Constrained Multi-Task Learning for Bridging Resolution

Hideo Kobayashi[1], Yufang Hou[2], Vincent Ng[1]

[1]University of Texas at Dallas, [2] IBM Research Europe - Ireland

Task Definition

- Bridging resolution aims to identify and resolve context-dependent but non-identical mentions
  - "**the customers**" requires the antecedent "*its*" (= "The Bakersfield Supermarket")
- More challenging than entity coreference resolution

S1: [The Bakersfield Supermarket]_new went bankrupt last May.

S2: [The business]_old closed when [its]_old old owner was murdered by robbers.

S3: [The murder]_old saddened [the customers]_bridging.

Existing Approaches

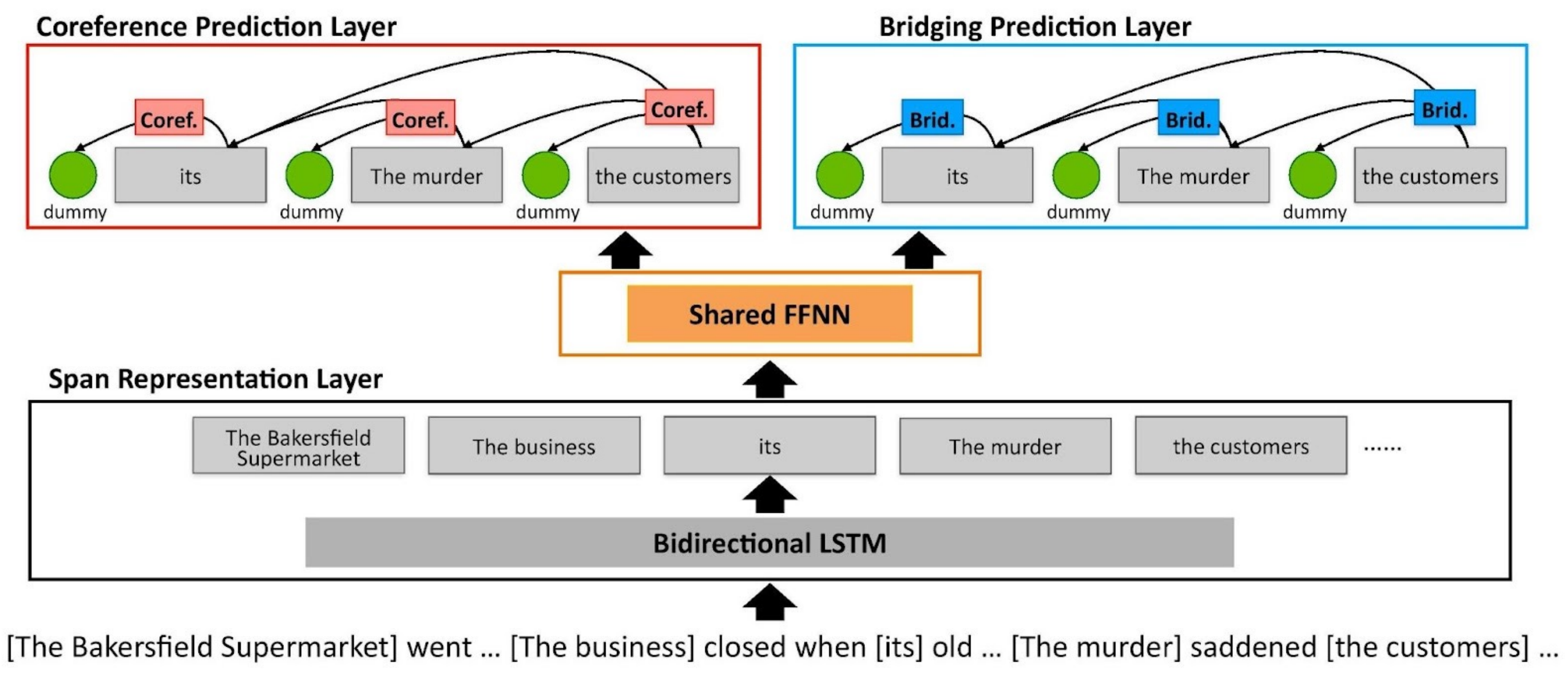
- Rule-based (Hou et al. 2014, Rosiger 2018): 18 rules; rulesets for different corpora are slightly different
- Neural (Yu and Poesio, 2020): Multi-task learning (MTL) of entity coreference and bridging resolution
- Hybrid (Kobayashi and Ng, 2021) [Current state of the art]: Applies Rules and MTL in a sequential manner

