

# Characterizing Interactions and Relationships between People

Farzana Rashid and Eduardo Blanco

Human Intelligence and Language Technologies Lab

University of North Texas

Denton, TX, 76203

farzanarashid@my.unt.edu, eduardo.blanco@unt.edu

## Abstract

This paper presents a set of dimensions to characterize the association between two people. We distinguish between interactions (when somebody refers to somebody in a conversation) and relationships (a sequence of interactions). We work with dialogue scripts from the TV show *Friends*, and do not impose any restrictions on the interactions and relationships. We introduce and analyze a new corpus, and present experimental results showing that the task can be automated.

## 1 Introduction

People interact with each other and as a result form relationships. These relationships range from weak (e.g., John talking to a waiter to order a drink) to strong (e.g., John and his best friend discussing career options). Traditionally, information extraction systems target, among others, relationships between people, e.g., PARENT, SIBLING, OTHER-PERSONAL, OTHER-PROFESSIONAL (Doddington et al., 2004).

Extracting a label describing the general relationship between two entities—often called *relation type*—is useful for tasks such as summarization (Jijkoun et al., 2004) and question answering (White et al., 2001). Only assigning a relation type, however, does not account for nuances in the relationship between two individuals. First, a relationship can be characterized beyond a relation type. For example, people having a PROFESSIONAL relationship may be *spatially near* or *distant* (working at the same or different offices), and have an *equal* or *hierarchical* relationship (two software developers or a developer and the CEO). Second, relationships are defined by multiple interactions, and the fine-grained characteristics of interactions do not necessarily mirror the characteristics of the corresponding relationship.

For example, software developers having a *cooperative* PROFESSIONAL relationship may have a *heated* interaction in a meeting that does not affect the long-term PROFESSIONAL relationship. Similarly, the same software developers having a *task-oriented* PROFESSIONAL relationship may have occasionally *pleasure oriented* interactions (e.g., when they go out for drinks on Fridays).

This kind of fine-grained characteristics of interactions and relationships are called *dimensions* in social science (Wish et al., 1976). Social scientists have also studied language usage and how people interact with each other depending on their relationship. For example, Gibbs (2000) studies irony (sarcasm, hyperbole, understatement, etc.) in communications among friends, and Snyder and Stukas Jr (1999) analyze the expectations in social interactions (e.g., interactions between strangers tend to be more formal) as well as the consequences of breaking expectations. In the social sciences, however, researchers mostly focus on how people act (e.g., how they talk to each other and about each other) and how people perceive interactions and relationships. In general, they do not attempt to automatically characterize interactions or relationships from language usage.

In this paper, we characterize the interactions between people and the resulting relationships. The main contributions of this paper are: (a) a set of dimensions to characterize interactions and relationships, including dimensions previously defined in the social sciences and two novel dimensions; (b) annotations of these dimensions for all interactions and relationships in 24 episodes of the TV show *Friends* (Season 1); (c) corpus analysis including agreements, label distributions and correlations; and (d) experimental results showing that classifiers grounded on language usage (and discarding the names of people being considered) are successful at automating the task.

## 2 Previous Work

Within natural language processing, there have been several recent efforts working with relationships between people. Voskarides et al. (2015) extract human-readable descriptions of relations in a knowledge graph. Unlike the work presented here, they experiment with a proprietary knowledge graph and rely heavily on features extracted from the graph. Iyyer et al. (2016) propose an unsupervised algorithm to extract relationship trajectories of fictional characters. Bracewell et al. (2012) introduce social acts (e.g., agreement, undermining) designed to characterize relationships exhibiting adversarial and collegial behavior (similar to our cooperative vs. competitive dimension). None of these works distinguish between interactions and relationships, characterize interactions and relationships with dimensions, or consider all interactions between two people.

In our previous work, we characterize interpersonal relationships with dimensions (Rashid and Blanco, 2017). In this paper, we improve upon our previous effort as follows. First, we distinguish between interpersonal interactions and relationships. Second, we work with dialogues thus the same people interact with each other many times.

There have been a few studies on analyzing language usage when people communicate. For example, Danescu-Niculescu-Mizil et al. (2012) study how power differences affect language style in online communities, and Prabhakaran and Rambow (2014) present a classifier to detect power relationships in email threads. Similarly, Gilbert (2012) explores how people in hierarchical relationships communicate through email, and Bramsen et al. (2011) focus on identifying power relationships in social networks. Politeness in online forums has also been studied (Danescu-Niculescu-Mizil et al., 2013). While power (similar to our equal vs. hierarchical dimension, Section 3) and politeness could be considered dimensions, these works exploit structural and linguistic features derived from communications between two individuals. Unlike all of them, we distinguish between and characterize interactions and relationships, and automate the task using only information derived from language usage.

