Dialogue Relation Extraction

Renato Vukovic

Dialog Systems and Machine Learning, HHU

2022-11-24

Relation Extraction (RE)

- Part of natural language understanding (NLU)
- Goal is to find the *semantic relation type* between entities in text such as professionally written and edited news reports.

Dialogue Relation Extraction

In **dialogue relation extraction** (DRE) the aim is to find relations in human conversations that are mostly not supported by any single utterance

Relation Extraction (RE)

- Part of natural language understanding (NLU)
- Goal is to find the *semantic relation type* between entities in text such as professionally written and edited news reports.

Dialogue Relation Extraction

In **dialogue relation extraction** (DRE) the aim is to find relations in human conversations that are mostly not supported by any single utterance
$\Longrightarrow$ Relations could add additional features and information for dialogue system tasks, e.g. for personalisation

Relation Extraction (RE)

Heinrich Heine Universität Düsseldorf

- Part of natural language understanding (NLU)
- Goal is to find the *semantic relation type* between entities in text such as professionally written and edited news reports.

Dialogue Relation Extraction

In **dialogue relation extraction** (DRE) the aim is to find relations in human conversations that are mostly not supported by any single utterance
$\Longrightarrow$ Relations could add additional features and information for dialogue system tasks, e.g. for personalisation
$\Longrightarrow$ Essential step in building ontologies from large-scale corpora automatically (Ji et al. 2010; Yu et al. 2020)

Dialogue Relation Extraction

- Dialogues have *lower information density* compared to formal literature (Wang and Yang Liu 2011).

Dialogue Relation Extraction

- Dialogues have *lower information density* compared to formal literature (Wang and Yang Liu 2011).
- Mostly cross-sentence relations (Chen et al. 2020)

Heinrich Heine Universität Düsseldorf

Dialogue Relation Extraction

- Dialogues have *lower information density* compared to formal literature (Wang and Yang Liu 2011).
- Mostly cross-sentence relations (Chen et al. 2020)
- Dialogues use simpler language and more pronouns than in written text

Dialogue Relation Extraction

- Dialogues have *lower information density* compared to formal literature (Wang and Yang Liu 2011).
- Mostly cross-sentence relations (Chen et al. 2020)
- Dialogues use simpler language and more pronouns than in written text
- Not made for an external reader

Dialogue Relation Extraction

- Dialogues have *lower information density* compared to formal literature (Wang and Yang Liu 2011).
- Mostly cross-sentence relations (Chen et al. 2020)
- Dialogues use simpler language and more pronouns than in written text
- Not made for an external reader
- Important to understand relations in dialogues in real-time with incoming utterances

Dialogue Relation Extraction

Heinrich Heine Universität Düsseldorf

- Dialogues have *lower information density* compared to formal literature (Wang and Yang Liu 2011).
- Mostly cross-sentence relations (Chen et al. 2020)
- Dialogues use simpler language and more pronouns than in written text
- Not made for an external reader
- Important to understand relations in dialogues in real-time with incoming utterances
  $\Longrightarrow$ Huge differences to relation extraction in other formal textual data!

Dialogue Relation Extraction

- Given:
  - Dialogue context $\mathcal{D}$ composed of utterances $\mathcal{D} = \{u_1 : s_1, ..., u_m : s_m\}$ where u_i, s_i denote the speaker ID and m is the dialogue length
  - Query pair q containing a subject entity and an object entity $q = (s, o)$ present in $\mathcal{D}$
- Goal:
  - Learn function f that finds the most possible relations between the given entities from a predefined relation set $\mathcal{R}$, $f(\mathcal{D}, q) \rightarrow \mathcal{R}$.
    $\longrightarrow$ Result is a *(subject, relation type, object)* triple (s, r, o)

Dialogue Relation Extraction

- Given:
  - Dialogue context $\mathcal{D}$ composed of utterances $\mathcal{D} = \{u_1 : s_1, ..., u_m : s_m\}$ where u_i, s_i denote the speaker ID and m is the dialogue length
  - Query pair q containing a subject entity and an object entity $q = (s, o)$ present in $\mathcal{D}$

- Goal:
  - Learn function f that finds the most possible relations between the given entities from a predefined relation set $\mathcal{R}$, $f(\mathcal{D}, q) \to \mathcal{R}$.
    $\longrightarrow$ Result is a *(subject, relation type, object)* triple (s,r,o)

- Closely connected to other dialogue tasks: Emotion recognition can be seen as a relation extraction task (Lee and Choi 2021)
  $\rightarrow$ the speaker of an utterance as subject, the utterance as object and the emotion as the relation between the arguments

DRE Data-sets

Heinrich Heine Universität Düsseldorf

DialogRE (Yu et al. 2020)

- First human-annotated dialogue-based relation extraction dataset
- 1,788 dialogues originating from the complete transcripts of the famous American television situation comedy "Friends".
- 36 possible relation types that exist between an argument pair in a dialogue.
- 10,168 relational triples are annotated
- 65.9% of relational triples involve arguments that never appear in the same turn
- Triggers are labelled here: useful hints for deciding on relations

DialogRE Relations

ID	Subject	Relation Type	Object	Inverse Relation	TR (%)
1	PER	per:positive_impression	NAME		70.4
2	PER	per:negative_impression	NAME		60.9
3	PER	per:acquaintance	NAME	per:acquaintance	22.2
4	PER	per:alumni	NAME	per:alumni	72.5
5	PER	per:boss	NAME	per:subordinate	58.1
6	PER	per:subordinate	NAME	per:boss	58.1
7	PER	per:client	NAME		50.0
8	PER	per:dates	NAME	per:dates	72.5
9	PER	per:friends	NAME	per:friends	94.7
10	PER	per:girl/boyfriend	NAME	per:girl/boyfriend	86.1
11	PER	per:neighbor	NAME	per:neighbor	71.2
12	PER	per:roommate	NAME	per:roommate	89.9
13	PER	per:children*	NAME	per:parents	85.4

Excerpt from the relations in DialogRE (Yu et al. 2020). TR denotes the percentage of relation triples of this type accompanied by an annotated trigger.

