Beyond Possession Existence: Duration and Co-Possession

Dhivya Chinnappa[*]
Center for Cognitive Computing
Thomson Reuters

Srikala Murugan and **Eduardo Blanco**
Human Intelligence and Language Technologies Lab
University of North Texas

Abstract

This paper introduces two tasks: determining (a) the duration of possession relations and (b) co-possessions, i.e., whether multiple possessors possess a possessee at the same time. We present new annotations on top of corpora annotating possession existence, and experimental results. Regarding possession duration, we derive the time spans we work with empirically from annotations indicating lower and upper bounds. Regarding co-possessions, we use a binary label. Cohen's kappa coefficients indicate substantial agreement, and experimental results show that text is more useful than the image for solving these tasks.

1 Introduction

Relation extraction is a core problem in natural language processing. Extracting relations is generally defined as linking two text chunks with a label. For example, relations such as PRESI-DENT_OF and MARRIED_TO are common in information extraction (Angeli et al., 2015). Within computational semantics, relations capture spatial and temporal knowledge (Kordjamshidi et al., 2018; McDowell et al., 2017), as well as many other meanings (Abend and Rappoport, 2017).

Approaches to relation extraction usually only determine the right label—often referred to as relation *name* or *type*—between two text chunks. Relation labels are certainly useful, but there is almost always complementary information that can be extracted. For example, relation labels do not give any hint about for how long the relation holds true or whether the relation is one-to-one or one-to-many. Many relations would benefit from having this additional information available, including LOCATED_AT (people have many locations over time) and AGENT (some events are carried out by

Figure 1: Sample tweet with text and an image. The author of the tweet possesses the *cup* for a few weeks or months. The tweet does not indicate a co-possession.

only one person but not all; the additional agents may not be explicitly named in a given text).

Possession relations are ubiquitous and understudied from a computational perspective. Possessions are defined as someone (the possessor) possessing something (the possessee), where *possessing* includes not only ownership but also control, kinship, physical and temporal proximity, and others (Section 2). From a computational perspective, previous work on extracting possessions targets possession existence (i.e., whether a possessor *x* possesses a possessee *y*) and limited temporal information using anchors, (e.g., at some point of time before or after an event, Section 2).

In this paper, we complement previous work targeting possession existence with two attributes: duration (for how long does the possession hold true?) and co-possession (are there other possessors possessing the possessee concurrently?). Consider the tweet in Figure 1. The possessee is the *cup*, and from the text we understand that it is reusable. Thus the author of the tweet is likely to have the cup for a few weeks or months. If the

---

[*]Work done at the University of North Texas

possessee were a paper cup, however, the author would probably have it for at most one hour. Similarly, if the possessee were a personal coffee mug, the author would have it for longer—probably years. On the other hand, if either the text or image indicated that the setting was a restaurant, the author most likely would only have the cup for at most a couple hours, and there would be a co-possession—the restaurant and the customer.

The main contributions of this paper are: (a) strategy to determine sound intervals for possession durations grounded on lower and upper temporal bounds; (b) corpus of possession relations annotated with durations and co-possessions;[1] (c) detailed corpus analysis; and (d) experimental results showing that both tasks can be automated. While we work with possessions, a similar approach could be used to determine the duration of any relation and distinguish between one-to-one and one-to-many relations.

2   Related Work

Most previous work on relation extraction does not identify the temporal bounds during which a relation holds true. There are, however, some exceptions that assign temporal information to relations (Ji et al., 2011; McClosky and Manning, 2012). Unlike these previous efforts, we work with durations that are rarely explicitly stated.

Previous works on extracting possession relations primarily fall under efforts to extract large relation inventories. The goal of these efforts is to identify which relation—out of a predefined inventory—holds between two arguments. For example, Tratz and Hovy (2013) investigate semantic relations realized by English possessive constructions, both Nakov and Hearst (2013) and Tratz and Hovy (2010) consider relations realized by noun compounds such as *family estate*, and Badulescu and Moldovan (2009) extract relations realized by English genitives. Recently, Blodgett and Schneider (2018) present a corpus of web reviews in which the s-genitive and of-genitive are annotated with semantic labels (or supersenses). Regardless of the lexico-syntactic pattern, possession relations are a minority of the relations targeted by these previous works (other relations include THEME, QUANTITY, CAUSE, ORIGINATOR, EXPERIENCER, etc.). In addition, they do not target possession duration or co-possession.

