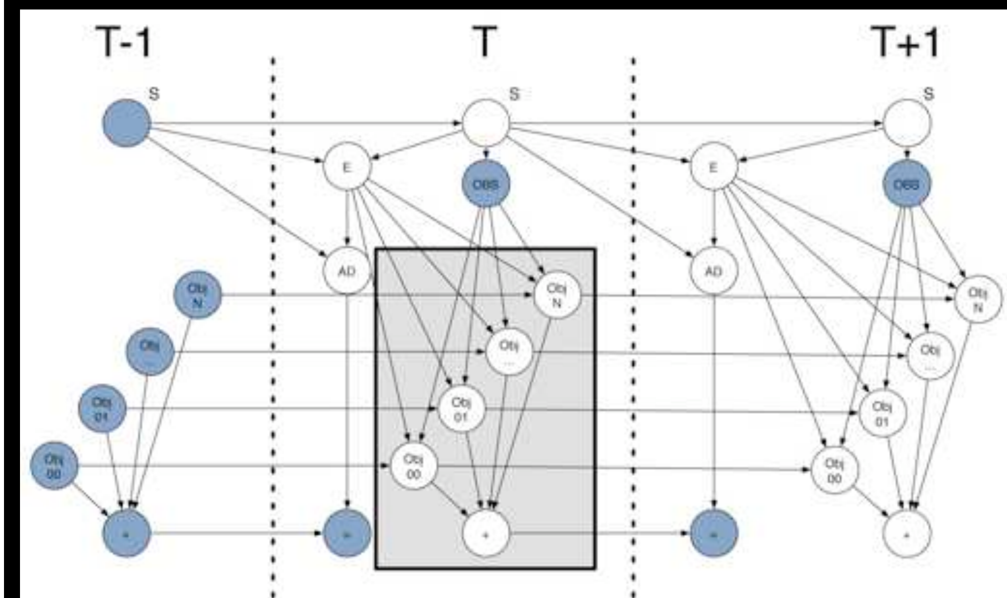
Cause of Errors

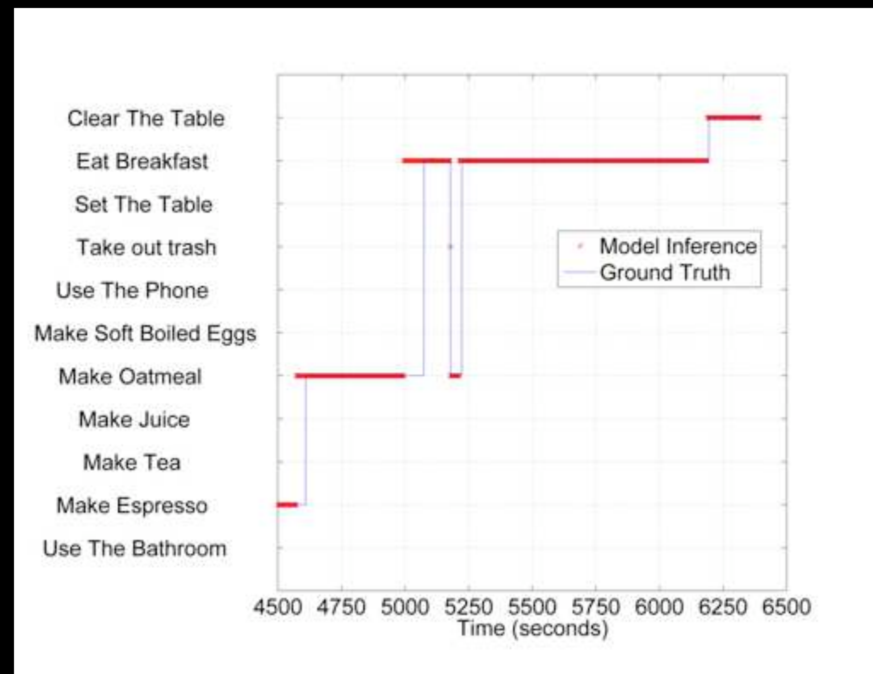
- Observations were types of objects
  - o Spoon, plate, fork …
- Typical errors: confusion between activities
  - o Using one object repeatedly
  - o Using different objects of same type
- Critical distinction in many ADL's
  - o Eating versus setting table
  - o Dressing versus putting away laundry

Aggregate Features

- HMM with individual object observations fails
  - No generalization!
- Solution: add aggregation features
  - Number of objects of each type used
  - Requires history of current activity performance
  - DBN encoding avoids explosion of HMM