DialogRE Example

S2: He didn't have a last name. It was just "Tag". You know, like Cher, or, you know, Moses.

S3: But it was a deep meaningful relationship.

S2: Oh, you know what - my first impression of you was absolutely right. You are arrogant, you are pompous ...

Arguments	Trigger	Relation
(Tag, S2)	a deep meaningful relationship	per:girl/boyfriend
(S2, S3)	arrogant	per:negative_impression

DialogRE Triggers

- A trigger is ”the smallest extent (i.e., span) of contiguous text in the dialogue that most clearly indicates the existence of the relation between two arguments“.
- If there are multiple possible triggers, only one is kept for a relation triple.
- 49.6% of all relational triples are annotated with triggers

S1:	Hey Pheebs.
S2:	Hey!
S1:	Any sign of your **brother**?
S2:	No, but he's always late.
S1:	I thought you only met him once?
S2:	Yeah, I did. I think it sounds y'know big sistery, y'know, 'Frank's always late.'
S1:	Well relax, he'll be here.

	Argument pair	**Trigger**	**Relation type**
R1	(Frank, S2)	brother	per:siblings
R2	(S2, Frank)	brother	per:siblings
R3	(S2, Pheebs)	none	per:alternate_names
R4	(S1, Pheebs)	none	unanswerable

DialogRE Evaluation

- **Standard setting:**
  - Entire dialogue $\mathcal{D}$ as document d
  - Predict relation r based on d, s, o as input
  - F1-score as evaluation metric

DialogRE Evaluation

- **Standard setting:**
  - Entire dialogue $\mathcal{D}$ as document d
  - Predict relation r based on d, s, o as input
  - F1-score as evaluation metric
- **Conversational setting:**
  - Only the first $i \leq m$ turns are considered as document d
  - given the first i turns in a dialogue, relation type r associated with s and o is evaluable if s, o and the trigger for r have all been mentioned in the turns so far.
  - The converstational F1-score is denoted by $F1_c$ based on the conversational recall and precision, R_c and P_c

DRE Data-sets

Heinrich Heine Universität Düsseldorf

DDRel (Jia et al. 2021)

- Data-set for interpersonal relation classification in dyadic dialogues (i.e. between two persons/speakers)
  $\Longrightarrow$ Based on one or more dialogues predict the relation between the speakers
- Crawled movie scripts from IMSDb
- 6,300 dyadic dialogue sessions between 694 pairs of speakers with 53,126 utterances in total based on movie scripts.
- Up to 13 pre-defined relationships based on ”Encyclopedia of Human Relationships“ (Harry Reis 2009).
- Provide dialogue systems with features for personalisation

Interpersonal Relation Classification

- Input: one or several dialogues between two speakers called session and pair-level task
  - Session-level: Given the j-th dialogue between the i-th pair of speakers D_j^i infer the most probable relation between them: $R_j^i = \underset{R}{\text{argmax}}\ f_s(D_j^i)$
  - Pair-level: Given all dialogues between the i-th pair of speakers $D^i = (D_1, ..., D_n)$ infer the most probable relation between them: $R^i = \underset{R}{\text{argmax}}\ f_s(D^i)$

DDRel Relation Classes

4 classes	6 classes	13 classes
Family	Elder-Junior	Child-Parent
		Child-Other Family Elder
	Family Peer	Siblings
		Spouse
Intimacy	Intimacy	Lovers
		Courtship
Others	Peer	Friends
		Neighbors
		Roommates
Official	Elder-Junior	Workplace Superior-Subordinate
		Colleague/Partners
	Official Peer	Opponents
		Professional Contact

More and less fine-grained relation classes in DDRel data-set (Jia et al. 2021).

DialogRE vs. DDRel relations

DDRel Relation	DialogRE Relation
Workplace Superior-Subordinate	per:boss
Workplace Superior-Subordinate	per:subordinate
Friends	per:friends
Lovers	per:girl/boyfriend
Neighbors	per:neighbor
Roommates	per:roommate
Child-Parent	per:children
Child-Other Family Elder	per:other family
Siblings	per:siblings
Spouse	per:spouse
Colleague/Partners	per:works
Courtship	-
Opponents	-
Professional Contact	-

Overlap between DDRel and DialogRE relations (Lin et al. 2022)

DDRel Example

Heinrich Heine Universität Düsseldorf

Session 1

A Well, hi!

B Uh, we just wanted to stop by and say that we really **enjoyed your sets**.

A Oh, yeah, really, oh!

B I though it was ... very musical, and I liked it a lot.

A Oh, neat ... oh, that's very nice, gosh, thanks a lot.

B Are you ... are you recording? Or do-

A No, no, no, not at all.

B Uh, well, I'd like to talk to you about that sometime, if you get a chance.

A Oh. What about?

......

Session 2

B Oh, well, I-if it's inconvenient, eh, we can't do it now ... that's fine, too. W-w-we'll do it another time.

A Hey

B Maybe if you're on the Coast, we'll get together and ... and we'll meet there.

A Oh.

B It was a wonderful set.

A Oh, gosh.

B I really enjoyed it.

Session 3

B We just need about six weeks, in about six weeks we could **cut a whole album**.

A I don't know, this is strange to me, you know.

......

Session 4

B Boy, this is really a nice **screening room**. It's really a nice room.

A Oh, and there's another thing about New York. See ... you-you wanna see a movie, you have to stand in a long line.

A Yeah.

B It could be freezing, it could be raining.

A Yeah.

B And here, you just

Dialogue sessions with label ”Professional Contact“. Possible classification cues highlighted in red.