Information extraction systems target, among others, relationships between people. There have been many evaluations (Grishman and Sundheim, 1996; Doddington et al., 2004; Kulick et al., 2014;

Surdeanu and Heng, 2014), and there are two main approaches. Traditionally, relationships are defined before training takes place (e.g., PAR-ENT, FRIENDS), and systems are trained using supervised machine learning (Yu and Lam, 2010; Nguyen et al., 2016; West et al., 2014). On the other hand, open information extraction (Wu and Weld, 2010; Angeli et al., 2015) has emerged as an unsupervised domain-independent approach to extract relations. Regardless of details, these previous works extract explicit relationships and do not attempt to characterize instances of relationships with dimensions. Additionally, they do not distinguish between interactions and relationships.

## 3 Interpersonal Interactions and Relationships

In this paper, we work with transcripts of conversations and define interaction and relationship as follows. An *interaction* between two people  $x$  and  $y$  exists for each conversation turn by either  $x$  or  $y$  referring to the other person. A *relationship* between two people  $x$  and  $y$  exists if there is at least one interaction between them. One could understand a relationship between  $x$  and  $y$  as the association defined by a sequence of interactions between  $x$  and  $y$ . Beyond these definitions, we do not impose any restriction on what constitutes an interaction or relationship: interactions occur each time two people refer to each other in their speech (even if they are not talking to each other), and one or more interactions constitute a relationship.

Interactions and relationships between people have been extensively studied in psychology and social sciences in general. The *right* set of dimensions is not agreed upon (Wish et al., 1976; Deutsch, 2011; Adamopoulos, 2012), and we argue that it depends on the domain of interest (e.g., personal diaries vs. news articles covering politics). We note that dimensions apply to interactions and relationships between specific people (i.e., instances of interactions and relationships), not relation types. For example, a KINSHIP relationship between  $x$  and  $y$  could be *intense* or *superficial* (depending on  $x$  and  $y$ ) and a particular interaction of that relationship may be *spatially near* or *distant* (even for the same  $x$  and  $y$ ).

The dimensions we work with are briefly summarized in Table 1 and described below. All but two dimensions are defined in previous work in the social sciences (see references in Table 1).

		Other descriptors	Ref.
Interaction	Cooperative vs. Competitive	friendly vs. hostile, promotive vs. contrient	[1]
	Active vs. Passive	direct vs. indirect, unequivocal vs. equivocal	[2]
	Concurrent vs. Nonconcurrent	convergent vs. divergent, synchronous vs. asynchronous	New
	Spatially Near vs. Distant	nearby vs. faraway, attached vs. detached	New
Relationship	Equal vs. Hierarchical	autonomy vs. control, submission, dominance	[1]
	Intense vs. Superficial	important vs. insignificant, influential vs. trivial	[1]
	Pleasure vs. Task Oriented	emotionally involved vs. detached	[1]
	Intimate vs. Unintimate	close vs. distant, indifference, randomness	[3]
	Temporary vs. Enduring	momentary vs. lasting, provisional vs. permanent	[4]

Table 1: Dimensions of interpersonal interactions and relationships. [1] stands for (Wish et al., 1976), [2] for (Kelley, 2013), [3] for (Adamopoulos, 2012), and [4] for (Deutsch, 2011). *New* indicates a dimension defined after analyzing several examples and doing pilot annotations.

**Dimensions of Interactions.** We consider four dimension of interactions between people, i.e., when a speaker refers to a person in a conversation:

- *Cooperative vs. Competitive.* A interaction is *cooperative* if both people (a) have a common interest or goal, (b) like each other, (c) benefit from the interaction, or (d) think alike or share similar views. Otherwise, the interaction is *competitive*.
- *Active vs. Passive.* An interaction is *active* if both people communicate directly. Otherwise, the interaction is *passive*. For example, when John talks to Bill about Mary, John and Bill have an active interaction, and John and Mary have a passive interaction.
- *Concurrent vs. Nonconcurrent.* An interaction is *concurrent* if both people are involved in an event or action at the same time (the event does not need to be a communicating event). Otherwise, the interaction is *nonconcurrent*. For example, when John talks to Bill about a trip with Mary, John and Mary have a *concurrent* interaction, but when John talks to Bill about Paul’s house, John and Paul have a nonconcurrent interaction.
- *Spatially Near vs. Distant.* An interaction is *spatially near* (or *near* for short) if the interaction is concurrent and both people are at the same location during the event in which they are involved. Otherwise, the interaction is *spatially distant* (or *distant* for short).

**Dimensions of Relationships.** We consider five dimensions of relationships between people:

- *Equal vs. Hierarchical.* A relationship is *equal* if both people (a) have the same social status, (b) are at the same level in the power

structure, (c) share similar responsibilities, or (d) have the same role. Otherwise, the relationship is *hierarchical*.