Dynamic Bayes Net with Aggregate Features



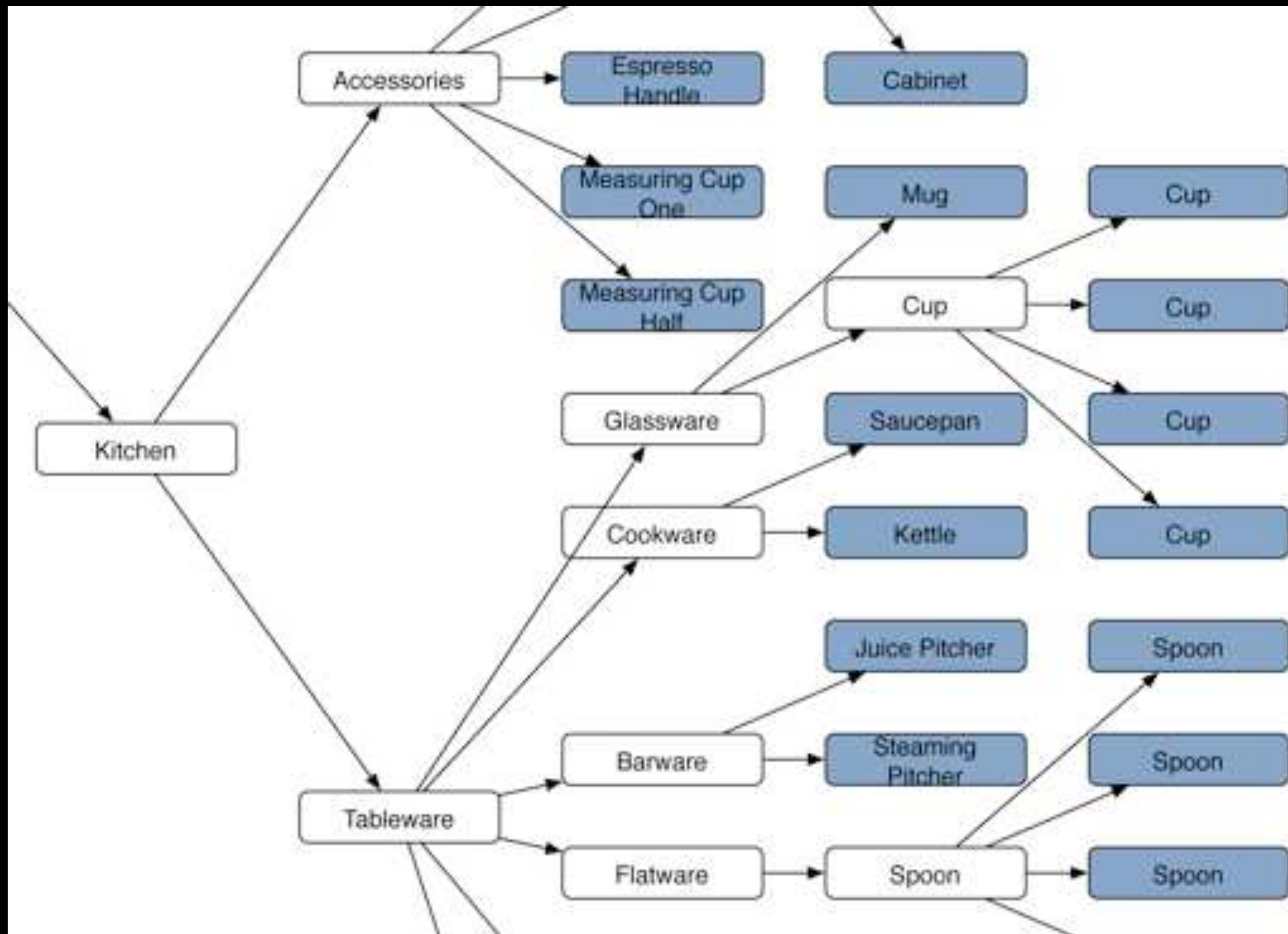


88% accuracy
6.5 errors per episode

Improving Robustness

- DBN fails if novel objects are used

- Solution: smooth parameters over abstraction hierarchy of object types



$$P(O_i \rightarrow O_j) = \frac{\exp(-\frac{Dist(O_i,O_j)}{2})}{\sum_j \exp(-\frac{Dist(O_i,O_j)}{2})}$$

Abstraction Smoothing

- Methodology:
  - Train on 10 days data
  - Test where one activity substitutes one object
- Change in error rate:
  - Without smoothing: 26% increase
  - With smoothing: 1% increase

Summary

- Activities of daily living can be robustly tracked using RFID data
  - o Simple, direct sensors can often replace (or augment) general machine vision
  - o Accurate probabilistic inference requires sequencing, aggregation, and abstraction
  - o Works for essentially all ADLs defined in healthcare literature

Inferring ADLs from Interactions with Objects, *Pervasive Computing*, 3 (4), 2004

This Talk

- Activity tracking from RFID tag data
  - o ADL Monitoring
- Learning patterns of transportation use from GPS data
  - o Activity Compass
- Learning to label activities and places
  - o Life Capture

Challenge

- Given a data stream from a GPS unit...
  - Infer the user’s mode of transportation, and places where the mode changes
    - Foot, car, bus, bike, …
    - Bus stops, parking lots, enter buildings, …
  - Learn the user’s daily pattern of movement
  - Predict the user’s future actions
  - Detect user errors

Application

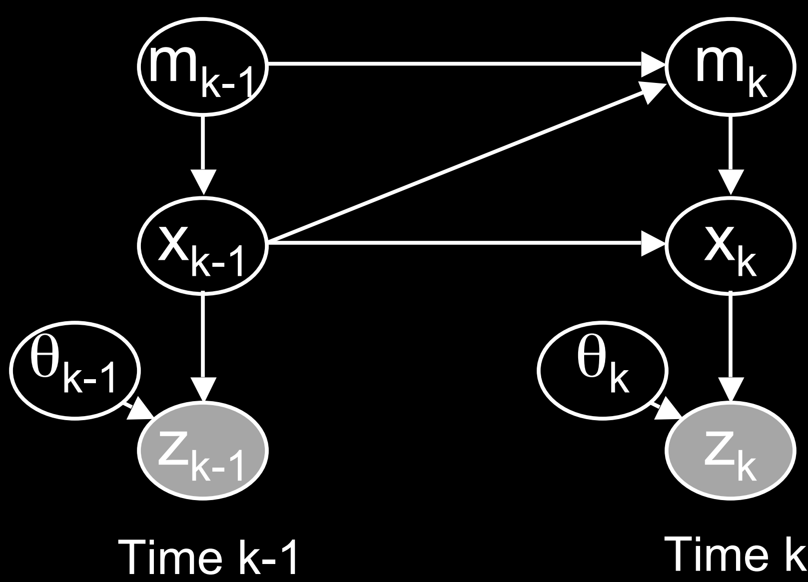
- Activity Compass
  - Personal guidance system for people with cognitive disabilities
  - Stroke, traumatic brain injury, mental retardation
  - Adaptive and proactive