DDRel Baseline Results

		4-class		6-class		13-class	
		Acc	F1-macro	Acc	F1-macro	Acc	F1-macro
Session-level	Random	23.0±3.56	22.67±3.71	17.33±2.62	15.80±3.00	8.33±2.62	6.63±2.12
	Majority	31.00±0.00	11.80±0.00	31.00±0.00	7.90±0.00	26.00±0.00	3.20±0.00
	LSTM	29.80±1.28	22.87±1.24	30.83±1.16	11.10±0.08	28.50±1.44	4.63±0.45
	CNN	42.67±2.93	33.27±6.63	37.80±1.31	31.40±6.67	32.33±2.46	9.20±4.97
	BERT	47.10±1.28	44.53±1.10	41.87±0.81	39.40±0.85	39.40±0.36	20.40±0.67
	Human	56.00±6.00	55.20±6.30	50.00±9.00	53.00±8.10	38.50±5.50	40.75±8.15
Pair-level	Random	28.20±9.30	26.90±9.24	17.93±7.89	16.2±7.54	6.43±2.76	5.73±2.64
	Majority	23.10±0.00	9.40±0.00	23.10±0.00	6.20±0.00	19.20±0.00	2.50±0.00
	LSTM	25.63±2.76	13.13±5.06	22.67±0.61	6.40±0.29	19.20±0.00	2.57±0.05
	CNN	47.47±2.76	35.03±5.80	38.47±4.21	30.40±9.06	22.20±6.08	7.07±6.04
	BERT	58.13±0.61	52.00±0.86	42.33±2.76	38.00±1.14	39.73±1.79	24.07±0.63
	Human	75.65±3.85	73.00±4.40	72.40±4.50	73.55±5.45	63.45±1.95	54.40±3.00

classification results on session-level and pair-level tasks with a gap to human performance

Heinrich Heine Universität Düsseldorf

DRE Approaches

- Graph-based
  - SocAoG (Qiu et al. 2021)
- BERT-based
  - D-REX (Albalak et al. 2022)
  - TREND (Lin et al. 2022)

SocAoG: Incremental Graph Parsing for Social Relation Inference in Dialogues

Liang Qiu[1], Yuan Liang[2], Yizhou Zhao[1], Pan Lu[1], Baolin Peng[3], Zhou Yu[4], Ying Nian Wu[1], Song-Chun Zhu[1]

[1]UCLA Center for Vision, Cognition, Learning, and Autonomy
[2]University of California, Los Angeles
[3]Microsoft Research, Redmond
[4]University of California, Davis

Heinrich Heine Universität Düsseldorf

SocAoG

- Model social relation as an attributed And-Or graph (Wu and Zhu 2011)
  $\Rightarrow$ Social And-Or graph (SocAoG) based on a dialogue
- Incremental graph parsing algorithm to jointly infer human attributes and social relations from a dialogue
  $\Rightarrow$ enable dynamic updates of the relational belief based on incoming dialogue utterance

SocAoG

- Apply *Markov Chain Monte Carlo* (MCMC) to sample from the posterior probability calculated by three complementary processes called α, β and γ
  $\Rightarrow$ Incrementally parse SocAoG to get the final set of relations
- Learn the SocAoG model with a *contrastive loss* (Hadsell et al. 2006) comparing the posterior of a positive parse graph against a negative one according to relation annotations
- Infer social relations with dialogue turns as input

Heinrich Heine Universität Düsseldorf

SocAoG

The social and-or graph is defined as a 5-tuple:
$\mathcal{G} = \langle S, V, E, X, P \rangle$ where

SocAoG

Heinrich Heine Universität Düsseldorf

The social and-or graph is defined as a 5-tuple:
$\mathcal{G} = \langle S, V, E, X, P \rangle$ where

- S is the root node for representing a society

SocAoG

The social and-or graph is defined as a 5-tuple:
$\mathcal{G} = \langle S, V, E, X, P \rangle$ where

- S is the root node for representing a society
- $V = V_{and} \cup V_{or} \cup V_T^e \cup V_T^a$ are all nodes, where V_{and} the And-node set, V_{or} the Or-node set, while V_T^e and V_T^a represent human members and attribute values respectively

SocAoG

The social and-or graph is defined as a 5-tuple:
$\mathcal{G} = \langle S, V, E, X, P \rangle$ where

- S is the root node for representing a society
- $V = V_{and} \cup V_{or} \cup V_T^e \cup V_T^a$ are all nodes, where V_{and} the And-node set, V_{or} the Or-node set, while V_T^e and V_T^a represent human members and attribute values respectively
- E is the set of edges describing social relations
- $X(v_i)$ are the attributes associated with node v_i and $X(\vec{e}_{ij})$ the

SocAoG

The social and-or graph is defined as a 5-tuple:
$\mathcal{G} = \langle S, V, E, X, P \rangle$ where

- S is the root node for representing a society
- $V = V_{and} \cup V_{or} \cup V_T^e \cup V_T^a$ are all nodes, where V_{and} the And-node set, V_{or} the Or-node set, while V_T^e and V_T^a represent human members and attribute values respectively
- E is the set of edges describing social relations
- $X(v_i)$ are the attributes associated with node v_i and $X(\vec{e}_{ij})$ the social relation type of edge $\vec{e}_{ij} \in E$ (For simplicity here, $X(v_i)$ is denoted as $\mathbf{v}_i$ and $X(\vec{e}_{ij})$as $\mathbf{e}_i j$ from now on)

SocAoG

The social and-or graph is defined as a 5-tuple:
$\mathcal{G} = \langle S, V, E, X, P \rangle$ where