To the best of our knowledge, there are three previous works on extracting possession relations. All of them introduce their own annotations and present experimental results. In our previous work (Chinnappa and Blanco, 2018), we consider possession relations between individuals (named entity *person* and personal pronouns) and concrete objects mentioned within the same sentence in the OntoNotes corpus. Regarding time, we indicate whether the possession held true before, during or after the event in the sentence. Banea and Mihalcea (2018) consider possessions between the author of a weblog (i.e., the possessor is fixed) and the possessees identified in the weblog. Regarding time, they exclusively target possessions that held true when the weblog was written—not before or after. More recently, we investigate the problem of determining whether authors of tweets possess the objects they tweet about, and use tweets consisting of text and images (Chinnappa et al., 2019). All of these previous efforts target possession existence (i.e., whether a possession relation holds true) and very limited temporal information. Unlike them, we go beyond possession existence and target possession duration and co-possession.

Finally, we note that theoretical works consider having temporary *control* of something as a type of possession (Tham, 2004). For example, ship captains and plane pilots have control possession of the ships and planes under their command, but usually not ownership or alienable possession. Similarly, office workers have control possession of their work desk and computer, but they do not own them. According to this definition, control possessions indicate co-possession. We note, however, that control possessions are only a subset of possessions thus they are insufficient to determine co-possession.

Event Durations. Our methodology to annotate possession durations is heavily inspired by previous work targeting event durations (Pan et al., 2011). The main difference is that we do not target events (e.g., How long did *met* in *John met his advisor on Thursday* last?) but possession relations. As we shall see, we derive sound time intervals for possession durations from lower and upper temporal bounds. To the best of our knowledge, we are the first to target the duration in which a semantic relation holds true. Not surprisingly, we find that possession durations tend to be longer than events. For example, events may last only a few seconds

8333

(*e.g., turn on a car*), but possessions last at least a few minutes and many last over a year.

3 Annotating Possession Duration and Co-possession

To the best of our knowledge, we are the first to go beyond possession existence and target possession duration and co-possession. More generally, we are the first to determine for how long a semantic relations holds true, and distinguish between one-to-one and one-to-many relations. Thus, we create a new corpus to tackle these tasks.

Source Corpora. Starting from plain text is a straightforward choice. Since existing corpora already annotate possession existence, however, it would be suboptimal. Thus we work with the corpora by Chinnappa and Blanco (2018), Banea and Mihalcea (2018), and Chinnappa et al. (2019), and enhance their possession existence annotations with possession duration and co-possession annotations. These source corpora contain 2,257 possession relations, a relatively small amount. We note, however, that the source corpora are diverse (Section 2) and include possession relations identified in formal (OntoNotes) and informal texts (weblogs, Twitter). Additionally, we work with possessions identified from not only text (OntoNotes and weblogs), but also tweets consisting of text and images.

The corpus by Chinnappa and Blanco (2018) contains 979 sentences, and we select the 358 intra-sentential possessions annotated in those sentences. The corpus by Banea and Mihalcea (2018) contains 799 possession relations. The possessor is always the author of a weblog, and the possessee is mentioned in the weblog and can be: (a) a concrete object, e.g., *car*, *notebook*; (b) an implicit concrete object associated with an event, e.g., *car* for *driving*, *cell phone* for *texting*; or (c) an abstract object, e.g., *wifi*, *idea*. The corpus by Chinnappa et al. (2019) contains 5,000 tweets (text + image). We select 1,100 tweets in which the author (the possessor) possesses a concrete object mentioned in the tweet (the possessee).

3.1 Annotation Process and Post-Processing

The annotations were done by two graduate students who fully annotated the whole corpus. Regarding possession duration, they annotate lower and upper bounds. Then, we post-process their annotations to obtain time intervals for possession durations. Regarding co-possession, they use a binary label and no post-processing takes place.

3.1.1 Possession Duration

How long do possession relations hold true for? The answer to this question is not obvious, and previous work has named temporal durations in general a significant issue for temporal reasoning (Allen and Ferguson, 1994). Intuitively, possessors have possession of some possessees for short periods of time (e.g., ice cream, pencils) and other possessees for long periods of time (e.g., cars). But there are exceptions, e.g., drivers have (relatively) short possessions of rental cars—at least compared to the cars they own. In addition, possession durations are almost never explicitly stated in text (e.g., *I got rid of this computer 5 years after buying it*), despite humans have no issues inferring some duration information.