- *Intense vs. Superficial.* A relationship is *intense* if both people interact with each other frequently, i.e., they are involved repeatedly. Otherwise, the relationship is *superficial*.
- *Pleasure vs. Task Oriented.* A relationship is *pleasure oriented* if both people interact socially and their relationship is not bound by professional rules or regulations. Otherwise, the relationship is *task oriented*.
- *Intimate vs. Unintimate.* A relationship is *intimate* if both people are emotionally close and warm to each other. Otherwise, the relationship is *unintimate*.
- *Temporary vs. Enduring.* A relationship is *temporary* if it lasts less than a day. A relationship is *enduring* if it last over a month. Otherwise (if it lasts more than a day and less than a month), this dimension is undefined.

#### 4 Annotating Dimensions of Interactions and Relationships

Annotating dimensions of interactions and relationships requires a corpus in which the same people interact several times. We augment an existing corpus of scripts from the TV show Friends (Chen and Choi, 2016). More specifically, we work with the 24 episodes from Season 1 because they:

- contain a large number of conversation turns (9,168, see counts per episode in Table 2);
- involve many characters (42 characters speak at least 100 conservation turns, see the characters that interact the most in Table 2, and the full list in the supplementary materials);

- include speaker information (i.e., we have access to who says what); and
- include annotations linking each mention of people in each conversation turn to the actual person (the name of the person).

Beyond size, the main motivation to use this corpus is the last item above: starting from scratch with another corpus of dialogues would require a substantially larger annotation effort. We refer the reader to the afore-cited paper for details, but the original corpus clusters mentions to people such as *guy*, *my brother* and *he* together with other mentions of the same person. The original corpus is publicly available,<sup>1</sup> and we release our annotations as stand-alone annotations.<sup>2</sup>

#### 4.1 Selecting Pairs of People

The corpus we start with makes it straightforward to select pairs of people whose interactions and corresponding relationships will be annotated. We consider as interactions all instances of somebody mentioning (or referring to) somebody else in a conversation turn. We consider as relationships individuals who interact at least once. Note that we do not (a) distinguish between  $x$  mentioning  $y$  and  $y$  mentioning  $x$ , and (b) consider as an interaction  $x$  talking to  $y$  unless the conversation turn contains a mention to  $y$  (the mention need not be the actual name, it could be a pronoun or any nominal mention). Our rationale is as follows. First, all dimensions of interactions are symmetric; and all dimensions of relationships are symmetric except *equal* vs. *hierarchical*. Second, the characters of Friends refer to each other explicitly at least once in most conversations and scenes, either using first names or the pronoun *you*. Thus we are considering as an interaction most verbal exchanges.

Table 2 shows basic counts per episode. We show the number of interactions, unique relationships (i.e., interactions between unique pairs of people), and the pair of people who interact the most. The supplementary materials include an extended table listing the number of times each pair of people interact per episode.

#### 4.2 Annotation Process

The annotations were done one episode at a time. Annotators were presented with the full transcript of the episode including speaker informa-

<sup>1</sup><https://github.com/emorynlp/character-mining>

<sup>2</sup><http://www.cse.unt.edu/~blanco/>

Episode	#Turns	#Inters	#Rels	People with most inters.
1	544	143	21	Monica-Rachel (23)
2	358	105	19	Rachel-Barry (14)
3	475	98	20	Joey-Chandler (12)
4	359	64	17	Monica-Rachel (15)
5	290	76	18	Chandler-Janice (14)
6	353	82	16	Chandler-Aurora (32)
7	291	76	15	Ross-Rachel (13)
8	371	64	19	Monica- Mrs. Geller (14)
9	375	75	18	Monica-Rachel (13)
10	273	69	19	Phoebe-David (22)
11	447	159	24	Chandler-Ross (22)
12	351	109	23	Phoebe-Paolo (14)
13	328	120	24	Joey-Mr.Tribbiani (21)
14	294	239	21	Chandler-Janice (28)
15	430	78	17	Joey-Ross (15)
16	470	95	31	Chandler-Nina (14)
17	526	109	17	Monica-Rachel(28)
18	306	42	14	Ross-Rachel (9)
19	183	29	10	Ross-Rachel (9)
20	467	111	20	Rachel-Mindy (30)
21	403	85	24	Monica-Fake Monica (23)
22	468	92	22	Phoebe-Chandler (32)
23	457	104	26	Ross-Susan (19)
24	325	107	18	Ross-Rachel (30)

Table 2: Basic corpus counts. We show the number of conversation turns, interactions (i.e., one person referring to another one), unique relationships (i.e., unique pairs of people who interact with each other), and the pairs of people with most interactions.

tion (who speaks what?) and the names of the individuals mentioned in each conversation turn (who do speakers talk about?). Annotators read each episode from the beginning, and annotate dimensions of interactions and relationships after each interaction. Regarding interactions, they were instructed to annotate dimensions taking into account the language of the current conversation turn. Regarding relationships, they were instructed to annotate dimensions taking into account all previous conversation turns within the same episode. For example, if previous turns state that *Rachel* and *Monica* are best friends, the relationship will continue to be annotated *intense* even if an interaction does not indicate so (until a turn indicates that they are not friends, if applicable).