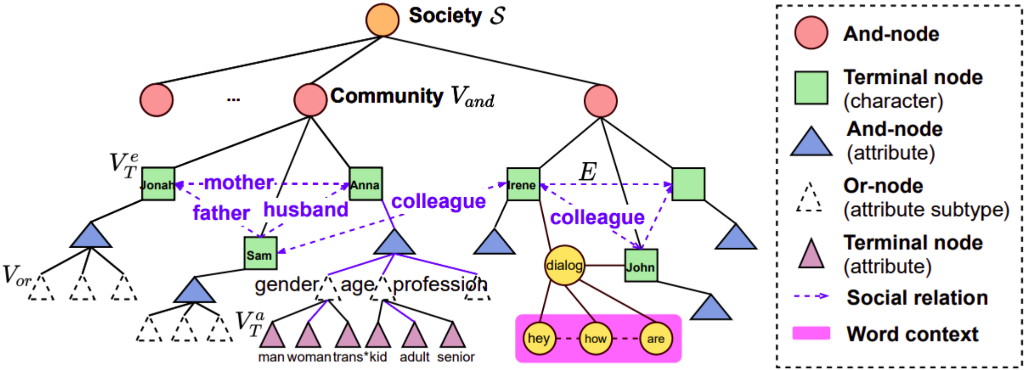
- S is the root node for representing a society
- $V = V_{and} \cup V_{or} \cup V_T^e \cup V_T^a$ are all nodes, where V_{and} the And-node set, V_{or} the Or-node set, while V_T^e and V_T^a represent human members and attribute values respectively
- E is the set of edges describing social relations
- $X(v_i)$ are the attributes associated with node v_i and $X(\vec{e}_{ij})$ the social relation type of edge $\vec{e}_{ij} \in E$ (For simplicity here, $X(v_i)$ is denoted as $\mathbf{v}_i$ and $X(\vec{e}_{ij})$as $\mathbf{e}_i j$ from now on)
- P is the probability model defined on SocAoG
  $\Rightarrow$ The parse graph pg for SocAoG is updated incrementally over turns to get the optimal parse graph pg^* to be:

$$pg^* = \underset{pg}{\operatorname{argmax}}(p(pg \mid D; \theta))$$

with dialogue D and inferred model parameters θ

Heinrich Heine Universität Düsseldorf

SocAoG Parse Graph



$\alpha - \beta - \gamma$ for Graph Inference

Heinrich Heine Universität Düsseldorf

- posterior probability for a parse graph *pg*:
$$p(pg \mid D; \theta) = \frac{1}{Z} \exp\{-\mathcal{E}(D \mid D; \theta) - \mathcal{E}(pg; theta)\}$$
where Z is the partition function and $\mathcal{E}(D \mid D; \theta)$ and $\mathcal{E}(pg; theta)$ are dialogue- and social norm-based energy potentials, measuring the cost of assigning graph instantiation
- For a dialogue as a sequence of words $D = \{w_1, ..., w_T\}$ the dialogue likelihood energy is given by:
$$\mathcal{E}(D \mid pg; \theta) = \sum_{t=1}^{T} - \log(p(w_t \mid c_t, pg))$$
where $c_t = [w_1, ..., w_{t-t}]$ is the context vector by a BERT model that gets the dialogue history and the current parse graph belief as input.
$p(w_t \mid c_t, pg)$ is the α process

$\alpha - \beta - \gamma$ for Graph Inference

- The social norm-based potential is composed of three potential terms:

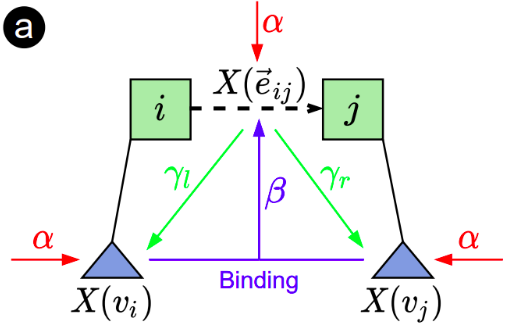
$$\mathcal{E}(pg;\theta) = -\beta \sum_{v_i, v_j \in V(pg)} log(p(\mathbf{e}_{ij} \mid \mathbf{v}_i, \mathbf{v}_j)) \tag{1}$$

$$-\gamma_l \sum_{\vec{e}_{ij} \in E(pg)} log(p(\mathbf{v}_i \mid \mathbf{e}_{ij})) \tag{2}$$

$$-\gamma_r \sum_{\vec{e}_{ij} \in E(pg)} log(p(\mathbf{v}_j \mid \mathbf{e}_{ij})) \tag{3}$$

where
$p(\mathbf{e}_{ij} \mid \mathbf{v}_i, \mathbf{v}_j)$ is called the β process where the relation edge is updated based on the node attributes
$p(\mathbf{v}_i \mid \mathbf{e}_{ij})$ and $p(\mathbf{v}_j \mid \mathbf{e}_{ij})$ the γ process in which the social relation edge is used to update the node attributes

$\alpha - \beta - \gamma$ for Graph Inference



SocAoG Inference

Heinrich Heine Universität Düsseldorf

Algorithm 1: Incremental SocAoG Parsing for Social Relation Inference

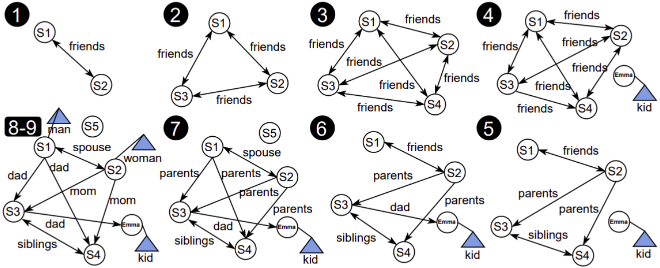
Input: dialogue $D_T = \{D^{(1)}, D^{(2)}, ..., D^{(T)}\}$, target argument pairs $\{a_1, a_2\}$.
Initialize $pg^{(0)}$. Initialize $\mathbf{v}_i$ and $\mathbf{e}_{ij}$.
for $t = 1, ..., T$ **do**
  for $s = 1, ..., S$ **do**
    Compute the posterior $p(pg|D^{(t)}; \theta)$.
    Make proposal moves with probabilities q_1, q_2 to get a new parse graph pg'.
    Compute the posterior $p(pg'|D^{(t)}; \theta)$.
    Compute acceptance rate $\alpha(pg'|pg, D^{(t)}; \theta)$.
    Accept/reject pg' according to the acceptance rate.
  end for
  return $\mathbf{e}_{a_1,a_2}$ from the average of accepted pg samples.
end for

