Incorporating User State in Task Based Information Retrieval and Recommendation of Educational Resources

Emine Yilmaz

Turing Fellow and Professor, University College London
Amazon Scholar, Amazon

Emine.Yilmaz@ucl.ac.uk
Why User State?

• In many search and recommendation settings users have complex needs

• Accurate representation of user state highly important for providing them with correct information

• Many applications of state based user modelling

• Our focus:
  – Task Based Information Retrieval
  – Recommendation of educational resources
• Users use IR systems to achieve some real world tasks
• Most of these tasks consist of several subtasks
• User navigates through the subtasks to complete the task
Users use IR systems to achieve some real world tasks

Significant effort required using existing systems [He & Yilmaz, ACM CHIIR ‘17]
-

- Incorporating state information to reduce effort for task completion
  - Form representations of tasks and associated subtasks
  - Map the user to their corresponding state in the task/subtask space
  - Provide the user with contextual task completion assistance, given their state
Queries previously issued by users contain information about tasks users use the search engines for.

Use ML techniques to infer task representations from query logs.
Forming Task Representations

- mortgage interest rates
  - credit history check
  - mortgage interest rates
  - solicitor fees moving
  - solicitor comparison
  - renovate house cost
  - renovate house interior
  - renovate house yourself
  - Interior designers
  - Interior designer cost
  - how to renovate a house
  - renovate house yourself
  - renovate house online
Most previous work represents tasks as flat structures [Li et al. KDD’14; Wang et al. WWW’13; Verma and Yilmaz CIKM’14; White et al. CIKM’14]

Most complex tasks involve multiple subtasks, which can themselves be complex tasks

Tasks can best be represented as a hierarchy
Constructing Task Hierarchies

• One possibility: Hierarchical clustering methods
  – Most construct binary tree representations of data
  – No guide on the correct number of clusters

• Need models that can represent trees with arbitrary branches
  – Complexity is a major problem
Hierarchical Task Extraction
[Mehrotra and Yilmaz, SIGIR ‘17]

Rose Trees
Hierarchical Task Extraction
[Mehrotra and Yilmaz, SIGIR ‘17]

Bayesian Rose Trees [Blundell et al. NIPS’13]

\[ p(D \mid T) = \sum_{\phi(T) \in \text{Part}(T)} p(\phi(T)) p(D \mid \phi(T)) \]
Hierarchical Task Extraction
[Mehrotra and Yilmaz, SIGIR ‘17]

• Bayesian Rose Trees (BRTs) for task extraction

• A tree (T) represented as a mixture of partitions $\phi$ over a group of queries (Q)
  – Each node of the tree corresponds to a task
  – Each task represented by a set of queries

• Goal: Find the tree structure that maximizes
  $$p(Q | T) = \sum_{\phi(T) \in Part(T)} p(\phi(T)) p(Q | \phi(T))$$
Hierarchical Task Extraction
[Mehrotra and Yilmaz, SIGIR ‘17]

- Need to compute
\[ p(Q | T) = \sum_{\phi(T) \in \text{Part}(T)} p(\phi(T)) p(Q | \phi(T)) \]

- Number of partitions consistent with T can be exponentially large

- Approximate using dynamic programming [Blundell et al. NIPS ‘13]
\[ P(Q | T) = \pi_T f(Q_T) + (1 - \pi_T) \prod_{T_i \in ch(T)} p(\text{leaves}(T_i) | T_i) \]
Hierarchical Task Extraction
[Mehrotra and Yilmaz, SIGIR ‘17]

• Need to compute

\[ p(Q \mid T) = \sum_{\phi(T) \in \text{Part}(T)} p(\phi(T)) p(Q \mid \phi(T)) \]

• Number of partitions consistent with \( T \) can be exponentially large

• Approximate using dynamic programming [Blundell et al. NIPS ’13]

\[ P(Q \mid T) = \pi_T f(Q_T) + (1 - \pi_T) \prod_{T_i \in \text{ch}(T)} p(\text{leaves}(T_i) \mid T_i) \]
Hierarchical Task Extraction
[Mehrotra and Yilmaz, SIGIR ‘17]

• Need to compute

\[ P(Q \mid T) = \sum_{\phi(T) \in \text{Part}(T)} p(\phi(T)) p(Q \mid \phi(T)) \]

• Number of partitions consistent with \( T \) can be exponentially large

• Approximate using dynamic programming [Blundell et al. NIPS ‘13]

\[ P(Q \mid T) = \pi_T f(Q_T) + (1 - \pi_T) \prod_{T_i \in \text{ch}(T)} p(\text{leaves}(T_i) \mid T_i) \]
How to Utilise Task Representations?