Patterson *et al,* Opportunity Knocks: a System to Provide Cognitive Assistance with Transportation Services, *UBICOMP-2004*

Approach

- Map is a directed graph G=(V,E)
- Location:
  - Edge e
  - Distance d from start of edge
  - Actual (displaced) GPS reading
- Movement:
  - Mode { foot, car, bus } determines velocity range
  - Change mode near bus stops & parking places
- Tracking (filtering): Given some prior estimate,
  - Update position & mode according to motion model
  - Correct according to next GPS reading

Dynamic Bayesian Network I



Transportation mode

Edge, velocity, position

Data (edge) association

GPS reading

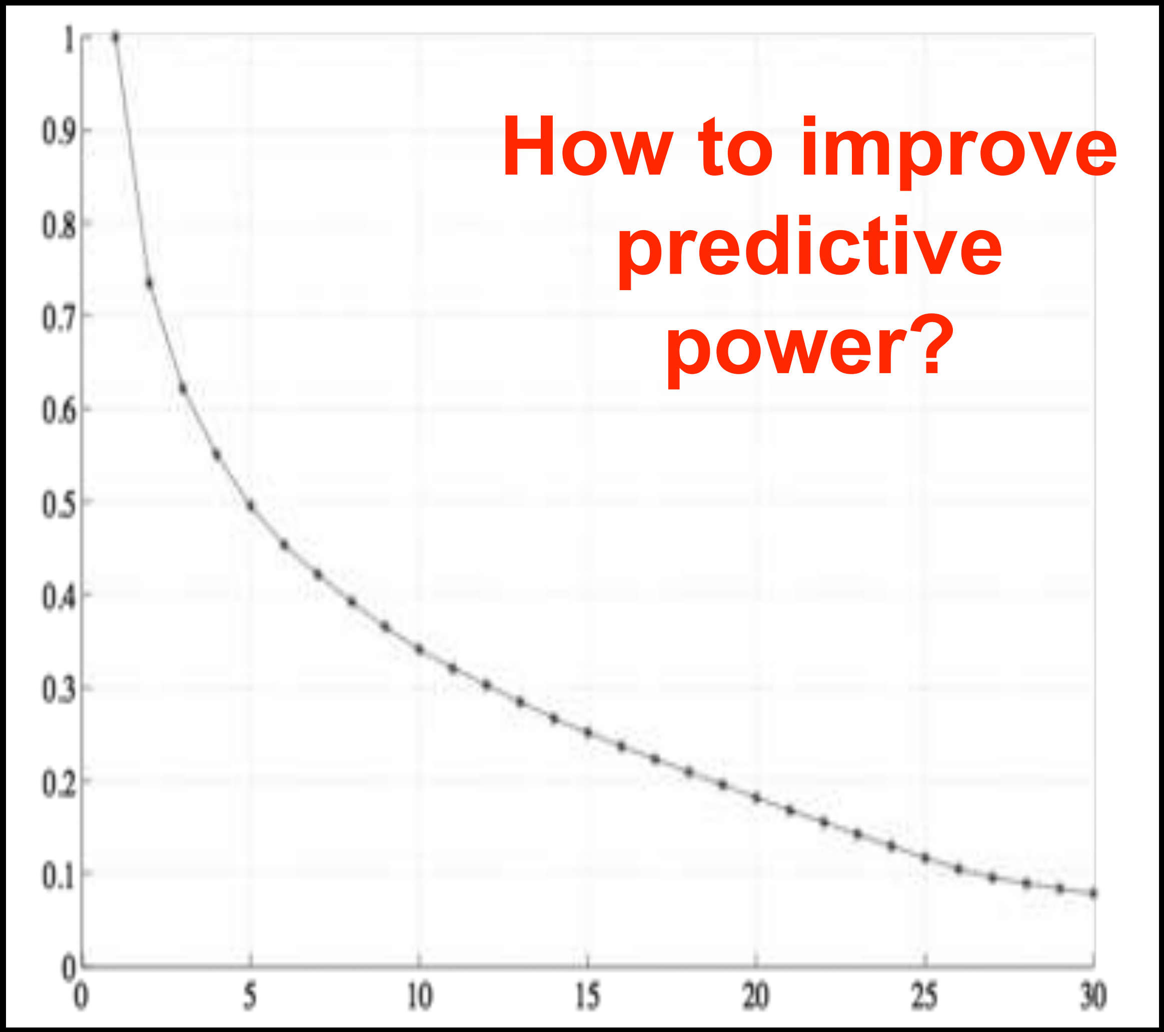
Mode & Location Tracking

■	Measurements
●	Projections
Green	Bus mode
Red	Car mode
Blue	Foot mode

Learning

- Prior knowledge – general constraints on transportation use
  - Vehicle speed range
  - Bus stops
- Learning – specialize model to particular user
  - 30 days GPS readings
  - Unlabeled
  - Learn edge transition parameters using expectation-maximization (EM)