- Acceptance rate α for updated parse graph pg':
  $\alpha(pg' \mid pg, D; \theta) = min(1, \frac{p(pg' | D; \theta)}{p(pg | D; \theta)})$
- $S = min\{w \times (KM + K(K-1)N, S_{max}\}$ for K entities, M attributes, N relations, and a sweep number of w

Heinrich Heine Universität Düsseldorf

SocAoG Inference Example

1	**S1, S2:**	Hi!
2	**S3:**	Hey!
3	**S4:**	So glad you came!
4	**S1:**	I can't believe Emma is already one!
5	**S2:**	I remember your first birthday! Ross was jealous of all the attention we were giving you. He pulled on his testicles so hard! We had to take him to the emergency room!
6	**S3:**	There's something you didn't know about your dad!
7	**S5:**	Hey Mr. and Mrs. Geller! Let me help you with that.
8	**S1:**	Thank you!
9	**S5:**	Oh man, this is great, uh? The three of us together again! You know what would be fun? If we gave this present to Emma from all of us!



SocAoG Results on DialogRE

Methods	DialogRE (V2) Dev $F1(\sigma)$	Dev $F1_c(\sigma)$	Test $F1(\sigma)$	Test $F1_c(\sigma)$
BERT (Devlin et al., 2018)	59.4 (0.7)	54.7 (0.8)	57.9 (1.0)	53.1 (0.7)
$BERT_S$ (Yu et al., 2020)	62.2 (1.3)	57.0 (1.0)	59.5 (2.1)	54.2 (1.4)
GDPNet (Xue et al., 2020b)	67.1 (1.0)	61.5 (0.8)	64.3 (1.1)	60.1 (0.9)
SimpleRE (Xue et al., 2020a)	68.2 (1.1)	63.4 (0.6)	66.7 (0.7)	63.3 (0.9)
$SocAoG_{reduced}$ (our method)	69.1 (0.4)	65.7 (0.5)	68.6 (0.9)	65.4 (1.1)
SocAoG (our method)	**69.5 (0.8)**	**66.1 (0.7)**	**69.1 (0.5)**	**66.5 (0.8)**

SocAoG Results on DialogRE

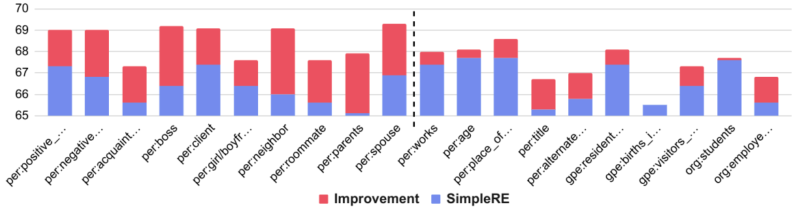


Figure: Improvement per relation type of SocAoG compared to BERT-based SimpleRE (Xue, Sun, Zhang, Ni, et al. 2020) where several [CLS] tokens from BERT are used to capture relations between multiple entity pairs.

D-REX: Dialogue Relation Extraction with Explanations

Alon Albalak[1], Varun Embar[2], Yi-lin Tuan[1], Lise Getoor[2], William Yang Wang[1]

[1] University of California, Santa Barbara
[2] University of California, Santa Cruz

Heinrich Heine Universität Düsseldorf

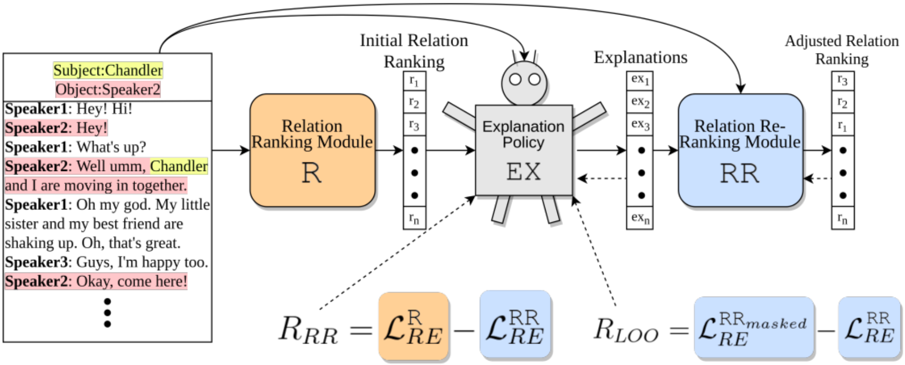
D-REX

- Combine dialogue relation extraction with *explanation extraction* (EE)
- Triggers as partial supervision signal for EE
- Multiple reward functions to explore the explanation space with policy gradient
  $\rightarrow$ learn meaningful explanations on data with less than 40% supervised triggers
- DRE as a ranking task with EE as intermediate step

Explanation Extraction

- **Given:**
  - Dialogue d consisting of n tokens $T_1, ..., T_n$
  - Relational triple $\langle s, r, o \rangle$
- **Goal:** predict span with start and end positions i,j in the dialogue, such that explanation $ex = [T_i, ..., T_j]$ indicates that r holds between s and o.