Goal

- Examine the extent to which supervised bridging resolvers can be improved by proposing four extensions to Yu & Poesio's (2020) multi-task learning framework

Yu & Poesio's (2020) Multi-Task Learning Framework

- A span-based MTL model of coreference and bridging
- Learn task-specific contextualized representations for text spans
- Create pair embeddings for each mention pair by combining span embeddings
- Pass them to shared FFNN layer and dedicated task prediction layers



Extension 1: Adding Information Status Classification in MTL

- Goal of Information Status (IS) classification: Assign IS to each discourse mention (Prince, 1981)
  - IS describes **the extent to which a noun phrase is available to the hearer**
  - **Old**: a mention has been explicitly mentioned before
  - **Mediated**: a mention has been implicitly mentioned before
    - A mention is **bridging** if it has not been introduced in the discourse directly, but it is inferable
  - **New**: a mention has not been mentioned before
- **Motivation**: Exploiting the close connection btw IS, bridging, and coreference resolution by introducing IS classification as third task in MTL
  - These tasks can benefit each other by having **a shared representation** in the span-based model

Extension 2: Incorporating Cross-Task Consistency Constraints

- **Motivation**: Merely exploiting the connection between tasks via **a shared representation** does not guarantee cross-task consistency constraints are enforced. Example constraints include:
  - If IS task predicts a span as **bridging**, then bridging task must predict it as **a bridging anaphor**.
  - If IS task predicts a span as **old**, then coreference task must predict it as **a coreference anaphor**.
  - If coreference task predicts a span as **a coreference anaphor**, then bridging task must NOT predict it as **a bridging anaphor**. and so on.

Extension 2: Incorporating Cross-Task Consistency Constraints (cont.)