	Speaker	2nd party	Cooperative (I)
1: Wait, does <u>he</u> eat chalk?	Phoebe	Paul	-1
2: Hey, hey, hey, that's not the rule and <u>you</u> know it.	Ross	Woman	-1
3: <u>She</u> is so peaceful.	Monica	Phoebe	1
	Speaker	2nd party	Concurrent (I)
4: Well, then can we meet <u>him</u> ?	Rachel	Alan	-1
5: Hey, <u>buddy</u> , what's up!	Chandler	Alan	1
6: <u>Ma</u> , I'm sorry.	Joey	Mrs. Tribbiani	1
	Speaker	2nd party	Equal (R)
7: Oh, ah-the <u>kid</u> has it.	Joey	the kid	-1
8: Happy birthday, <u>pal</u> !	Chandler	Ross	1
9: Hey <u>dad</u> , what's up?	Monica	Mr. Geller	-1
	Speaker	2nd party	Intense (R)
10: Um, has uh <u>Dr. Franzblau</u> been by?	Rachel	Dr. Franzblau	-1
11: There's a beautiful woman at eight, nine, ten o'clock!	Chandler	Aurora	-1
12: Uh, <u>Rach</u> , it's the Visa card people.	Monica	Rachel	1

Table 3: Annotation examples. We show examples of contrasting values for selected dimensions. The first party is always the speaker, and the second party is underlined. I stands for interactions, and R for relationship.

		raw	$\kappa$
Interaction	Cooperative	93.4%	0.82
	Active	95.4%	0.89
	Concurrent	93.7%	0.84
	Spat. Near	95.7%	0.87
Relationship	Equal	94.6%	0.85
	Intense	88.8%	0.80
	Pleasure Or.	97.4%	0.77
	Intimate	92.3%	0.85
	Temporary	94.8%	0.78

Table 4: Inter-annotator agreement (raw agreement and Cohen's  $\kappa$ ).  $\kappa$  values between 0.6 and 0.8 indicate *substantial* agreement,  $\kappa$  values over 0.8 indicate *perfect* agreement (Artstein and Poesio, 2008).

We discovered during pilot annotations that the value for a dimension sometimes cannot be determined. For example, if the first interaction between *Rachel* and *Monica* is *Rachel: How are [you]<sub>Monica</sub> doing?*, we cannot tell if the relationship is *temporary* or *enduring*. We note, however, that all interaction after we find out that they are best friends (as long as they remain best friends) will be annotated *enduring*. Hereafter, we refer to dimensions by the first descriptor in Table 1, and use 1 if the first descriptor of a dimension is true, -1 if the second descriptor is true, and 0 if neither the first nor the second descriptor can be chosen.

**Annotation Quality.** The annotations were done by two graduate students in computational linguistics.

First, they did pilot annotations to better define the dimensions (Section 3). After several iterations, both of them annotated 3 episodes (15% of all interactions). Table 4 presents the inter-annotator agreements. Cohen's  $\kappa$  range between 0.77 and 0.89, and most (7 out of 9) are above 0.80, which is considered perfect agreement. Values between 0.60 and 0.80 are considered substantial (Artstein and Poesio, 2008). The remaining episodes were annotated once.

### 4.3 Annotation Examples

We present annotation examples in Table 3. The interactions in conversation turns 1 and 2 are *competitive*: Phoebe is ridiculing Paul by asking Monica if he eats chalk, and Ross is confronting an unnamed woman. In turn (3), Monica refers to Phoebe with affection (as the latter sleeps), thus the interaction is *cooperative*. Turns (4–6) exemplify *concurrent* vs. *nonconcurrent*. In (4), Rachel is inquiring whether she can meet Alan (Monica's boyfriend), and Rachel and Alan are not involved in the same event (at this point, the meeting may or may not happen). In (5–6), however the speaker and second party are involved directly in a communication event. In examples (4–6), the values for *active* are the same as for *concurrent*.

Turns 7–12 present examples for dimensions of relationships. Examples 7 and 9 are fairly straightforward: previous interactions reveal that *Joey* is