D-REX Architecture



D-REX Ranking Modules

- R denotes the ranking module, which is a BERT (Devlin et al. 2019) or RoBERTa (Yinhan Liu et al. 2019) language model fine-tuned on relation extraction as a classification task
  - Input is the Dialogue d, the subject s and the object o, output is the relation $R(s,o,d)$
  - Is trained separately before being put into D-REX and not updated anymore
- RR denotes the re-ranking module with the same model architecture as R
  - Gets the explanation as additional input $\rightarrow RR(ex, s, o, d)$
  - updated with cross-entropy loss, condition its ranking on explanations from the explanation module EX

D-REX Explanation Module

- Input is (s,o,d) plus $R(s,o,d)$
- Output is an extracted phrase from the dialogue d, denoted as $EX(r,s,o,d)$
- First train on supervised triggers with cross-entropy loss, then with policy gradient to include unlabelled examples
- Predict explanations for the top-k ranked relations by R.

DialogRE V2

Dialogues	Relations	Relational Triples (train/dev/test)	Triggers (train/dev/test)
1788	36	6290/1992/1921	2446/830/780

Number of triggers vs. number of relational triples. Only supervised training on annotated triggers

D-REX Explanation Policy

- Action space of *EX* is the set of spans in the dialogue
  $\rightarrow$ predict start and end token of explanation and receive feedback from environment with two reward functions $\mathcal{R}_{RR}$ and $\mathcal{R}_{LOO}$
- The loss is then calculated as:
  $\mathcal{L}_{EX_{PG}} = -(\log(P^S_{T_i}) + \log(P^E_{T_j})) \cdot (\mathcal{R}_{RR} + \mathcal{R}_{LOO})$
  where $P^S_{T_i}$ and $P^E_{T_j}$ denote the probability that a token is a start or end token of an explanation respectively

D-REX Explanation Rewards

- **Re-ranking reward:** ensure that *EX* predicts explanations that benefit *RR* by subtracting the RE loss from *RR* from the *R* loss:

$$\mathcal{R}_{RR} = \mathcal{L}^{R}_{RE}(s,o,d) - \mathcal{L}^{RR}_{RE}(ex,s,o,d)$$

Because R is stationary, *EX* minimises $\mathcal{L}^{RR}_{RE}$ by improving explanations *EX*.

D-REX Explanation Rewards

- **Re-ranking reward:** ensure that *EX* predicts explanations that benefit *RR* by subtracting the RE loss from *RR* from the *R* loss:

  $$\mathcal{R}_{RR} = \mathcal{L}^{R}_{RE}(s,o,d) - \mathcal{L}^{RR}_{RE}(ex,s,o,d)$$

  Because R is stationary, *EX* minimises $\mathcal{L}^{RR}_{RE}$ by improving explanations *EX*.
- **Leave-one-out Reward:** direct EX in finding phrases which are essential to correctly classifying the relation between an argument-pair:

  $$\mathcal{R}_{LOO} = \mathcal{L}_{RE}(s,o,d_{mask}(ex)) - \mathcal{L}_{RE}(s,o,d)$$

  where $d_{mask}(ex)$ is the dialogue d with the predicted explanation ex masked.
  $\rightarrow$ The model needs to maximise the masked loss, such that the explanation contains everything important for relation extraction

D-REX Prediction Example

Speaker 1: Could you please get the key off the back of the door for me.

Speaker 2: Oh yeah! Yeah!

Speaker 1: You tell your friend Chandler that we're definitely broken up this time.

Speaker 2: Okay!

Subject	Object	Initial Predicted Relation	D-REX Predicted Explanation	D-REX Predicted Relation
Speaker 2	Chandler	girl/boyfriend	your friend	friends

D-REX Results on DialogRE

Model	F1(σ)
R_{BERT}	59.2(1.9)
$Joint_{\text{BERT}}$	59.4(1.7)
D-REX$_{\text{BERT}}$	**59.9**(0.5)
R_{RoBERTa}	64.2(1.6)
$Joint_{\text{RoBERta}}$	65.2(0.3)
D-REX$_{\text{RoBERTa}}$	**67.2**(0.3)
*GDPNet	60.2(1.0)
*TUCORE-GCN$_{\text{BERT}}$	65.5(0.4)
†SocAoG	**69.1**(0.5)

$F1_c$ of D-REX compared to other models

D-REX EE Evaluation

D-REX$_{\text{RoBERTa}}$ vs.	Win(%)	Tie(%)	Lose(%)
Random (*NL*)	79.9	10.4	9.8
Joint$_{\text{RoBERTa}}$ (*NL*)	38.5	52.3	9.2
Ground truth (*DL*)	12.1	44.3	43.7

D-REX human preference on examples with no labelled trigger (NL) and where explanations differ from the label (DL).

TREND: Trigger-Enhanced Relation Extraction Network for Dialogues

Po-Wei Lin, Shang-Yu Su, Yun-Nung Chen

National Taiwan University, Taipei, Taiwan

TREND

Heinrich Heine Universität Düsseldorf

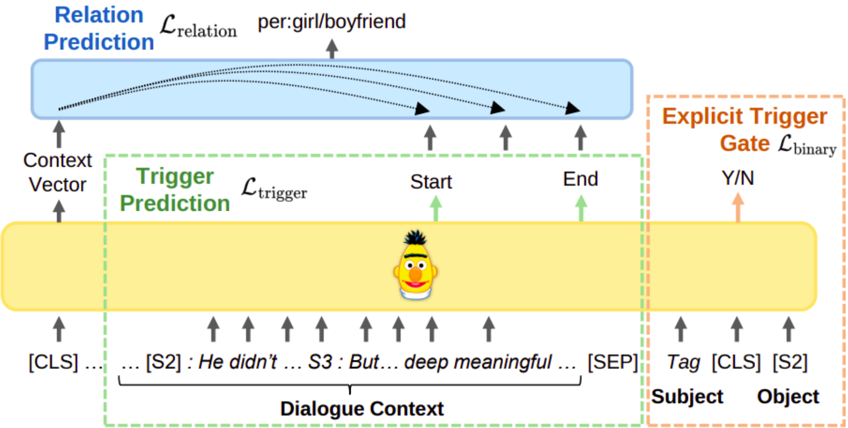
- **Core idea:** identify trigger spans to be used for relation extraction improvement
- Multitasking model with attentional relation extractor
- General capability of finding triggers
  $\Longrightarrow$ Transferability to unseen relations