To address the inherent difficulties of annotating temporal durations, we follow previous work on determining event durations (Pan et al., 2011). Specifically, we ask annotators to provide lower and upper bounds for the duration of the possession relation between possessor and possessee (recall that we already know whether a possession exists). Lower and upper bounds consist of an integer followed by a unit of time (seconds, minutes, hours, days, weeks, months or years). These annotations are rather open and we do not expect to obtain high agreements. As we shall see, however, a simple post-processing allows us to obtain sound time intervals for possession duration, where *sound* means empirically driven and with substantial agreements (Section 3.2).

We argue that any predefined duration intervals (e.g., less than five minutes, between five minutes and a day, more than a day and less than a month, over a month) would be arbitrary—at least to a certain degree. Additionally, we would have to go back and forth annotating and redefining the predefined intervals until we obtain (a) a reasonable distribution of duration intervals (e.g., avoid 95% of possessions assigned to a single interval) and (b) substantial agreements. Asking annotators for lower and upper bounds and the proposed post-processing bypasses all these issues.

Post-Processing Possession Durations. We post-process the annotations of lower and upper bounds for possession durations following two steps:

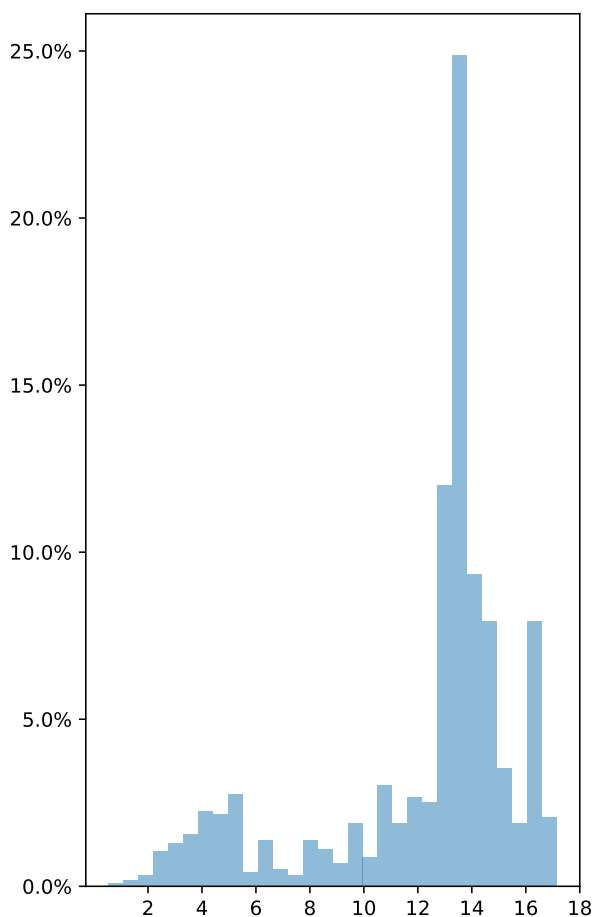1. Convert lower and upper bounds to minutes and calculate the mean.

Figure 2: Distribution of mean possession durations after post-processing (i.e., after converting to minutes and calculating the natural logarithm). We determine duration labels after identifying changes in frequency at 6 (6 hours) and 13 (10 months).

2. Calculate the natural logarithm of the mean duration from Step (1).

Converting to minutes allows us to measure time with a single unit and facilitates further post-processing and calculating agreements (Section 3.2). We convert to minutes (as opposed to, for example, seconds) because the annotators never chose less than a minute as a lower bound. Calculating the logarithm is useful to account for the fact that temporal differences must be calculated in relative terms. For example, the differences between (a) 5 minutes and 10 minutes and (b) 5 years and 10 years should be roughly the same. On the other hand, the differences between (b) 5 years and 10 years and (c) 5 years and 5 minutes, and 10 years and 10 minutes should be close to zero.

Figure 2 plots the frequency of mean possession durations after post-processing. The distribution shows a drop at 6 (equivalent to 6 hours) and a rise at 13 (equivalent to 10 months). Based on this

observations, we define the following intervals to specify possession durations:

- `short`: possessions lasting less than 6 hours,
- `medium`: possessions lasting at least 6 hours and less than 10 months; and
- `long`: possessions lasting at least 10 months.

The annotations we release include (a) lower and upper bounds and (b) the 3-way labels for each possession existence. Except to discuss agreements, however, in the remaining of this paper we work with the three duration labels.