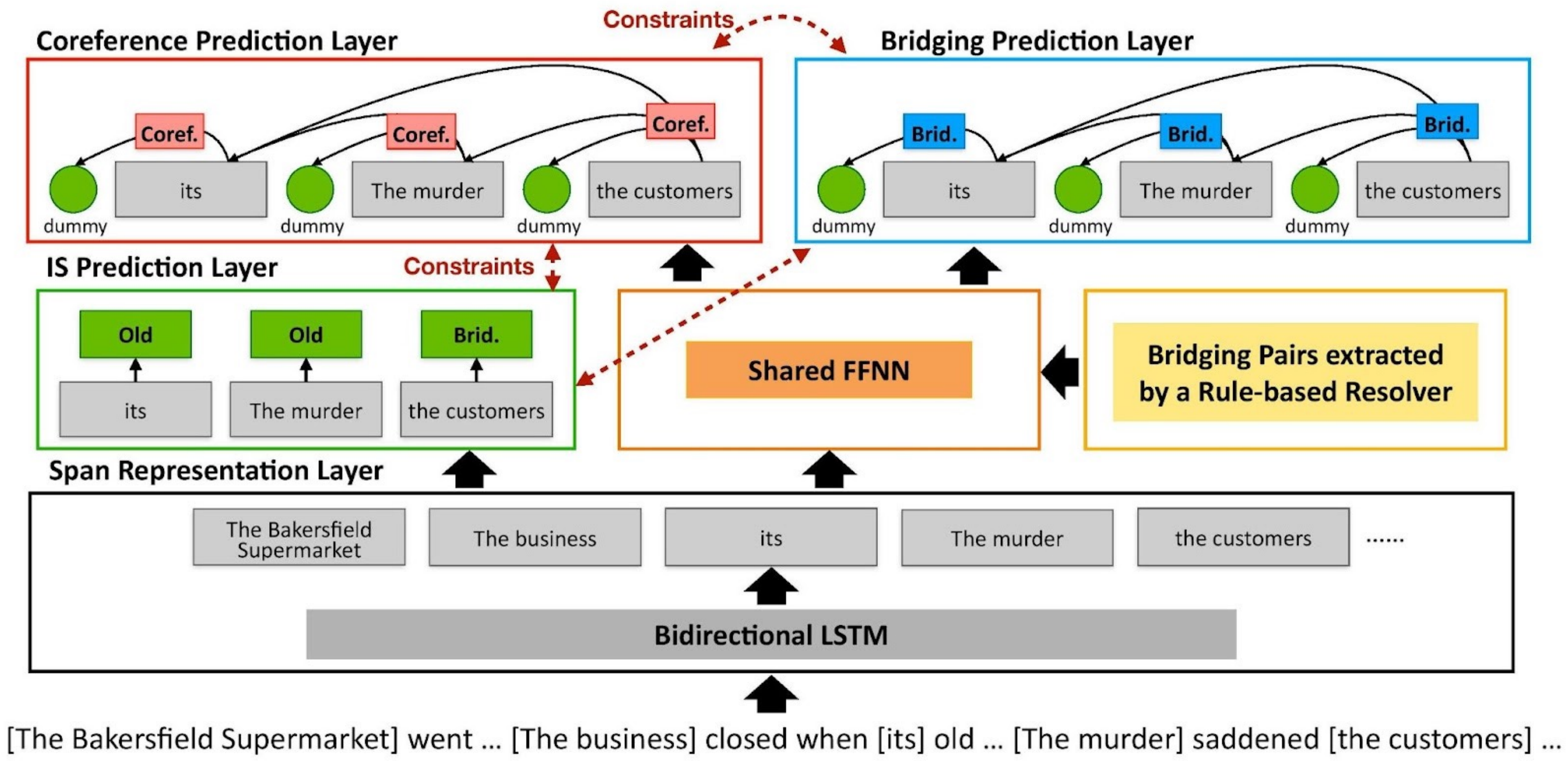
- Use **soft** rather than hard constraints since hard constraints could hurt performance
  - We define a penalty function for each consistency constraint that reduces the score of a mention pair if it violates a constraint
    - E.g., IS module predicts "the customers" as bridging but Bridging module does not
      - Bridging score will be reduced

Extension 3: Pre-Training Coreference

- **Motivation**: even soft constraints could still hurt performance
  - How can pre-training help address error propagation?
    - Reduce errors in the coreference module
      - Leverage the large amount of coreference data in OntoNotes

Extension 4: Incorporating Prior Knowledge (Rules)

- **Motivation**: annotated bridging data is small, so use rules to provide knowledge of bridging absent in data
- Incorporate prior knowledge (Rules) into the coreference and bridging modules as **a feature**
  - Encode whether a mention pair is predicted by a rule-based resolver
- Rules are **precision-** rather than recall-oriented
  - Add **a rule loss** to encourage the model to exploit rule information