TREND

Two modules:

1. Multi-tasking BERT forcontext encoding and explicit trigger identification
2. Relation predictor with a feature combination of the dialogue and the automatically identified trigger

The TREND Model



TREND Modules

Heinrich Heine Universität Düsseldorf

- **Explict trigger gate:** A binary classifier as a gate to identify if the explicit triggers exist
- **Trigger prediction:** Fine-tune a language model on predicting the start and end token of explicit triggers in the dialogue
- **Relation prediction:** Feed the context vector as query and trigger words as keys and values into an attention mechanism with a classification head on top for relation classification

TREND Modules

- **Explict trigger gate:** A binary classifier as a gate to identify if the explicit triggers exist
- **Trigger prediction:** Fine-tune a language model on predicting the start and end token of explicit triggers in the dialogue
- **Relation prediction:** Feed the context vector as query and trigger words as keys and values into an attention mechanism with a classification head on top for relation classification

  $\Rightarrow$ Train these models jointly on DialogRE Data-set

TREND Modules

- **Explicit trigger gate:** A binary classifier as a gate to identify if the explicit triggers exist
- **Trigger prediction:** Fine-tune a language model on predicting the start and end token of explicit triggers in the dialogue
- **Relation prediction:** Feed the context vector as query and trigger words as keys and values into an attention mechanism with a classification head on top for relation classification
  $\Rightarrow$ Train these models jointly on DialogRE Data-set
  $\Rightarrow$ transfer the trigger-finding capability to DDRel data-set where the model trained on DialogRE is fine-tuned on relation extraction without triggers

TREND Results on DialogRE

Model	**F1**
BERT	60.6
GDPNet	64.3
SimpleRE (single entity pair)	60.4
D-REX$_{\text{BERT}}$	59.2
TUCORE-GCN$_{\text{BERT}}$	65.5
TREND$_{\text{BERT-Base}}$	**66.8**
TREND$_{\text{BERT-Large}}$	**67.8**
SimpleRE (multiple entity pairs)	66.7
SocAoG (multiple entity pairs)	*69.1*
TREND$_{\text{BERT-Base}}$ (ground-truth triggers)	75.3

Conversational $F1_c$ of TREND and other models.
The trigger prediction has no more than 50% exact match which is why TREND with ground-truth triggers performs better

TREND Results on DDRel

Model	4-class		6-class		13-class	
	Acc	Macro-F	Acc	Macro-F	Acc	Macro-F
BERT	47.1 / 58.1	44.5 / 52.0	41.9 / 42.3	39.4 / 38.0	39.4 / 39.7	20.4 / 24.1
TUCORE-GCN$_{\text{BERT}}$	43.8 / 60.3	41.9 / 56.6	36.9 / **52.6**	38.7 / 54.2	29.5 / 44.9	20.5 / **36.9**
TREND$_{\text{BERT-Base}}$	51.5 / **65.4**	**46.5 / 61.2**	40.3 / **52.6**	**43.0 / 55.0**	**40.5 / 46.2**	21.2 / 34.7
w/o binary gate	52.5 / 53.8	45.3 / 49.7	37.0 / 43.6	41.8 / 45.9	36.6 / 43.6	**26.4** / 36.3
TREND$_{\text{BERT-Large}}$	**51.6** / 60.3	**46.5** / 54.0	**42.5** / 46.2	**43.0** / 48.2	34.4 / 43.6	19.9 / 36.3
w/o binary gate	41.5 / 47.4	40.3 / 44.9	39.0 / 42.3	43.1 / 42.9	38.5 / 34.6	17.3 / 21.1

DDRel performance in session-level/pair-level settings and different granularity settings

TREND Unseen DDRel Relations

DDRel Relation	DialogRE Relation
Workplace Superior-Subordinate	per:boss
Workplace Superior-Subordinate	per:subordinate
Friends	per:friends
Lovers	per:girl/boyfriend
Neighbors	per:neighbor
Roommates	per:roommate
Child-Parent	per:children
Child-Other Family Elder	per:other family
Siblings	per:siblings
Spouse	per:spouse
Colleague/Partners	per:works
Courtship	-
Opponents	-
Professional Contact	-

Overlap between DDRel and DialogRE relations

DDRel Relation	Seen	Unseen
BERT	23.77	9.94
TUCORE-GCN	23.39	10.81
TREND	28.30	13.13

S1: **Fuck** me!
S2: Want a drink? Okay... I'm not good at this sort of thing, but we don't have a lot of time, so I'll just go ahead and get started.

Argument	Relation (Unseen)	Trigger
(S1, S2)	Opponent	fuck

Performance on DDRel relations seen and not seen on DialogRE and predicted trigger and relation on unseen DDRel relation

Conclusion

- Dialogue relation extraction is important for finding relevant structures for dialogue systems

Heinrich Heine Universität Düsseldorf

Conclusion

- Dialogue relation extraction is important for finding relevant structures for dialogue systems
- Capturing relations between people is important for personalising dialogue systems and adjusting language to the user

Conclusion

- Dialogue relation extraction is important for finding relevant structures for dialogue systems
- Capturing relations between people is important for personalising dialogue systems and adjusting language to the user
- Graph-based and language model based approaches perform reasonably well on DRE

Conclusion

- Dialogue relation extraction is important for finding relevant structures for dialogue systems
- Capturing relations between people is important for personalising dialogue systems and adjusting language to the user
- Graph-based and language model based approaches perform reasonably well on DRE
- It is possible to model dialogue relations with a graph