3.1.2 Co-Possession

Annotating co-possession is relatively straightforward. Knowing that a possession relation exists between a possessor *x* and a possessee *y*, annotators use a binary label to indicate whether an additional possessor *x'* has possession of *y* concurrently with *x*. *x'* must not be named explicitly, as otherwise an explicit possession relation would exist. Co-possession can sometimes be determined based on the possessee. For example, commercial plane pilots have control possession of the planes they fly, but usually there are concurrent possessors (e.g., co-pilot, owner). Determining many co-possessions, however, requires context. For example, consider a blogger writing down *I was using the wifi at the coffee shop*. There is a possession relation between the author of the blog and *wifi*, and that is a co-possession because other people are concurrent possessors (e.g., the owners of the coffee shop, other clients).

3.2 Inter-Annotator Agreement

Possession Duration: `short`, `medium` and `long`. We use unweighted Cohen's kappa (κ) to calculate the inter-annotator agreement with the three possession duration labels: `short`, `medium` and `long`. The κ coefficient is 0.63, which is consider *substantial*. Interpreting κ coefficient is somewhat subjective, but over 0.8 would be considered *nearly perfect* (Artstein and Poesio, 2008). We also note that a weighted version of agreement would yield higher agreements.

Possession Duration: Lower and Upper Bounds. Calculating agreement between the lower and upper bounds for possession duration is not straightforward. For example, the agreement between *at least 30 minutes and at most 12 hours* and *at least 1 hour and at most 1 day* should be considerable despite the lower and upper bounds differ by a sizable amount (half and double respec-
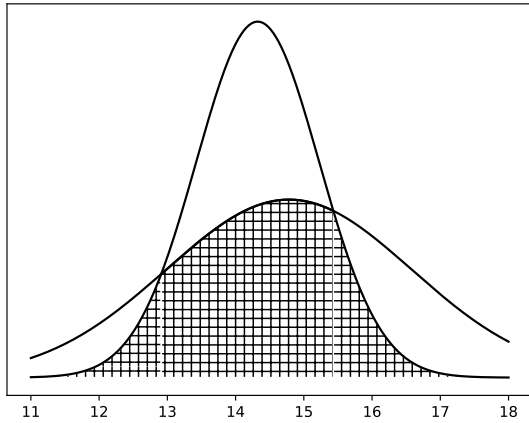
Figure 3: Observed agreement for the POSSESSION(*x*, *y*) in *[We]$_x$ brought the kids rod and reels from [home]$_y$ so they could fish*. The first annotator chose 6 months and 50 years as lower and upper bound (steeper curve), and the second annotator chose 1 year and 100 years (flatter curve). Observed agreement is the overlap between both curves, which is 0.64.

Only text (source: OntoNotes and weblogs)		
Possession duration	short	15.1%
	medium	6.2%
	long	78.7%
Co-Possession	no	56.5%
	yes	43.5%
Text + image (source: Tweets)		
Possession duration	short	4.3%
	medium	38.0%
	long	57.7%
Co-Possession	no	72.7%
	yes	27.3%

Table 1: Label distributions. Top block: possessions identified in text (from OntoNotes and weblogs); bottom block: possessions identified in text and image (from tweets).

tively). Cohen's κ is usually used for categorical labels and not directly applicable to ranges of durations defined by lower and upper bounds. We follow previous work on event durations to calculate the agreement (Section 2).

The formula for Cohen's κ is $\kappa = \frac{P(A)-P(E)}{1-P(E)}$, where $P(A)$ is the observed agreement between annotators and $P(E)$ is the expected agreement. We assume that possession durations follow a normal distribution, and that the lower and upper bounds account for 80% of the distribution. Under these assumptions, the lower (x_{lower}) and upper(x_{upper}) bounds are 1.28 standard deviations (σ) from the mean (μ), thus $\sigma = \frac{x_{upper}-\mu}{1.28} = \frac{x_{lower}-\mu}{-1.28}$ and $\mu = \frac{x_{upper}-x_{lower}}{2}$.

We calculate observed agreement between annotations (P(A)) as the overlap between their normal distributions, as exemplified in Figure 3. We calculate expected agreement (P(E)) as the average overlap between each annotation and the global distribution. In other words, the expected agreement would result from annotations that follow perfectly the global normal distribution.