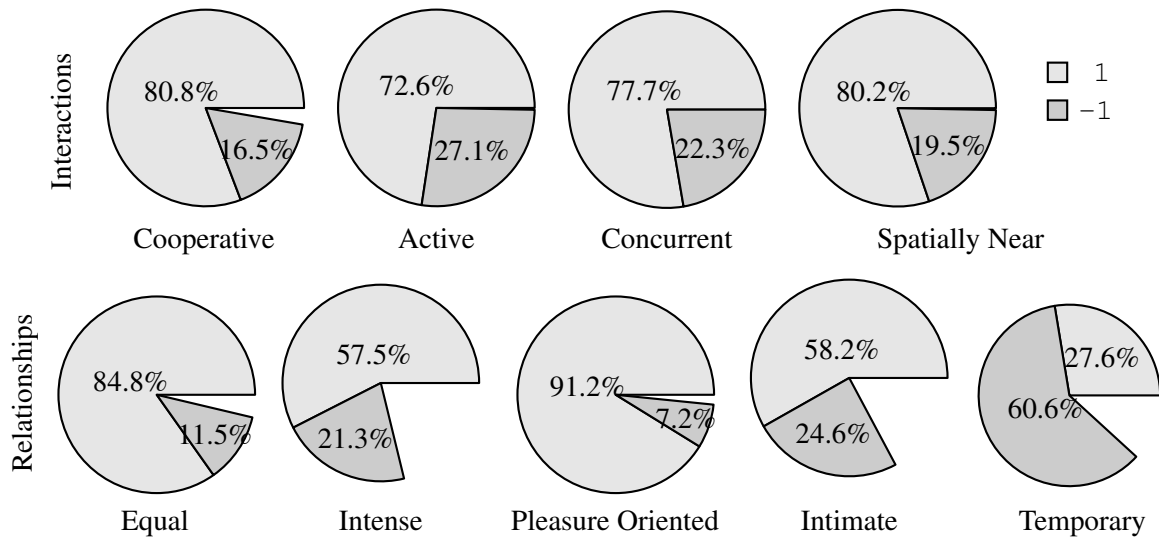


Figure 1: Label percentages per dimension. The missing portion of pie charts correspond to label 0 (unknown).

an adult and Monica’s father is indeed *Mr. Geller*. Example 8 is annotated *equal*, as Chandler and Ross are friends based on previous interactions (the use of *pal* also helps). Examples 10–12 require more explanation, as additional information beyond the current conversation turn is required (recall that dimensions of relationships are annotated taking into account the previous turns within the same episode, Section 4.2). Dr. Franzblau is the doctor of a friend’s ex-wife, so Monica and him have a *superficial* relationship (Turn 10). At the point Turn (11) is spoken by Chandler, she and *Aurora* are strangers, so they have a *superficial* relationship. In (12), previous conversations reveal that Monica and Rachel are close friends, and they interact often (*intense*).

## 5 Corpus Analysis

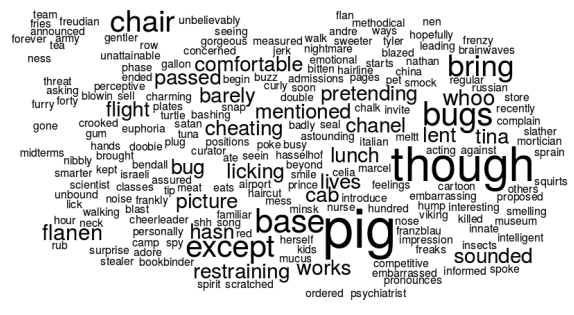
The pie charts in Figure 1 present the label distributions per dimension. Regarding interactions, we note that (a) values for all dimensions can be determined almost always (the percentages of 0 (unknown) are almost zero), and (b) the first descriptor is much more common in all dimensions. These percentages do not represent the distribution of interactions between people in general: the scripts of the TV show *Friends* mostly contain conversation between friends. Regarding relationships, we observe a larger percentage of 0 (unknown) although values of all dimension can be determined most of the time (labels 1 and -1, indicating that the first or second descriptor apply). Most dimensions are biased towards 1 (the only exception is *temporary*, as most relationships are

	Cooperative (I)	Active (I)	Concurrent (I)	Spat. Near (I)	Equal (R)	Intense (R)	Pleasure Or. (R)	Intimate (R)
Active	.06	–						
Conc.	.02	.81	–					
Near	.01	.71	.89	–				
Equal	-.01	.08	.06	.05	–			
Intense	.08	.28	.24	.22	.28	–		
Pleasure	-.05	.01	.08	.08	.50	.13	–	
Intimate	.12	.28	.26	.22	.32	.75	.28	–
Temp.	-.02	-.33	-.31	-.28	-.24	-.52	-.27	-.54

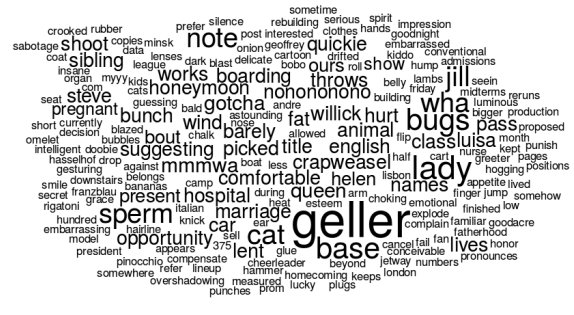
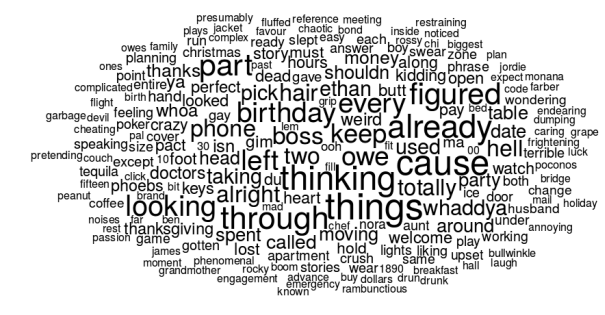
Table 5: Pearson correlations between pairs of dimensions of interactions and dimensions.

