

# Neural Snowball for Few-Shot Relation Learning

清华大学 高天宇

# I Background

- **Relation Extraction** (RE): extracting relational facts from text
  - E.g., **Bill Gates** founded **Microsoft** -> (**Bill Gates**, *founder*, **Microsoft**)
- Most RE systems target **pre-defined** relations
- However, our world undergoes **open-ended growth of relations**

# I Background

- **Supervised RE**

- Large-scale labeled data  $\rightarrow$  pre-defined relations

- **Semi-supervised RE**

- Medium-scale labeled data + large-scale unlabeled data  $\rightarrow$  pre-defined relations

- **Few-shot RE**

- Large-scale labeled data + few-shot labeled data  $\rightarrow$  new relations

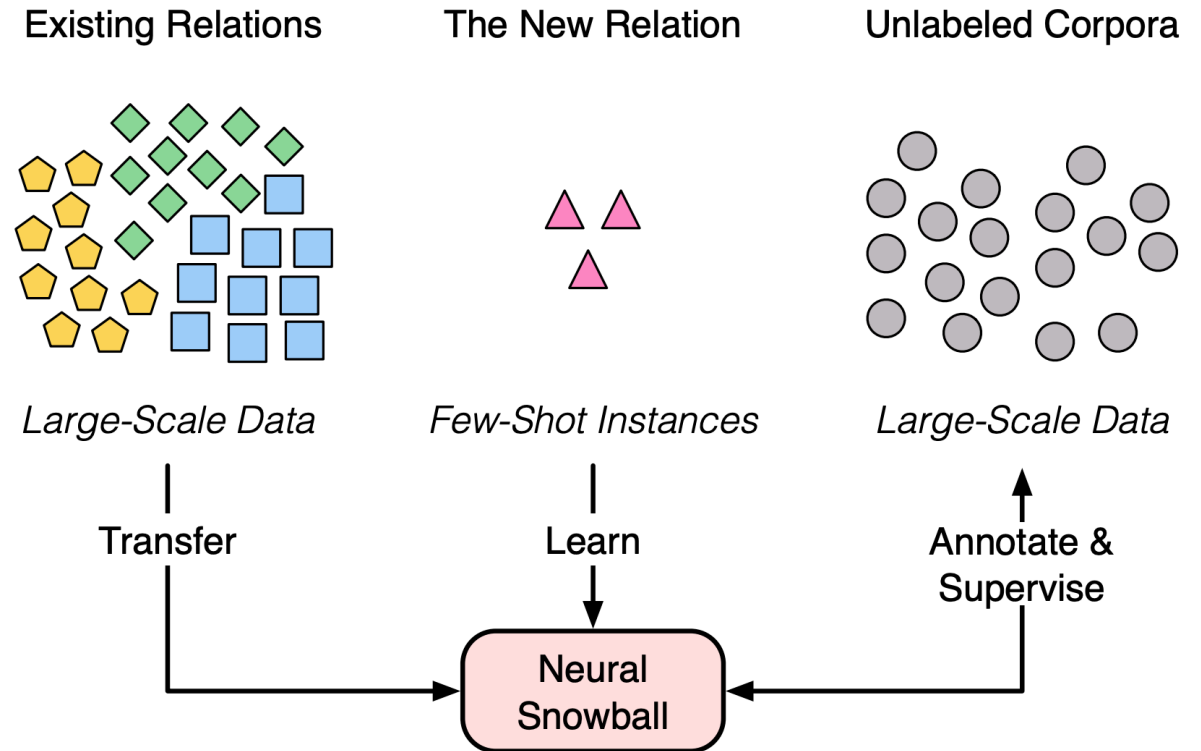
- **Bootstrapping RE**

- Few-shot labeled data + large-scale unlabeled data  $\rightarrow$  new relations