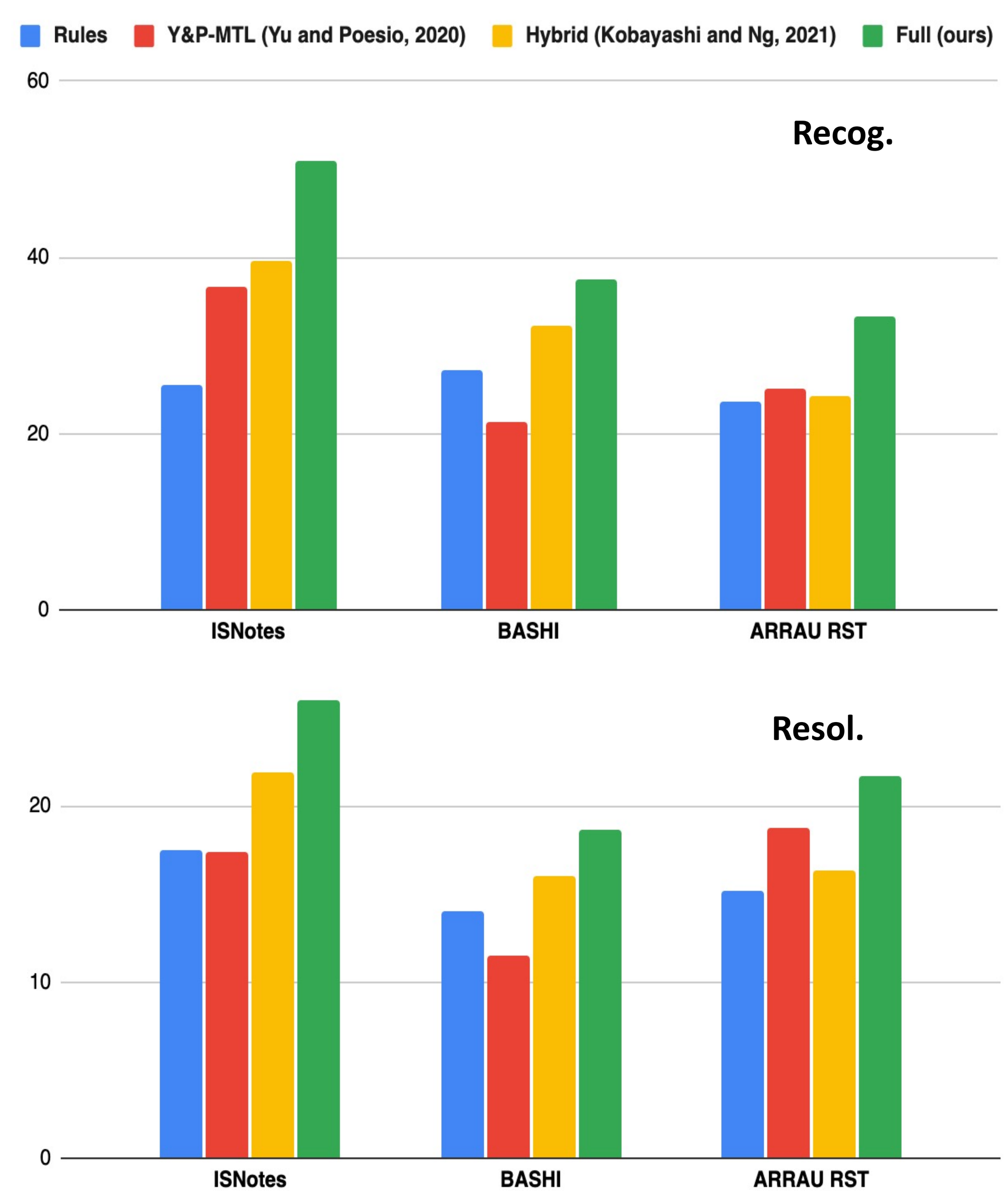


Evaluation Setup

- **Corpora**: ISNotes (50 WSJ news articles, 663 anaphors), BASHI (50 WSJ news articles, 459 anaphors), ARRAU RST (413 news docs, 3777 anaphors)
- **Setting**: Full bridging resolution, Input: gold mentions
  - Tasks: 1) Recognition: identify bridging anaphors, 2) Resolution: resolve them to their antecedents
- **Evaluation metrics**: Precision, recall, and F-score for recog. and resol.

Results

- Full model achieves the state of the art on resol. and recog. on all datasets
- Outperforms best baseline by **5.2-11.3%** (recog) and by **2.6-4.1%** (resol.)



Ablation Results

	Recog.	Resol.
Full Model	**50.9**	**26.0**
Consistency constraints	-4.8	-1.7
Soft→Hard	-3.2	-2.9
Rule loss	-1.1	-1.1
Rule loss + feature	-2.8	-2.4
Pre-training Coref.	-1.3	-5.8
Coref. task	-3.5	-3.4
IS task	-6.7	-3.3