*enduring*), especially *pleasure oriented* and *equal* (91.2% and 84.8%). Again, these distributions would be different if we worked with other sources of dialogue than the TV show *Friends*.

Many of the dimensions we consider in this work are intuitively correlated. For example, *concurrent* interactions must be *active*, and *pleasure oriented* interactions are probably also *equal*. We note, however, that interactions can be *passive* and *concurrent*, e.g., in (*Monica talking to Joey*) [*He*]<sub>Paul</sub> is just [*a guy*]<sub>Paul</sub> I am dating!, Monica and Paul have a *passive* and *concurrent* interaction (they are dating, but they are not talking to each other). Table 5 shows Pearson correlations between all dimensions of interactions and relationships. Most correlations are under 0.3 (29 out



Spatially Near vs. Spatially Distant (Interaction)



Intense vs. Superficial (Relationship)

Figure 2: Most salient words (calculated with tf-idf) for dimensions *spatially near* and *intense*.

of 36), although some pairs do have high correlations. In particular, *active* interactions tend to be both *concurrent* (0.81) and *spatially near* (0.71), and *spatially near* interactions tend to be concurrent (0.89). Regarding relationships, *intimate* correlates with *intense* (0.75), *pleasure oriented* with *equal* (0.50), and *temporary* with both *superficial* (0.52) and *intimate* (0.54).

Finally, Figure 2 shows the most salient words of dimensions *spatially near* and *intense*. We calculated salience using tf-idf (Schütze et al., 2008). Interactions containing derogatory words (e.g., *pig*, *bugs*, *pretending*, *cheating*) tend to be *distant*, and *near* interactions contain mostly neutral and nicer words such as *friends*, *sweatheart* and *please*. We also note that cognitive verbs and nouns (e.g., *thinking*, *figured (out)*, *looking (into)*, *cause*), as well as important events (*birthday*, *thanksgiving*) and slang usage (e.g., *whad-dya*) signal *intense* relationships.

## 6 Experiments and Results

We experimented with SVM classifiers with RBF kernel to predict dimensions of interactions and relationships. We divided the 24 episodes into train (episodes 1–20) and test (21–24), and trained one classifier per dimension using scikit-learn (Pe-

		P	R	F
Majority Baseline		0.54	0.73	0.62
SVM	first word	0.62	0.71	0.65
	BOW current	0.66	0.73	0.67
	+ sentiment	0.65	0.72	0.66
	+ other	0.71	0.75	0.70
	+ BOW previous	0.73	0.76	0.72

Table 6: Results obtained with the full set with several systems (average of all dimensions). *Previous* refers to the previous conversation in which the same pair of people interacted not the immediately previous turn).

dregosa et al., 2011). Each classifier is trained with three labels: 1 (1st descriptor), -1 (2nd descriptor) and 0 (unknown). The SVM parameters (C and  $\gamma$ ) were tuned using 10-fold cross-validation with the train split, and results are reported using the test split.

Note that different pairs of people interact more or less in each episode (Table 2). Thus, the classifiers are grounded on general language usage and not modeling who talks and who is talked about. We also experimented with LSTMs taking as input the current conversation turn and previous turns, but do not report results because SVM classifiers yielded better results.

		1 (1st descriptor)			0 (unknown)			-1 (2nd descriptor)			All		
		P	R	F	P	R	F	P	R	F	P	R	F
Interaction	Cooperative	0.83	0.96	0.89	0.00	0.00	0.00	0.31	0.08	0.13	0.73	0.80	0.75
	Active	0.92	0.90	0.92	n/a	n/a	n/a	0.75	0.76	0.75	0.87	0.87	0.87
	Concurrent	0.92	0.92	0.92	n/a	n/a	n/a	0.70	0.69	0.70	0.87	0.87	0.87
	Spat. Near	0.89	0.94	0.91	n/a	n/a	n/a	0.69	0.53	0.60	0.85	0.86	0.85
Relationship	Equal	0.86	0.95	0.90	0.00	0.00	0.00	0.15	0.07	0.10	0.76	0.83	0.79
	Intense	0.70	0.81	0.75	0.28	0.11	0.16	0.38	0.46	0.41	0.56	0.60	0.57
	Pleasure Or.	0.82	1.00	0.90	0.00	0.00	0.00	1.00	0.02	0.04	0.82	0.83	0.75
	Intimate	0.61	0.85	0.71	0.26	0.05	0.09	0.42	0.36	0.39	0.48	0.56	0.49
	Temporary	0.53	0.42	0.47	0.67	0.17	0.27	0.62	0.85	0.72	0.60	0.60	0.57

Table 7: Results obtained per dimension with the best system (all features, Table 6). The results under *All* the weighted averages for all labels, recall that the label distribution is biased (Figure 1).