# I Utilizing All the Data We Have

- Ours

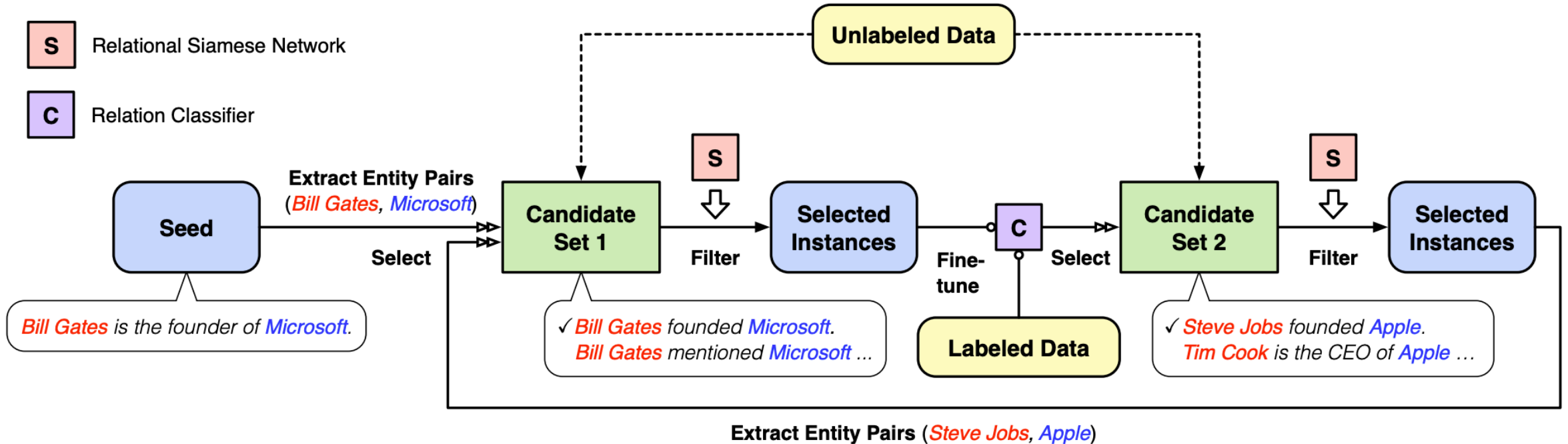
- Utilizing all three kinds of data
- **Large-scale labeled data**
  - Learning good relational metrics
- **Few-shot labeled data**
  - As golden instances
- **Large-scale unlabeled data**
  - Mining useful instances



# I Pipeline

- **Inputs:** A few annotated instances for the new relation (**seed**)
- **Target:** Learn a (binary) classifier for the new relation
- **How to:** Mine instances of the target relation from unlabeled data

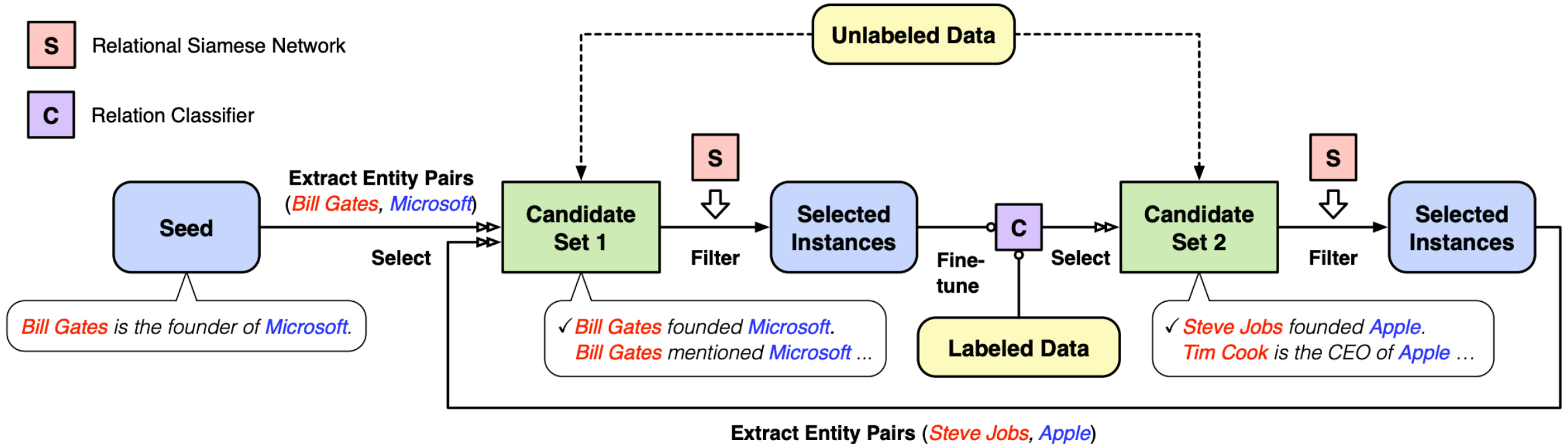
# I Pipeline



- **Step 1:** Use distant supervision to get candidates
- **Step 2:** Use the new relation classifier to get candidates