Predictive Accuracy



Probability of correctly predicting the future

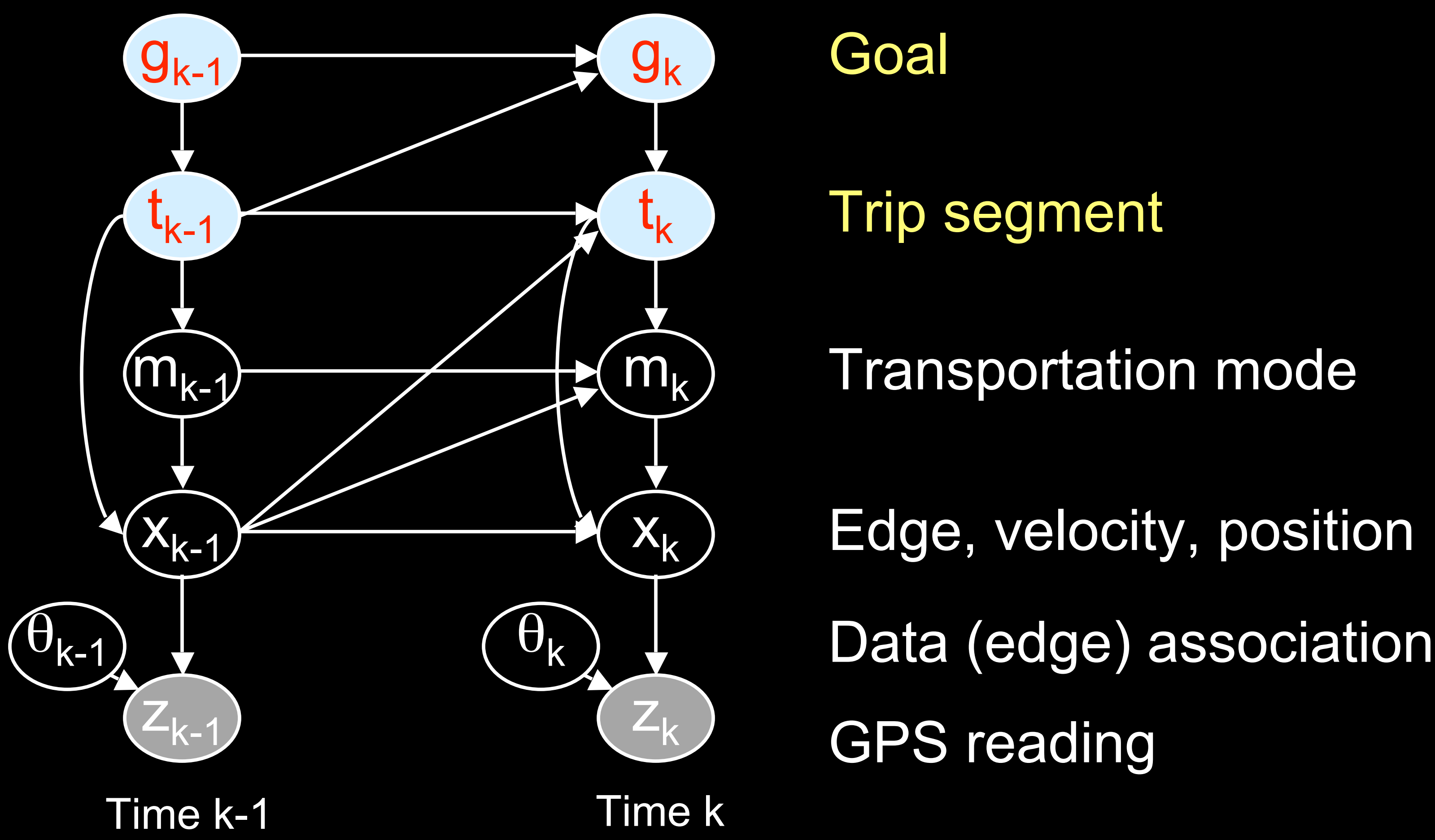
City Blocks

Transportation Routines

Home **A** **B** **Workplac e**

- Goal: intended destination
  - Workplace, home, friends, restaurants, …
- Trip segments: *<start, end, mode>*
  - **Home** to **Bus stop A** on **Foot**
  - **Bus stop A** to **Bus stop B** on **Bus**
  - **Bus stop B** to **workplace** on **Foot**

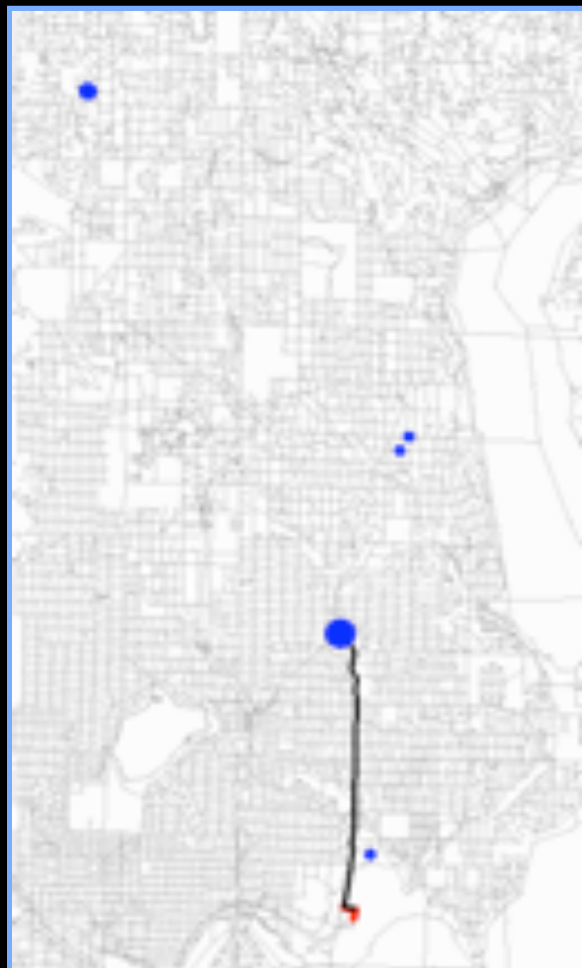
Dynamic Bayesian Net II



Unsupervised Hierarchical Learning

- Use previous model to infer:
  - o Goals
    - locations where user stays for a long time
  - o Transition points
    - locations with high mode transition probability
  - o Trip segments
    - paths connecting transition points or goals
- Learn transition probabilities
  - o Expectation-Maximization

