Greedy Algorithms for Sequential Sensing Decision

Hannaneh Hajishirzi*
Afsaneh Shirazi
Jaesik Choi
Eyal Amir

University of Illinois at Urbana-Champaign

* speaker
Motivation

- Search engines:
  - Updated repository of web pages
  - Not monitor every page every time

→ Important to detect change of web pages ASAP with minimum cost
- News page changes frequently
- Decide when to sense “Obama” and “Clinton” pages
Problem

- **Input:**
  - Web pages + change dependency diagram

- **Goal:**
  - Detect change ASAP
    - Minimum (or an approximate) number of sensing actions
Our Approach

- Model the problem as a **Partially Observable Markov Decision Processes**
  - Efficient algorithm using the special structure of our problem

- **Contribution:** tractable solution for sensing decisions
Decision Making as a POMDP

- **Belief state:** \{C, NC\}
  - C = 1: change occurs at this time
  - NC = 1: Not-captured change occurred since the last sensing action

- **Action:** a
  - Sense, idle

- **Reward function**
  - Positive reward for correct sensing
  - Penalty for sensing late
  - Error for wrong sense
Optimal Sensing Decision

- General POMDP approach
  - Assign value to each belief state
    - Immediate reward + expected future rewards
  - Optimal decision:
    - Maximize value function

- Our approach:
  - Change value function representation
    - Use properties of this new representation
  - Tractable solution for sensing decision
Value Function Representation

- Sum of rewards up to the first sensing action + value of a fixed belief state

\[
V(b_t) = \max_x \left\{ P(ch < t)(\text{Reward when } ch < t) \\
+ P(t < ch < t + x)(\text{Reward when } t < ch < t + x) \\
+ P(ch > t + x)(\text{Reward when } ch > t + x) \\
+ \gamma^{x+1}V(b^*) \right\}
\]
Value Function Property

- **Theorem:**
  - Let $V(b) = \max_x f(sense@t + x)$
  - If $f(sense@t + 1) < f(sense@t)$ then
    $$f(sense@t + m) < f(sense@t) \text{ for every } m > 1$$

- **Implies**
  - The first time that $f(.)$ starts to decrease is the time to sense
Proof Intuition

- Expand: \[ f(sense @ t + x) - f(sense @ t) \]

- Notice:
  1) \[ f(sense @ t + 1) < (sense @ t) \]
  2) \[ P(ch > t + x) < P(ch > t + 1) \]
Algorithm (One Page No Observation)

- **Greedy**: Decide whether to sense or not at the current time step $t$

  - **Algorithm**:
    
    If $f(sense@t) - f(sense@t+1) < 0$
    
    sense
    
    else : stay idle
Complexity

\[ f(sense@t + 1) - f(senes@t) \]

- \( P(nc_t, nc_{t+1}) \) : One step progression

- \( V(b^*) \) : Computed offline
  - \( b^* \) : belief state that we have sensed at previous time
Model:

- K: hidden and observable nodes
- Fully observable nodes in K
- C has no parent from K nodes
- Example: HMM
Algorithm (One Page with Observations)

- Approximation of the optimal policy
  - Assumption: decision about the immediate action is independent of the future observations
  - Optimal solution given the current observations
Algorithm
(One Page with Observations)

- Greedy algorithm:
  - Sense the page if value function starts to decrease
  - All the probabilities are conditioned given observations

- Theoretical Validation:
  - Still, \( P(ch > t + x \mid \text{obs}) < P(ch > t + 1 \mid \text{obs}) \)
  - Therefore,

\[
\text{If } f(sense \@ t + 1) < f(sense \@ t) \text{ then } \\
f(sense \@ t + m) < f(sense \@ t) \text{ for every } m > 1
\]
Multiple Sensing Variables

Example: factored HMMs

Sensing variables

O: observation node
Approximate Algorithm (Multiple Pages)

- **Goal:**
  - Approximate the optimal composite policy
    - Find subset of pages to sense

- **Algorithm:**
  1. Find the policy for each page
  2. Merge the results
Theoretical Verifications

- Value function for the composite policy = sum of value functions for single policies
  \[ V(b) = V_1(b_1) + V_2(b_2) \]

- Proof Intuition:
  - \( V^k(b) \) value function for the k-step policy
    \[ V^k(b) = V_1^k(b_1) + V_1^k(b_2) \]
  - Limit \( k \to \infty \)
Theoretical Validation

- Greedy algorithm works for each single policy

- Proof Intuition:

\[ P(ch > t + x, \text{observations}) < P(ch > t + 1, \text{observation}) \]
Experiments

- Simulated data
  - Randomly generate change prior, observation model, reward
  - Randomly generate a sequence of change and observations
  - Compute sum of rewards for the sensing decisions

- Wikipedia pages
Running Time Comparison

- **MOCHA**
- **Perseus**
- **Witness**
- **Our algorithm**

The graph shows the running time comparison for different algorithms as the number of states increases.
- Higher value for the policy returned by our algorithm vs. two other POMDP algorithms
• A (strong) assumption: the set of observable pages (nodes) is a vertex cover of $G$. 

A connectivity graph  

Our assumption  

A set of Factored HMMs  

Data from Wikipedia
Data from Wikipedia

- Three wikipedia pages
  - Observable: Democratic Party presidential primaries, 2008
  - Hidden: Barack Obama, Hillary Clinton

- History of each page
  - 10,000 updates for since 2004
  - Descretize time (1 hour period)
Results on Wikipedia

- (Value) sum of rewards vs. time step
- Oracle: perfect scenario (captures change whenever occurs)
Conclusions

- Formalization to the problem of detecting change using POMDPs
- Tractable algorithm for sensing decisions

- Limitations:
  - Do not know the bound of approximation compared to the oracle
Future Work

- Include actions that change the world (e.g. moving)
- Extend to the case of multiple sensing variables with no constraint over dependencies
- Learn the model
Value Function Representation

\[ V(b_t) = \max_x ( \]
\[ P(ch < t). ( \text{Reward for capturing change + penalty for sensing late}) \]
\[ + P(t < ch < t + x)(\text{Reward for capturing change + expected penalty for sensing late}) \]
\[ + P(t + x < ch) \text{ (Penalty for error sensing)} \]
\[ + \gamma^{x+1} V(b^*) \]

) \text{ b*: Belief state of the system after the sensing action}
Complexity

\[ f(\text{sense}@t + 1) - f(\text{senes}@t) = \]
\[ P(C' = 1)(\text{constant}) \]
\[ + P(C_t = 0, C_{t+1} = 1)(\text{constant}) \]
\[ + P(C_t = 0, C_{t+1} = 0)(\text{constant}) \]
\[ + (\gamma - 1)V(b^*) \]

- One step progression

- \( b^* \): belief state that we have sensed at previous time
  - \( V(b^*) \) Computed offline