# I Pipeline

Filtered by the **Relational Siamese Network** (RSN)

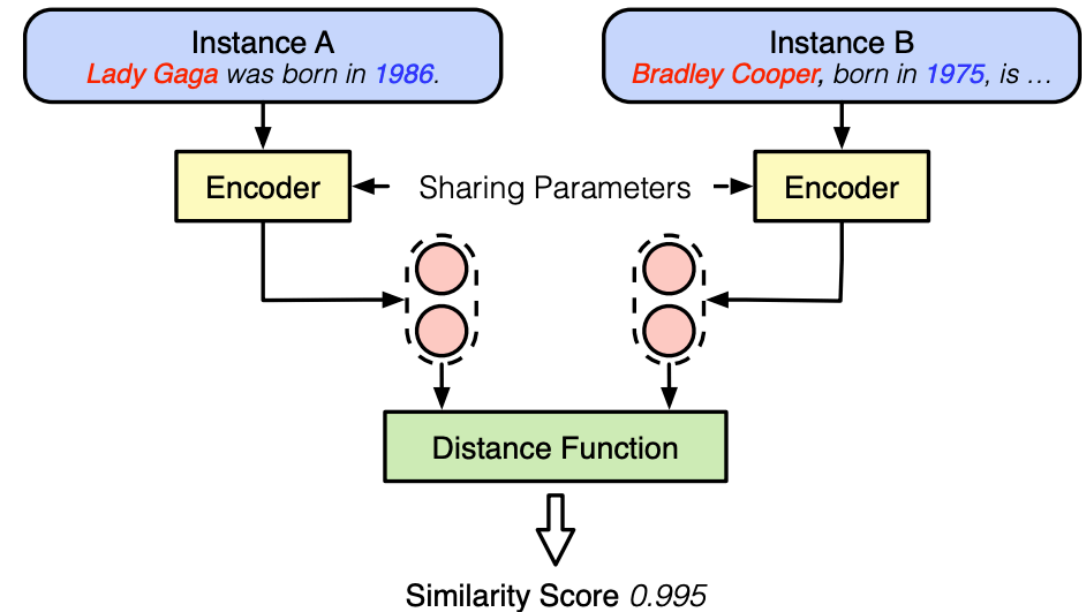


- **Step 1:** Use distant supervision to get candidates
- **Step 2:** Use the new relation classifier to get candidates

# I Relational Siamese Network

- Learn generalizable relational metrics
  - Pre-trained on large-scale data of **existing relations**
  - Experiment shows good transfer learning abilities

Relation Set	P@5	P@10	P@20	P@50
Train	83.60	80.66	76.03	61.98
Test	82.15	78.64	72.57	55.10





