**THE TASK**
- Find the **intent** and **slots** in a user utterance - core task in digital voice assistants like Alexa, Siri, Cortana etc.
- Utterances can be **simple** with flat slots, or **complex** with hierarchical slots

Play the song **don’t stop believin** by **Journey**

**OUR APPROACH**
- Use sequence to sequence models to solve this task
- Use a Pointer Generator Network for source tokens
- Easy to see how complex utterances can be framed

**TRADITIONAL APPROACHES**
- **Slot Filling mechanisms** for simple utterances – CRFs, Bi-LSTMs with CRFs, Capsule networks
- **Recurrent Neural Network Grammars** for complex utterances – Shift-reduce parsers
- All these models only work on **specific grammars**
- These models cannot process slots that are overlapping or non-consecutive
- Our model is **unified** and can process slots in any kind of tagging format

**RESULTS**
- Evaluate on 5 datasets – Alexa Music, Alexa Video, SNIPS, ATIS, and Facebook TOP
- Metric – Exact match accuracy. The entire semantic parse should match
- Best previous baseline models – BI-LSTM with CRF for Alexa, **Capsule Network** for SNIPS & ATIS, **Shift-Reduce ensemble with ELMo and SVM reranking** for TOP

**CONCLUSION**
- We achieve improved performance on all five datasets
- We have SOTA performance for 3 public datasets – SNIPS, ATIS, and Facebook TOP
- Our model is unified for both simple and complex utterances without any grammar specifications
- We can also handle overlapping and non-consecutive slots

**REFERENCES**