Deep RL Meets Structured Prediction
Structured Prediction Often Needs to Do Search

Which fits well with RL’s ability to form good search policy

- dialogue
- semantic parsing
- program synthesis
- architecture search
- machine translation
- summarization
- image caption
- knowledge graph reasoning
- information extraction
- ...

...
RL is attractive:

Directly Optimizing The Expected Reward
Which can be very useful for structured predictions

- **ML** optimizes the log likelihood of target sequences

\[ J^{ML}(\theta) = \sum_q \log P(a_{0:T}^{best}(q)|q, \theta) \]

- **RL** optimizes the expected reward under a stochastic policy

\[ J^{RL}(\theta) = \sum_q \mathbb{E}_{P(a_{0:T}|q,\theta)}[R(q, a_{0:T})] \]

[Williams 1992]
[Sutton & Barto 1998]
[Liang+ 2017]
RL has challenges:

Which we need to be aware of

- Large search space (sparse rewards)
  - Supervised pretraining (MLE)
  - Systematic exploration [Houthooft+ 2017]
  - Curiosity [Schmidhuber 1991][Pathak2017]

- Credit assignment (delayed reward)
  - Bootstrapping
    - E.g., AlphaGo uses a value function to estimate the future reward
  - Rollout n-steps

- Train speed & stability (optimization)
  - Trust region approaches (e.g., PPO)
  - Experience replay
<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:30 AM</td>
<td>Poster setup (posters will be up in the room all day.)</td>
</tr>
<tr>
<td>9:45 AM</td>
<td>Opening remarks</td>
</tr>
<tr>
<td>9:50 AM</td>
<td><strong>Jessica B. Hamrick</strong>: Structured Computation and Representation in Deep Reinforcement Learning</td>
</tr>
<tr>
<td>10:25 AM</td>
<td>Advertising contributed talk &amp; Best paper award (Zhiting Hu, 5min)</td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Coffee &amp; Poster (sync w/ ICLR conference)</td>
</tr>
<tr>
<td>11:00 AM</td>
<td><strong>Anima Anandkumar</strong>: Infusing Structure into Machine Learning Algorithms</td>
</tr>
<tr>
<td>11:35 AM</td>
<td><strong>Graham Neubig</strong>: What can Statistical Machine Translation teach Neural Machine Translation about Structured Prediction?</td>
</tr>
<tr>
<td>12:10 PM</td>
<td><strong>Mohammad Norouzi</strong>: Beyond Off-the-shelf Reinforcement Learning for Structured Prediction</td>
</tr>
<tr>
<td>12:45 PM</td>
<td>Advertising contributed talk (Wouter Kool, Zafarali Ahmed, Osbert Bastani, 3x5min)</td>
</tr>
<tr>
<td>13:00 PM</td>
<td>Lunch &amp; Poster</td>
</tr>
<tr>
<td>14:20 PM</td>
<td>Workshop on Deep Generative Models for Highly Structured Data</td>
</tr>
</tbody>
</table>