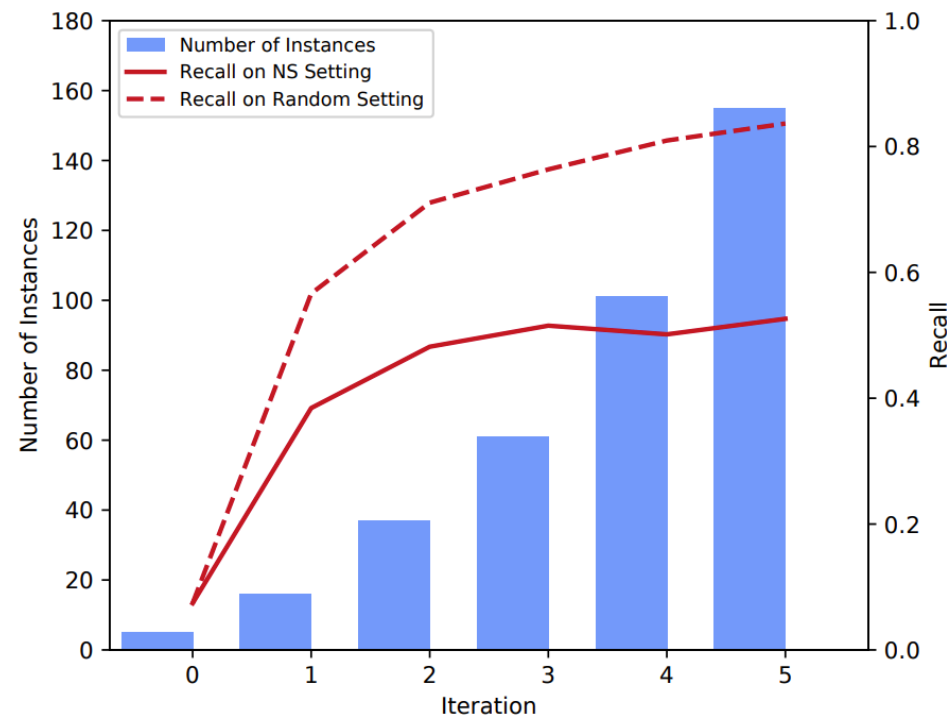
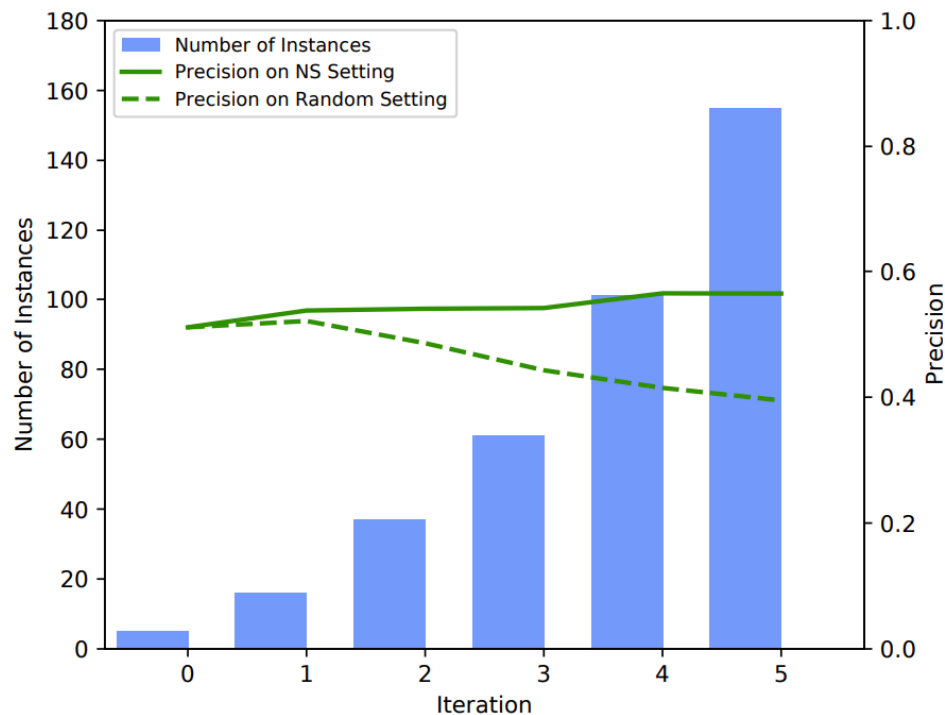
# I What is our edge

- Neural methods bring good **generalizability**
- RSN makes sure we have good **precision**

# | Experiment

Model	5 Seed Instances			10 Seed Instances			15 Seed Instances		
	P	R	F1	P	R	F1	P	R	F1
BREDS	33.71	11.89	17.58	28.29	17.02	21.25	25.24	17.96	20.99
Fine-tuning (CNN)	46.90	9.08	15.22	47.58	38.36	42.48	74.70	48.03	58.46
Relational Siamese Network (CNN)	45.00	31.37	36.96	46.42	30.68	36.94	49.32	30.46	37.66
Distant Supervision (CNN)	44.99	31.06	36.75	42.48	48.64	45.35	43.70	54.76	48.60
<b>Neural Snowball (CNN)</b>	48.07	36.21	41.30	47.28	51.49	49.30	68.25	58.90	63.23
Fine-tuning (BERT)	50.85	16.66	25.10	59.87	55.19	57.43	<b>81.60</b>	58.92	68.43
Relational Siamese Network (BERT)	39.07	<b>51.39</b>	44.47	42.42	54.93	47.87	44.10	52.73	48.03
Distant Supervision (BERT)	38.06	51.18	43.66	38.45	<b>76.12</b>	51.09	35.48	<b>80.33</b>	49.22
<b>Neural Snowball (BERT)</b>	<b>56.87</b>	40.43	<b>47.26</b>	<b>60.50</b>	62.20	<b>61.34</b>	78.13	66.87	<b>72.06</b>

# | Case Study



- By using RSN, we have achieved **high precision**
- However, neural snowball is stuck in the **comfort zone**

| Q & A