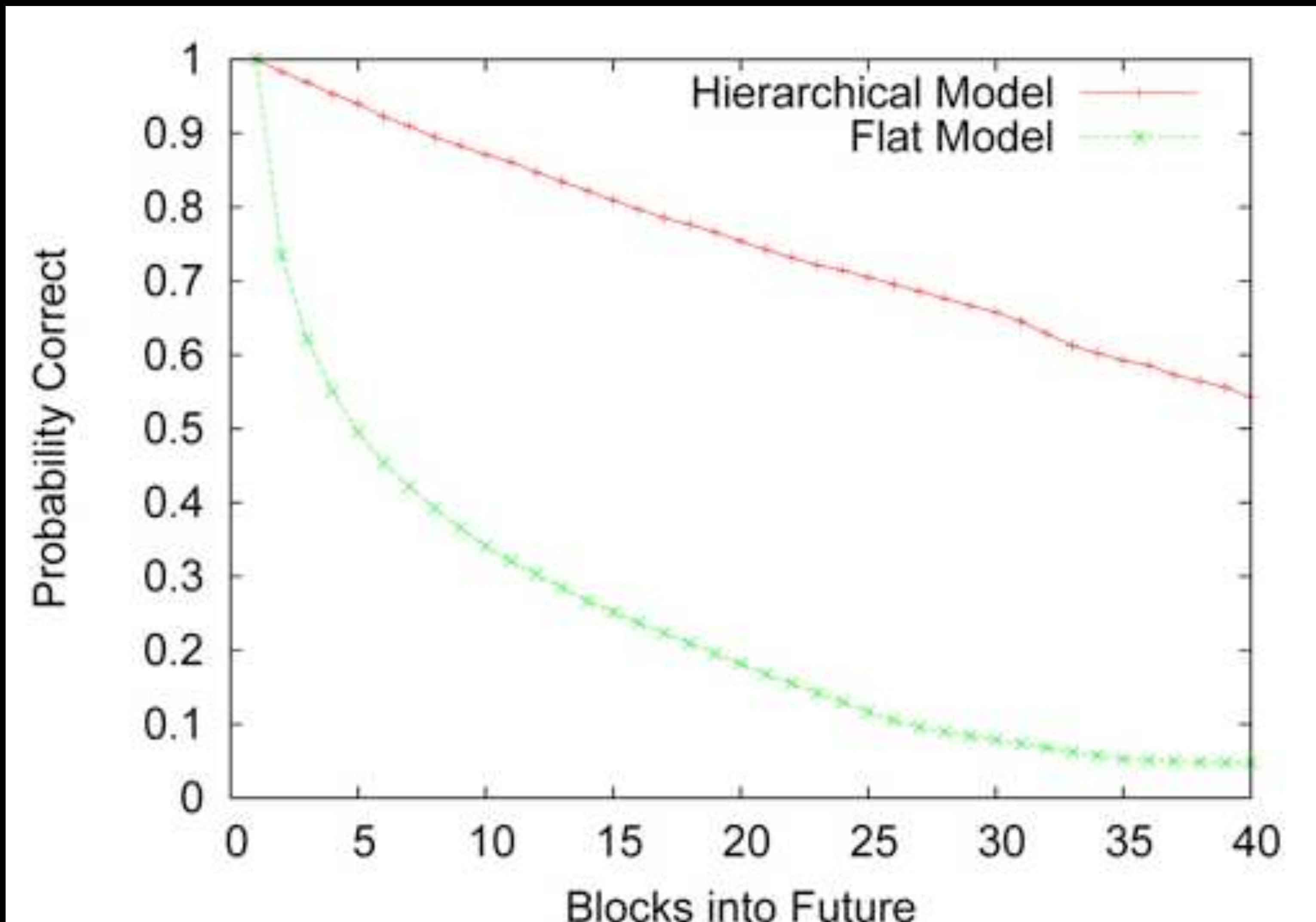
Predict Goal and Path



Predicted goal

Predicted path

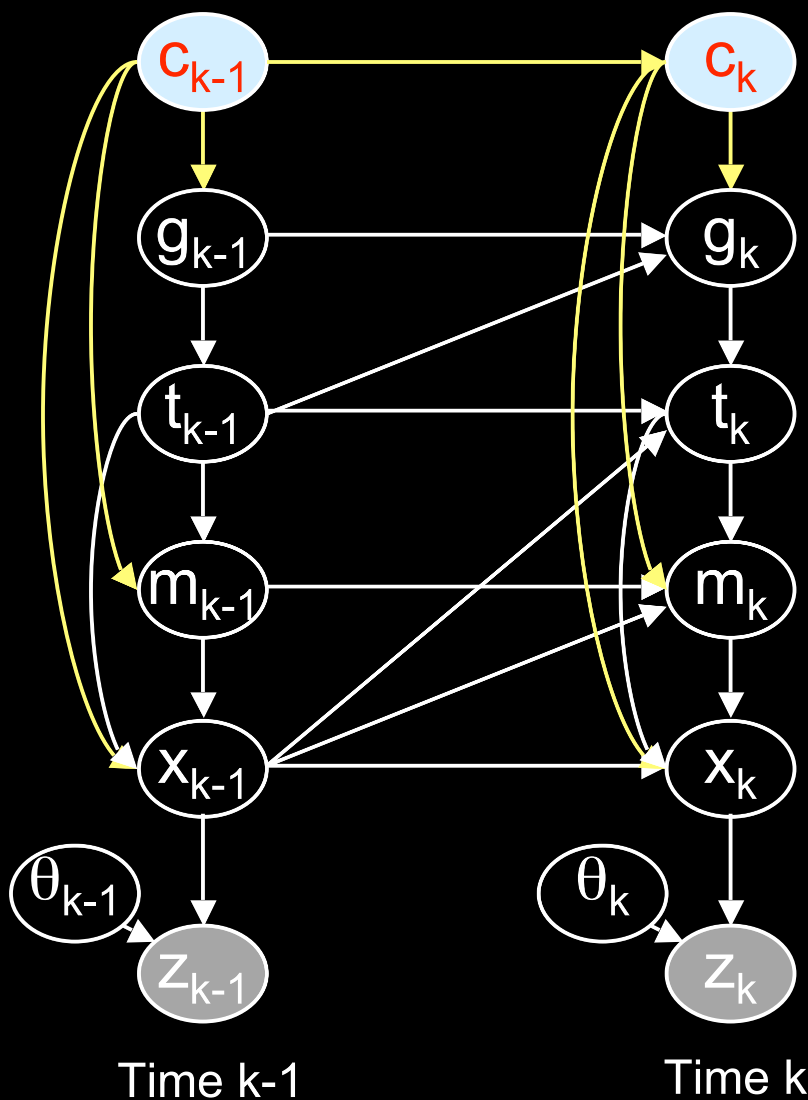
Improvement in Predictive Accuracy



Detecting User Errors

- Learned model represents typical correct behavior
  - o Model is a poor fit to user errors
- We can use this fact to detect errors!
- Cognitive Mode
  - o Normal: model functions as before
  - o Error: switch in prior (untrained) parameters for mode and edge transition

Dynamic Bayesian Net III



Cognitive mode
{ normal, error }

Goal

Trip segment

Transportation mode

Edge, velocity, position

Data (edge) association

GPS reading

Detecting User Errors

Status

- Development funded by National Institute for Disability and Rehabilitation Research
  - o UBICOMP 2003, 2005
  - o AAAI 2004 Best Paper Award
- Current work
  - o Usability studies
  - o Audio interface
  - o Learning effective prompting strategies

This Talk

- Activity tracking from RFID tag data
  - ADL Monitoring
- Learning patterns of transportation use from GPS data
  - Activity Compass
- Learning to label activities and places
  - Life Capture

Task

- Learn to label a person's
  - Daily activities
    - working, visiting friends, traveling, …
  - Significant places
    - work place, friend's house, usual bus stop, …
- Given
  - Training set of labeled examples
  - Wearable sensor data stream
    - GPS, acceleration, ambient noise level, …