Conclusion

- Dialogue relation extraction is important for finding relevant structures for dialogue systems
- Capturing relations between people is important for personalising dialogue systems and adjusting language to the user
- Graph-based and language model based approaches perform reasonably well on DRE
- It is possible to model dialogue relations with a graph
- Adding an intermediate explanation step improves performance and explainability

Heinrich Heine Universität Düsseldorf

Conclusion

- Dialogue relation extraction is important for finding relevant structures for dialogue systems
- Capturing relations between people is important for personalising dialogue systems and adjusting language to the user
- Graph-based and language model based approaches perform reasonably well on DRE
- It is possible to model dialogue relations with a graph
- Adding an intermediate explanation step improves performance and explainability
- Additional annotation like triggers can improve the performance on relation extraction with better generalisability

Further Reading

- Dialogue Relation Extraction with Document-Level Heterogeneous Graph Attention Networks (DHGAT) (Chen et al. 2020)
- An Embarrassingly Simple Model for Dialogue Relation Extraction (SimpleRE) (Xue, Sun, Zhang, Ni, et al. 2020)
- GDPNet: Refining Latent Multi-View Graph for Relation Extraction (Xue, Sun, Zhang, and Chng 2021)
- Graph Based Network with Contextualized Representations of Turns in Dialogue (TUCORE-GCN) (Lee and Choi 2021)

Thank you for your Attention!

Any questions?

References I

Albalak, Alon et al. (May 2022). "D-REX: Dialogue Relation Extraction with Explanations". In: *Proceedings of the 4th Workshop on NLP for Conversational AI*. Dublin, Ireland: Association for Computational Linguistics, pp. 34–46. DOI: 10.18653/v1/2022.nlp4convai-1.4. URL: https://aclanthology.org/2022.nlp4convai-1.4.

Chen, Hui et al. (2020). "Dialogue relation extraction with document-level heterogeneous graph attention networks". In: *arXiv preprint arXiv:2009.05092*.

Devlin, Jacob et al. (June 2019). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Minneapolis, Minnesota: Association for Computational Linguistics, pp. 4171–4186. DOI: 10.18653/v1/N19-1423. URL: https://aclanthology.org/N19-1423.

References II

- Hadsell, R., S. Chopra, and Y. LeCun (2006). "Dimensionality Reduction by Learning an Invariant Mapping". In: *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*. Vol. 2, pp. 1735–1742. DOI: 10.1109/CVPR.2006.100.
- Harry Reis, Susan Sprecher (2009). *Encyclopedia of Human Relationships*. Sage Publications. DOI: 10.4135/9781412958479. URL: https://sk.sagepub.com/reference/humanrelationships.
- Ji, Heng et al. (2010). "Overview of the TAC 2010 Knowledge Base Population Track". In.
- Jia, Qi, Hongru Huang, and Kenny Q. Zhu (May 2021). "DDRel: A New Dataset for Interpersonal Relation Classification in Dyadic Dialogues". In: vol. 35. 14, pp. 13125–13133. DOI: 10.1609/aaai.v35i14.17551. URL: https://ojs.aaai.org/index.php/AAAI/article/view/17551.

References III

Lee, Bongseok and Yong Suk Choi (Nov. 2021). "Graph Based Network with Contextualized Representations of Turns in Dialogue". In: *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Online and Punta Cana, Dominican Republic: Association for Computational Linguistics, pp. 443–455. DOI: 10.18653/v1/2021.emnlp-main.36. URL: https://aclanthology.org/2021.emnlp-main.36.

Lin, Po-Wei, Shang-Yu Su, and Yun-Nung Chen (Sept. 2022). "TREND: Trigger-Enhanced Relation-Extraction Network for Dialogues". In: *Proceedings of the 23rd Annual Meeting of the Special Interest Group on Discourse and Dialogue*. Edinburgh, UK: Association for Computational Linguistics, pp. 623–629. URL: https://aclanthology.org/2022.sigdial-1.58.

Liu, Yinhan et al. (2019). "RoBERTa: A Robustly Optimized BERT Pretraining Approach". In: *CoRR* abs/1907.11692v1. arXiv: 1907.11692. URL: https://doi.org/10.48550/arXiv.1907.11692.

References IV

- Qiu, Liang et al. (Aug. 2021). "SocAoG: Incremental Graph Parsing for Social Relation Inference in Dialogues". In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Online: Association for Computational Linguistics, pp. 658–670. DOI: 10.18653/v1/2021.acl-long.54. URL: https://aclanthology.org/2021.acl-long.54.
- Wang, Dong and Yang Liu (June 2011). "A Pilot Study of Opinion Summarization in Conversations". In: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. Portland, Oregon, USA: Association for Computational Linguistics, pp. 331–339. URL: https://aclanthology.org/P11-1034.
- Wu, Tianfu and Song Zhu (June 2011). "A Numerical Study of the Bottom-Up and Top-Down Inference Processes in And-Or Graphs". In: vol. 93, pp. 226–252. DOI: 10.1007/s11263-010-0346-6.

References V

- Xue, Fuzhao, Aixin Sun, Hao Zhang, and Eng Siong Chng (May 2021). "GDPNet: Refining Latent Multi-View Graph for Relation Extraction". In: *Proceedings of the AAAI Conference on Artificial Intelligence* 35.16, pp. 14194–14202. DOI: 10.1609/aaai.v35i16.17670. URL: https://ojs.aaai.org/index.php/AAAI/article/view/17670.
- Xue, Fuzhao, Aixin Sun, Hao Zhang, Jinjie Ni, et al. (2020). *An Embarrassingly Simple Model for Dialogue Relation Extraction*. DOI: 10.48550/ARXIV.2012.13873. URL: https://arxiv.org/abs/2012.13873.
- Yu, Dian et al. (July 2020). "Dialogue-Based Relation Extraction". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics, pp. 4927–4940. DOI: 10.18653/v1/2020.acl-main.444. URL: https://aclanthology.org/2020.acl-main.444.