**Feature Set.** We use a combination of features extracted directly from the conversation turn, sentiment lexica and context. Specifically, we extract (a) the first word in the conversation turn, (b) bag-of-words features (binary flags and tf-idf scores), and (c) the root verb, and flags indicating the presence of exclamation, question marks and negation cues from (Morante and Daelemans, 2012) (other). Regarding sentiment, we extract flags indicating whether the turn has a positive, negative or neutral word in the list by Hamilton et al. (2016), the sentiment score of the turn (summation of sentiment scores per token over number of tokens in the turn), and a flag indicating whether the turn contains a negative word from the list by Hu and Liu (2004). Regarding context, we extract bag-of-words features from the previous conversation turn in which the same people interact (not necessarily the preceding turn).

## 6.1 Results

Table 6 shows the overall results (average of all dimensions) obtained with the majority baseline and several feature combinations. All feature combinations outperform the baseline. Sentiment features are not beneficial, leading to the conclusion that sentiment does not correlate with dimensions of interactions and relationships between people. This may look surprising at first sight, but recall that our dimensions capture much more than if two people get along (Table 1). Finally the bag-of-words features from the previous turn in which the same people interacted bring a small improvement (F: 0.70 vs. 0.72).

We show results per dimension for the best feature combination (all) in Table 7. Despite the label

distributions are biased (Figure 1), the system predicts most labels for most dimensions except the very biased ones (*cooperative*, *equal* and *pleasure oriented*). Note that 0 (unknown) does not allow us to determine the value of a dimension, and the low results with this label are not a concern.

## 7 Conclusions

This paper presents the task of characterizing interactions and relationships between people. We work with dialogue transcripts, and define an *interaction* as a speaker referring to somebody else, and a *relationship* as a sequence of one or more interactions. Unlike previous work (Section 2), we target all interactions and relationships, and use dimensions that are applicable to any interaction or relationship regardless of the underlying type (e.g., SIBLINGS, FRIENDS, DOCTOR-PATIENT).

We have presented an annotation effort on 24 episodes of the popular TV show Friends (Season 1). The total number of conversation turns is 9,168, and the total number of interactions is 2,331. The label distribution per dimension shows that the labels are unbalanced, but a relatively straightforward SVM is able to outperform the majority baseline (F: 0.62 vs. 0.72, Table 6). Features extracted using well-known sentiment lexica yield no improvements, leading to the conclusion that the dimensions we work with capture information beyond whether two people get along.

Crucially, values for the dimensions we work with can be determined most of the time (Figure 1, labels 1 and -1). Since we do not impose any restriction on the interactions or relationships we work with, we conclude that these dimensions may be universally applicable.



## References

- John Adamopoulos. 2012. The emergence of social meaning: A theory of action construal. In *Handbook of Social Resource Theory*, pages 255–272. Springer.
- Gabor Angeli, Melvin Jose Johnson Premkumar, and Christopher D. Manning. 2015. Leveraging linguistic structure for open domain information extraction. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 344–354, Beijing, China. Association for Computational Linguistics.
- Ron Artstein and Massimo Poesio. 2008. Inter-coder agreement for computational linguistics. *Comput. Linguist.*, 34(4):555–596.
- David B. Bracewell, Marc T. Tomlinson, Mary Brunson, Jesse Plymale, Jiajun Bracewell, and Daniel Boerger. 2012. Annotation of adversarial and collegial social actions in discourse. In *Proceedings of the Sixth Linguistic Annotation Workshop, LAW VI '12*, pages 184–192, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Philip Bramsen, Martha Escobar-Molano, Ami Patel, and Rafael Alonso. 2011. Extracting social power relationships from natural language. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 773–782, Portland, Oregon, USA. Association for Computational Linguistics.
- Yu-Hsin Chen and Jinho D. Choi. 2016. Character identification on multiparty conversation: Identifying mentions of characters in tv shows. In *Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 90–100. Association for Computational Linguistics.
- Cristian Danescu-Niculescu-Mizil, Lillian Lee, Bo Pang, and Jon Kleinberg. 2012. Echoes of power: Language effects and power differences in social interaction. In *Proceedings of the 21st International Conference on World Wide Web, WWW '12*, pages 699–708, New York, NY, USA. ACM.
- Cristian Danescu-Niculescu-Mizil, Moritz Sudhof, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. A computational approach to politeness with application to social factors. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 250–259, Sofia, Bulgaria. Association for Computational Linguistics.
- Morton Deutsch. 2011. Interdependence and psychological orientation. In *Conflict, Interdependence, and Justice*, pages 247–271. Springer.
- George Doddington, Alexis Mitchell, Mark Przybocki, Lance Ramshaw, Stephanie Strassel, and Ralph Weischedel. 2004. The automatic content extraction (ace) program tasks, data, and evaluation. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC-2004)*, Lisbon, Portugal. European Language Resources Association (ELRA). ACL Anthology Identifier: L04-1011.
- Raymond W Gibbs. 2000. Irony in talk among friends. *Metaphor and symbol*, 15(1-2):5–27.
- Eric Gilbert. 2012. Phrases that signal workplace hierarchy. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work, CSCW '12*, pages 1037–1046, New York, NY, USA. ACM.
- Ralph Grishman and Beth Sundheim. 1996. Message understanding conference-6: A brief history. In *Proceedings of the 16th Conference on Computational Linguistics - Volume 1, COLING '96*, pages 466–471, Stroudsburg, PA, USA. Association for Computational Linguistics.
- William L. Hamilton, Kevin Clark, Jure Leskovec, and Dan Jurafsky. 2016. Inducing domain-specific sentiment lexicons from unlabeled corpora. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 595–605, Austin, Texas. Association for Computational Linguistics.
- Minqing Hu and Bing Liu. 2004. Mining and summarizing customer reviews. In *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '04*, pages 168–177, New York, NY, USA. ACM.
- Mohit Iyyer, Anupam Guha, Snigdha Chaturvedi, Jordan Boyd-Graber, and Hal Daumé III. 2016. Feuding families and former friends: Unsupervised learning for dynamic fictional relationships. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1534–1544, San Diego, California. Association for Computational Linguistics.
- Valentin Jijkoun, Maarten de Rijke, and Jori Mur. 2004. Information extraction for question answering: Improving recall through syntactic patterns. In *Proceedings of the 20th International Conference on Computational Linguistics, COLING '04*, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Harold H Kelley. 2013. *Personal relationships: Their structures and processes*. Psychology Press.
- Seth Kulick, Ann Bies, and Justin Mott. 2014. Inter-annotator agreement for ere annotation. In *Proceedings of the Second Workshop on EVENTS: Definition, Detection, Coreference, and Representation*, pages 21–25, Baltimore, Maryland, USA. Association for Computational Linguistics.

