1. Background

- In a nutshell, submodular functions are the class of functions that exhibit diminishing returns. As such, many machine learning applications fall under the umbrella of submodularity:

\[ f(A \cup \{v\}) - f(A) \geq f(B \cup \{v\}) - f(B) \]

- Mathematically, a function is said to be submodular if for all sets \( A \subseteq B \) and all elements \( v \in V \setminus B \):

\[ f(A \cup \{v\}) - f(A) \geq f(B \cup \{v\}) - f(B) \]

- In other words, the marginal value of any item is non-increasing as our set grows. For example, suppose we want to summarize a set of images about Spain. Once we already have one image of Barcelona, additional images of Barcelona will be much less valuable.

2. Problem Statement

- Our goal is to maximize submodular functions where the order of selected items matters. We follow the setup of Tschiatschek et al. (2017).

- Our input consists of:
  - a directed hypergraph \( H = (V, E) \), where the vertices \( V \) represent the items, and the hyperedges \( E \) encode the additional value of selecting items in a particular order.
  - a monotone submodular function \( h: 2^E \rightarrow \mathbb{R}_{\geq 0} \).
  - a cardinality constraint \( k \).

- Our output is a sequence \( \sigma \) of \( k \) unique vertices that maximizes the objective function:

\[ f(\sigma) = h(E(\sigma)) \]

where

\[ E(\sigma) = \{ e \in E \mid \sigma \cap V(e) = e \} \]

Informally, \( E(\sigma) \) contains an edge \( e \in E \) if and only if all vertices of \( e \) appear in \( \sigma \) in the proper order.

- For example, the graph below encodes the idea that the order in which one watches the Lord of the Rings franchise matters. In this case, \( h(E(\sigma)) = |E(\sigma)| \):

\[ f(F) = h((F, F)) = 1 \]
\[ f(T) = h((T, T)) = 1 \]
\[ f(F, T) = h((F, F), (F, T), (T, T)) = 3 \]
\[ f(T, F) = h((F, F), (T, T)) = 2 \]

3. Algorithm and Theoretical Results

- Our algorithm greedily selects the most valuable valid hyperedge at each step:

1. Let \( \sigma \leftarrow () \).
2. while \( |\sigma| \leq k - r \) do
   1. Let \( E = \{ e \in E \mid \sigma \cap V(e) \text{ is a prefix of } e \} \).
   2. if \( E = \emptyset \) then Exit the loop.
   3. \( e^* = \arg \max_{e \in E} h(e \mid E(\sigma)) \).
   4. for every \( v \in e^* \) in order do
      1. if \( v \notin \sigma \) then \( \sigma = \sigma \cup v \).
   5. return \( \sigma \).

- The theoretical guarantees of our algorithm are comparable to \( \text{OMeGA} \), the only other known algorithm for this scenario. However, our guarantees apply to any general hypergraph, rather than just directed acyclic graphs:

<table>
<thead>
<tr>
<th>Algorithm Name</th>
<th>Approximation Guarantee</th>
<th>Running Time</th>
<th>Graph Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>OMeGA</td>
<td>( 1 - e^{-\frac{1}{\Delta}} ) ( O(\Delta mk \log k) )</td>
<td>Directed Acyclic Graphs</td>
<td></td>
</tr>
<tr>
<td>Hyper Sequence-Greedy</td>
<td>( 1 - e^{-\left(1-\frac{1}{r}\Delta+1\right)} ) ( O(\Delta m k) )</td>
<td>General Hypergraphs</td>
<td></td>
</tr>
</tbody>
</table>

- \( d_{\text{in}} \) is the maximum in-degree, \( d_{\text{out}} \) is the maximum out-degree, and \( \Delta = \min\{d_{\text{in}}, d_{\text{out}}\} \).
- \( m = |E| \), the number of edges in the graph.
- \( r \) is the maximum hyperedge size.

4. Applications

- Movie Recommendation
  - We use the Movielens 1M dataset (Harper and Konstan, 2015), which contains 1,000,209 anonymous, time-stamped ratings made by 6,040 users for 3,706 different movies.
  - Given the first \( B \) movies that a user has reviewed, our goal is to predict the next \( k \) movies she will review.

- Massive Open Online Course Planning
  - We use the open edX dataset (Ho et al., 2014), which contains 641,139 registrations from 476,532 unique users across 13 different online courses.
  - Our goal is to select a sequence of 4 courses that will be as valuable as possible to users interested in those fields.

5. Acknowledgements

We acknowledge support from DARPA YFA (D16AP00046), AFOSR YIP (FA9550-18-1-0160), ISF grant 1357/16, and ERC StG SCADAPT.