Constrained Inference for Bridging the Distributional Gap in Natural Language Processing

Tao Meng
Advisor: Kai-Wei Chang
Distributional Gap

- Distribution in model predictions is deviated from that in the ground truth.
- Widely exists in real-world applications
  - Training data is not drawn from the same distribution of test
    - Test domain is naturally low-resource
    - E.g., cross-lingual transfer, English → Urdu
  - NLP models are large. One model is applied in different tasks and domains.
    - BERT-Large, GPT-3, etc.
    - Training is computational expensive
- Causing systematic errors
Solution: constraints

- Constraints are a set of rules for the model predictions
  - Compiling domain knowledge
  - Predications satisfy the constraints

Constraint: In an ADP-NOUN arc in Hindi, ADP is more likely to be on the right.

- Boost performance (English transfer to Hindi: 34% → 53%)
Research Questions

• What kind of constraints can better bridge the gaps

• How to formulate constraints

• How to incorporate constraints with natural language processing models
Goal

• In results:
  • Constraints guide the model to make predictions in our expected distribution

• In method:
  • Various of NLP models are able to incorporate constraints efficiently (e.g., without retraining / fine-tuning the models)

• In practicality:
  • Constraints are easy to acquire
  • Constraints are effective yet controllable
Progress and Future Plan

• What I have done…
  • Word order constraints in cross-lingual transfer dependency parsing[1]
  • Gender ratio constraints in gender bias amplification mitigation[2]
  • A integer linear programming framework for mining constraints from data[3]

• Future plan:
  • Incorporate constraints in natural language generation (NLG)
  • Automatically learn general constraints from data