- Roser Morante and Walter Daelemans. 2012. Conandoyle-neg: Annotation of negation cues and their scope in conan doyle stories. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC-2012)*, pages 1563–1568, Istanbul, Turkey. European Language Resources Association (ELRA). ACL Anthology Identifier: L12-1077.
- Dat Quoc Nguyen, Kairit Sirts, Lizhen Qu, and Mark Johnson. 2016. Stranse: a novel embedding model of entities and relationships in knowledge bases. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 460–466, San Diego, California. Association for Computational Linguistics.
- F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- Vinodkumar Prabhakaran and Owen Rambow. 2014. Predicting power relations between participants in written dialog from a single thread. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 339–344, Baltimore, Maryland. Association for Computational Linguistics.
- Farzana Rashid and Eduardo Blanco. 2017. Dimensions of interpersonal relationships: Corpus and experiments. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2297–2306, Copenhagen, Denmark. Association for Computational Linguistics.
- Hinrich Schütze, Christopher D Manning, and Prabhakar Raghavan. 2008. *Introduction to information retrieval*, volume 39. Cambridge University Press.
- Mark Snyder and Arthur A Stukas Jr. 1999. Interpersonal processes: The interplay of cognitive, motivational, and behavioral activities in social interaction. *Annual review of psychology*, 50(1):273–303.
- Mihai Surdeanu and Ji Heng. 2014. Overview of the english slot filling track at the tac2014 knowledge base population evaluation. In *Proceedings of the TAC-KBP 2014 Workshop*.
- Nikos Voskarides, Edgar Meij, Manos Tsagkias, Maarten de Rijke, and Wouter Weerkamp. 2015. Learning to explain entity relationships in knowledge graphs. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 564–574, Beijing, China. Association for Computational Linguistics.
- Robert West, Evgeniy Gabrilovich, Kevin Murphy, Shaohua Sun, Rahul Gupta, and Dekang Lin. 2014. Knowledge base completion via search-based question answering. In *Proceedings of the 23rd International Conference on World Wide Web, WWW '14*, pages 515–526, New York, NY, USA. ACM.
- Michael White, Tanya Korelsky, Claire Cardie, Vincent Ng, David Pierce, and Kiri Wagstaff. 2001. Multidocument summarization via information extraction. In *Proceedings of the First International Conference on Human Language Technology Research, HLT '01*, pages 1–7, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Myron Wish, Morton Deutsch, and Susan J Kaplan. 1976. Perceived dimensions of interpersonal relations. *Journal of Personality and Social Psychology*, 33(4):409.
- Fei Wu and Daniel S. Weld. 2010. Open information extraction using wikipedia. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, ACL '10*, pages 118–127, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Xiaofeng Yu and Wai Lam. 2010. Jointly identifying entities and extracting relations in encyclopedia text via a graphical model approach. In *Proceedings of the 23rd International Conference on Computational Linguistics: Posters, COLING '10*, pages 1399–1407, Stroudsburg, PA, USA. Association for Computational Linguistics.