**Summary**

- **Motivation:** Domain Constraint & Inflexible Intent Schema
  - Intents are usually predefined and inflexible to expand and transfer across domains, where re-designing intent semantic schemes requires human annotation and model re-training.
- **Approach:** Learning Intent Embedding
  - Applying CDSSM to learn high-level semantic representations to address the semantic relation across domains for intent expansion (e.g. “find movie” and “find weather” belong to different domains, but they share the semantics about “find”).
- **Result**
  - CDSSM is capable of performing zero-shot learning effectively, e.g. generating embeddings of previously unseen intents, and therefore expand to new intents without re-training, and outperforms other semantic embeddings.

**1. Framework**

- Training Data
  - `<change_note>`: “postpone my meeting to five pm”
  - `<change setting>`: “temperature...”
  - `<change_calender>`: “what’s the weather...”

- CDSSM
  - Embedding Generation
  - New Intent

**2. Convolutional Deep Structured Semantic Models (CDSSM)**

- **Model Architecture**
  - **Semantic Layer:** feed-forward neural network layers outputs the final non-linear semantic features
  - **Max Pooling Layer:** only retain the most prominent local features by applying the max operation over each dimension of $I_t$ to keep the max activation of hidden topics across the whole word sequence

- **Convolutional Layer:** contextual features $c_t$ for each target word
  - **Convoluation Matrix:** $W_i$ one-hot vector $\rightarrow$ tri-letter vector (e.g. “email” $\rightarrow$ “emai”, “ema”, “mai”, “ail”)
  - **Word Sequence:** $x$ user utterance / intent

**3. Experiments**

- Dataset: collected via the Microsoft Cortana (> 100 intents)
  - Segmented into seen and unseen intents
    - Unseen: randomly chose 7 intents with different verbs: ~100K utterances
    - Seen: ~1M annotated utterances (2/3 for training CDSSM, 1/3 for testing)
- **Intent Prediction**
  - For each utterance vector, the semantic similarity can be estimated using vectors for both seen and unseen intents.
  - The unseen intent vectors can be generated from CDSSM by feeding the tri-letter vectors of the new intent as input without model re-training.
- **Evaluation Metrics:** Mean average precision at K (MAP@K)

**Conclusion**

- A convolutional deep structured semantic model (CDSSM) is applied to perform zero-shot learning of intent embeddings to bridge the semantic relation across domains.
- The experiments of intent expansion show that CDSSM can
  - capture the semantics borrowed from other domains and can be used to expandly extend the intents through high-level representations
  - carry the crucial high-level semantics and can be applied to different domains for easy adaptation and extension
  - generate more flexible intent embeddings without training samples and model re-training, removing the domain constraint in dialogue systems for practical usage.