Tactical Rewind:

- Liyiming “Kay” Ke (kayke@uw.edu), Xiujun Li, Yonatan Bisk, Ari Holtzman, Zhe Gan, Jingjing Liu, Jianfeng Gao, Yejin Choi, Siddhartha Srinivasa

INTRO
- The Vision-and-Language Navigation task, VLN, builds robots to navigate in houses from language instructions.
- Many approaches formulate it as a seq2seq task, build neural network and use either greedy or beam search decoding.
- We invented a plug-n-play procedure, FAST, that lets us reuse all the existing or new neural models – but more successfully and efficiently.

METHOD
- **Frontier-Aware Search with backtracking (FAST)**
- As the agent explores, it asks:
  - Did I reach the target?
  - Am I lost?
  - Should I backtrack?
  - Where to backtrack to?
- To answer, we leverage existing neural models but only change the decoding schema.
- View the search running:

Search as a **global** framework, **Neural Network as a local heuristic.**

```
<table>
<thead>
<tr>
<th>Model</th>
<th>TL</th>
<th>VR</th>
<th>SR</th>
<th>SPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMNA</td>
<td>16.04</td>
<td>5.67</td>
<td>0.48</td>
<td>0.35</td>
</tr>
<tr>
<td>+ FAST</td>
<td>14.82</td>
<td>6.62</td>
<td>0.35</td>
<td>0.28</td>
</tr>
<tr>
<td>Beam Search</td>
<td>22.08</td>
<td>5.14</td>
<td>0.54</td>
<td>0.41</td>
</tr>
<tr>
<td>SMNA</td>
<td>37.09</td>
<td>4.48</td>
<td>0.61</td>
<td>0.02</td>
</tr>
<tr>
<td>+ FAST</td>
<td>1,257.30</td>
<td>4.87</td>
<td>0.53</td>
<td>0.01</td>
</tr>
<tr>
<td>Beam Search</td>
<td>196.53</td>
<td>6.29</td>
<td>0.61</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Haman: 11.35, 1.61, 0.86, 0.76
```

Scan for code & paper

**RESULT**
- The simple change to decoding enables **immediate gains without modifications** to existing models.
- We took the previous SoTA and achieved the **new state-of-the-art** in SPL at the time of submission. Shown on test unseen.

**DISCUSSION**
- FAST **generalizes** between greedy decoding and beam search.
- Compared with Beam Search, FAST is more **efficient**.
- Compared with Greedy decoding, FAST is more **robust**.
- FAST also offers a simple knob that one can tune to trade between success rate and efficiency.

The Frontier

“Exit the living room, pass the hallway and stop in the bedroom.”

NN jointly scores the partial or completed trajectory

Scan for code & paper