Application

- Life Capture
  - o Automated diary
  - o On-duty log book

Time	Activity and transportation
8:15am - 8:34am	Drive from home 1 to parking lot 2, walk to workplace 1;
8:34am - 5:44pm	Work at workplace 1;
5:44pm - 6:54pm	Walk from workplace 1 to parking lot 2, drive to friend's place 3;
6:54pm - 6:56pm	Pick up/drop off at friend 3's place;
6:56pm - 7:15pm	Drive from friend 3's place to other place 5;
9:01pm - 9:20pm	Drive from other place 5 to friend 3's place;
9:20pm - 9:21pm	Pick up/drop off at friend 3's place;
9:21pm - 9:50pm	Drive from friend 3's place to home 1;
9:50pm - 8:22am	Sleep at home 1.

Conditional Models

- HMMs and DBNs are generative models
  - o Describe complete joint probability space
- For labeling tasks, conditional models are often simpler and more accurate
  - o Learn only P(label | observations)
  - o Fewer parameters than corresponding generative model

Things to be Modeled

- Raw GPS reading (observed)
- Actual user location
- Activities (time dependent)
- Significant places (time independent)
- Soft constraints between all of the above (learned)

Conditional Random Field

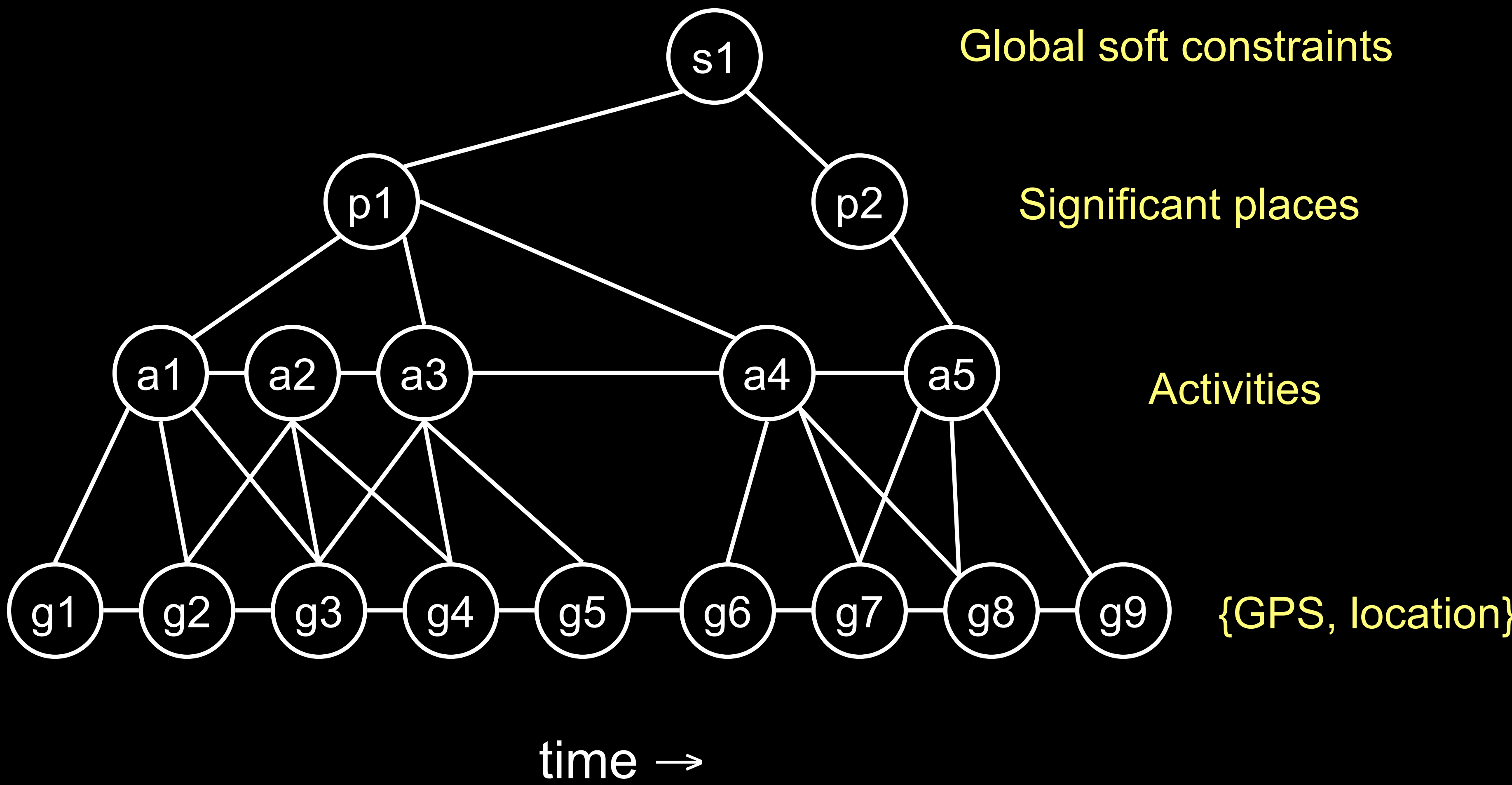
- Undirected graphical model
  - o Feature functions defined on cliques
  - o Conditional probability proportional to exp(weighted sum of features)
  - o Weights learned by maximizing (pseudo) likelihood of training data