• Given an input from the user, map the user to their corresponding state in the task & sub-task space

• Incorporate this information to
  • Improve search results
  • Make better query suggestions
  • Provide more targeted response in conversational systems
  • Provide proactive results
  • Provide better personalization
How to Utilise Task Representations?
Current Work: Incorporating State Transitions

- Users usually transition through several subtasks
- There is usually an order between these tasks
- Incorporate this order information to infer user state
State Based User Modelling in Recommendation of Education Resources

• More and more learning material becomes available online

• How to recommend correct material to the learner at the correct time?

• Need to have a model of
  – Learner’s knowledge state
    • Learner has to have enough prior knowledge to understand the material
  – Contents of the learning material
Predicting User Engagement from User State

• A Family of Bayesian Algorithms inspired by the TrueSkill algorithm [Herbrich et. al (NIPS’07)]

• Retains a humanly-intuitive learner representation

• Infers learner knowledge state

• Incorporates novelty of educational resources

[Bulathwela, Perez-Ortiz, Yilmaz and Shawe-Taylor (AAAI ’20a)]
Inspiration: TrueSkill

A graphical model aiming to capture the skills of players exclusively from the output of (two-player) games [Herbrich et. al (NIPS’07)].
Recommendation as a Competition

• Form a representation of the learner’s knowledge (skills)

• Form a representation of the learning material (difficulty, coverage)

• Cast the problem as a competition
  – Learner wins (is engaged) if the material is within the skill level of the learner
    • Not too easy (has enough novelty)
    • Not too difficult

• Update the skill level of the learner if the learner is engaged
Hypotheses to model learner’s engagement as a function of knowledge:

• If the learner is engaged, they have enough background to make use of the resource.

• No assumption can be made from the non-engaged cases (the learner might not be engaged for a myriad of reasons).

• If the learner is engaged, they must have the appropriate background and the content must also be novel to them (within an engagement margin).
Hypotheses tested in TrueLearn [AAAI 2020]
Hypotheses tested in TrueLearn [AAAI 2020]

\[ P(\text{engagement}|\theta_\ell, d_r) = f(|\theta_\ell - d_r| \leq \varepsilon) \]
Representing State

- Represent educational resources and learner state using **knowledge components (KCs)**
  - Knowledge components (KCs): **atomic units of learnable concepts**
- KCs obtained through Wikification
  - Identification of Wikipedia concepts (articles) present in the resource by connecting text to Wikipedia concepts via entity linking [SiKDD ’17]
Representing State

• Knowledge components obtained through Wikification

• Matching done using two features
  • Pagerank score of the concept
  • Cosine similarity of the resource to the concept

• Cosine similarity used to represent resource depth (difficulty)
Experimental Results

- Dataset: VideoLectures.Net from December 2016 - February 2018
  - 3,884 Lectures
  - 18,933 unique users and 248,643 video view events

- Representations obtained based on video text
  - Text representation partitioned into fragments of ~5000 characters (~5 mins)
    - Can record engagement of learners in a more fine-grained manner
    - Can associate more specific KCs to the parts of lectures
  - Top 5 Wikipedia concepts associated with an individual lecture fragment

- Binarize normalised learner engagement [Guo et. al (L@S’14)]
<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vanilla TrueSkill</td>
<td>.610</td>
<td>.541</td>
<td>.472</td>
<td>.480</td>
</tr>
<tr>
<td>TrueLearn DynamicDepth</td>
<td>.454</td>
<td>.530</td>
<td>.431</td>
<td>.418</td>
</tr>
<tr>
<td>TrueLearn FixedDepth</td>
<td>.736</td>
<td>.610</td>
<td>.558</td>
<td>.573</td>
</tr>
<tr>
<td>TrueLearn Novelty</td>
<td>.649</td>
<td>.603</td>
<td>.835</td>
<td>.677</td>
</tr>
</tbody>
</table>
Current Work

• Present the learner state to the learner, enable the learner to update the state

• Incorporate winning margin

• Exploit the Wikipedia hierarchy structure

• Model engagement on the full video resource
Generalisation to Information Gathering Tasks

• Similar models could be used for information gathering tasks in search and conversational systems

• Users start with some initial information, which gets updates as the user gathers more information

• Need a way to incorporate user’s prior knowledge before using the system
Conclusions

• Accurate representation of user state highly important for providing users with correct information

• Many applications of state based user modelling

• Our focus:
  • Task Based Information Retrieval
  • Recommendation of educational resources