$$p(\mathbf{y} \mid \mathbf{x}) \quad = \quad \frac{1}{Z(\mathbf{x})} \prod_{C \in \mathcal{C}} \prod_{\mathbf{v}_C \in C} \exp\{\mathbf{w}_C^T \cdot \mathbf{f}_C(\mathbf{v}_C)\}$$

Relational Markov Network

- First-order version of conditional random field
- Features defined by feature templates
  - All instances of a template have same weight
- Examples:
  - Time of day an activity occurs
  - Place an activity occurs
  - Number of places labeled "Home"
  - Distance between adjacent user locations
  - Distance between GPS reading & nearest street

RMN Model


s1
Global soft constraints
p1
p2
Significant places
a1
a2
a3
a4
a5
Activities
g1
g2
g3
g4
g5
g6
g7
g8
g9
{GPS, location}
time →

Significant Places

- Previous work decoupled identifying significant places from rest of inference
  - o Simple temporal threshold [Ashbrook & Starner 2003; Liao, Kautz, & Fox 2004; Dechter et al. 2005]
  - o Misses places with brief activities
- RMN model integrates
  - o Identifying significant place
  - o Labeling significant places
  - o Labeling activities

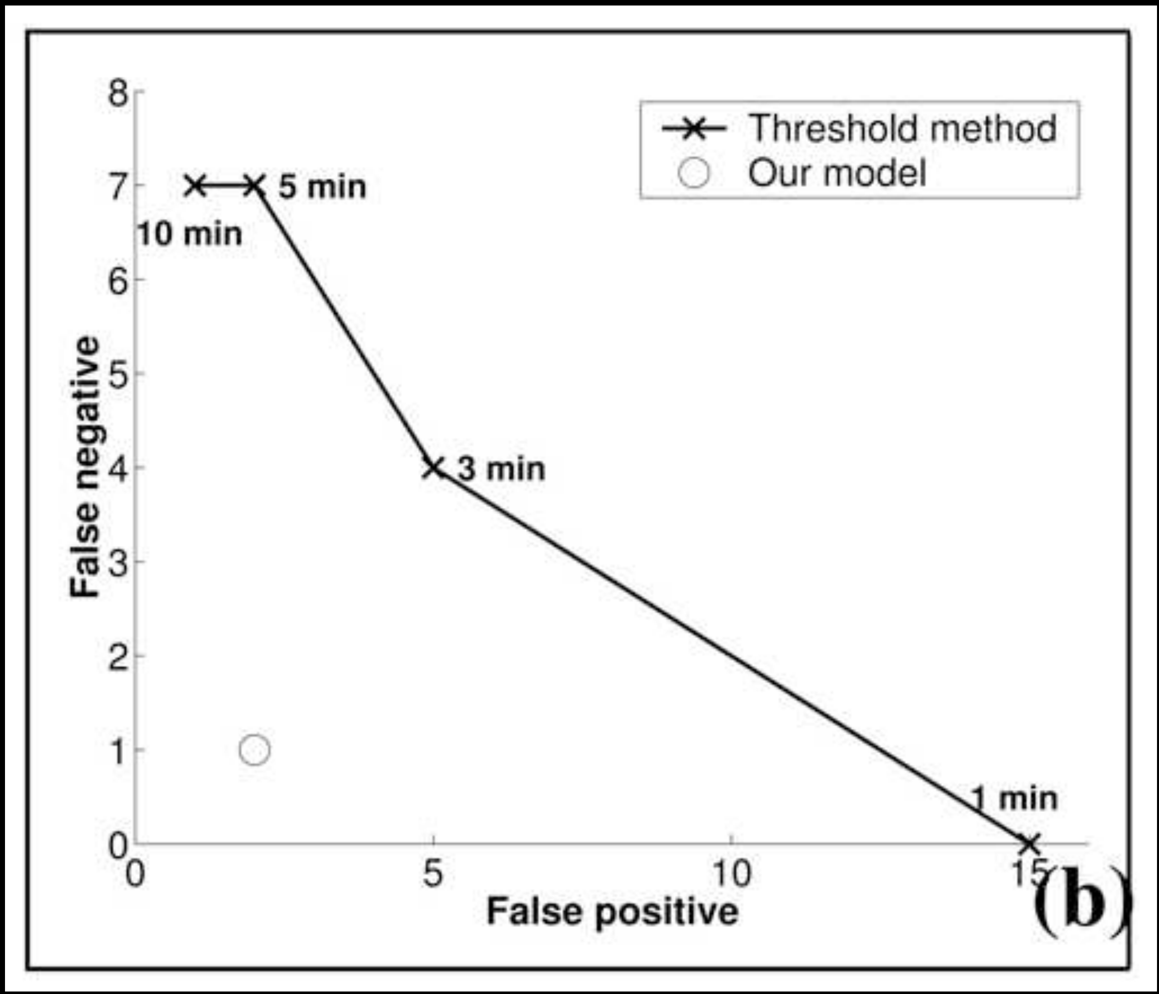
Efficient Inference

- Some features are expensive to handle by general inference algorithms
  - E.g. belief propagation, MCMC
- Can dramatically speed up inference by associating inference procedures with feature templates
  - Fast Fourier transform (FFT) to compute "counting" features
  - $O(n \log^2 n)$ versus $O(2^n)$

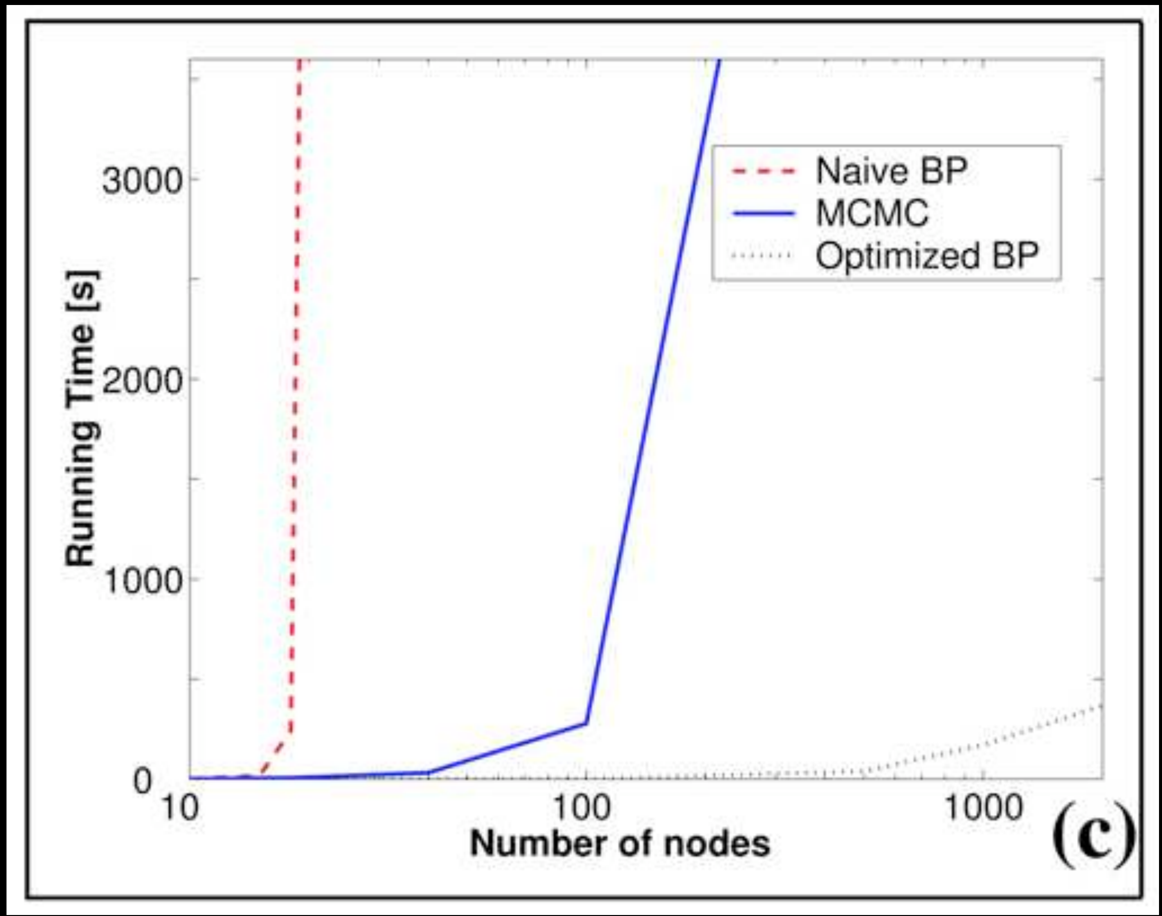
Results: Labeling

- One user, 1 week training data, 1 week testing data
  - Number of (new) significant places correctly labeled: 18 out of 19
  - Number of activities correctly labeled: 53 out of 61
  - Number of activities correctly labeled, if counting features not used: 44 out of 61

Results: Finding Significant Places



Results: Efficiency



Summary

- We can learn to label a user's activities and meaningful locations using sensor data & state of the art relational statistical models
- Many avenues to explore:
  - o Transfer learning
  - o Finer grained activities
  - o Structured activities
  - o Social groups

Conclusion: Why Now?

- An early goal of AI was to create programs that could understand ordinary human experience
- This goal proved elusive
  - o Missing probabilistic tools
  - o Systems not grounded in real world
  - o Lacked compelling purpose
- Today we have the mathematical tools, the sensors, and the motivation

Credits

- Graduate students:
  - o Don Patterson, Lin Liao
- Colleagues:
  - o UW CSE: Dieter Fox, Gaetano Borriello
  - o UW Rehabilitation Medicine
  - o Intel Research Seattle
- Funders:
  - o NIDRR, Intel, NSF, DARPA