The κ coefficient for lower and upper bounds is low, 0.37. We note, however, that (a) it would be larger if we assumed that annotators annotate less than 80% of the duration distribution, and (b) previous work on event durations obtained 0.08 κ under the same assumptions. Additionally, we experiment with the three duration intervals described above (κ: 0.63); our rationale to annotate lower

and upper bounds is to derive sound intervals.
Co-Possession. The Cohen's kappa (κ) coefficient for co-possession (two labels: yes and no) is 0.65, which again is considered *substantial*.

4 Corpus Analysis

Table 1 presents the label distribution in our corpus. We distinguish between possessions identified in text (Chinnappa and Blanco, 2018; Banea and Mihalcea, 2018), and those identified in tweets consisting of text and an image (Chinnappa et al., 2019). Regarding possession duration, most possessions are long (over 10 months, 78.7% and 57.7%). Possessions identified in tweets are much more likely to have medium length (38.0%) than those identified in text (6.2%), and the opposite it true about short durations: 4.3% vs. 15.1%. Regarding co-possession, yes and no are roughly uniformly distributed with possessions identified in text (yes: 56.5% and no: 43.5%). In tweets consisting of text and an image, however, no dominates yes (72.7% vs. 27.3%).

We present label distributions based on the WordNet synset and number of the possessee in Table 2. The majority (96.5%) of possessees are nouns. The top 4 most frequent WordNet synsets (*container*, *device*, *vehicle*, and *covering*) show interesting patterns. First, *vehicles* (e.g., *car*, *truck*) and *containers* (e.g., *handbag*, *spoon*) are most of the times part of long possessions. Second, *devices* (e.g., *comb*, *cell phone*) are twice as likely to be part of a medium length possession. Third, *coverings* (e.g., *jacket*, *pants*, *shirt*) are (b.1) almost never part of short possessions and (b.2) almost always (80%) part of long possessions Pos-

		WordNet Synsets (top 4 most frequent)				Number			
		Container	*Device*	*Vehicle*	*Covering*	Singular	Plural	Not noun	*%*
Possession duration	short	1.5	2.3	2.5	0.6	6.9	2.1	0.7	9.7
	medium	4.8	6.1	4.0	1.7	14.4	7.3	0.4	22.1
	long	18.4	9.5	10.8	8.0	49.0	16.8	2.4	68.2
	All	24.7	17.9	17.3	10.3	70.3	26.2	3.5	100.0
Co-Poss.	no	13.5	11.8	11.2	4.5	43.1	19.5	2.0	64.6
	yes	11.2	6.1	6.1	5.8	27.2	6.7	1.5	35.4
	All	24.7	17.9	17.3	10.3	70.3	26.2	3.5	100.0

Table 2: Label distribution of duration and co-possession labels depending on the WordNet synset and number of the possessee. All numbers are percentages in the whole corpus (text and text + images).

	Sentence with possessor *x* and possessee *y*	Duration	Co-Poss.
1	Everything served cold, with [ice cream]$_y$, fruit salad and strawberry yoghurt pudding for dessert [...] (*x*: the author of the weblog)	short	no
2	"At least [we]$_x$ have the decency to drop [bombs]$_y$ from airplanes", he said.	medium	no
3	I had to get out my [phone]$_y$ for a couple pics. (*x*: the author of the weblog)	long	no
4	[We]$_x$ took a [taxi]$_y$ along the path of the highway that heads toward Disney, trying to experience this mysterious park from close by.	short	yes
5	The first two months of the summer, I drove Andrew's wrapped car that has his face all over it (lucky me), and then the last two months, the dealership was able to provide me with a [loaner car]$_y$. (*x*: author of the weblog)	medium	yes
6	[They]$_x$ kept my father's [car]$_y$ for a year without writing a confiscation order.	long	yes

Table 3: Annotation examples on selected possessions identified in text. (*x*: possessors, *y*: possessees).

sesses not present in WordNet (e.g., *Garmin*, *dupioni*) and those not subsumed by the top 4 most frequent synsets have roughly the same distribution than all possessees (Table 1). Regarding co-possession, *devices* (e.g., *computer*, *watch*) and *vehicles* (e.g., *plane*, *truck*) follow a similar distribution: co-possession is roughly twice as likely. The distribution of other synsets indicate that possessees are unlikely to have co-possessors, but to a lesser degree. The right-hand side of Table 2 shows the label distributions depending on the possessee number. Plural and singular nouns follow a similar distribution with possession duration, but plural nouns are less likely to have concurrent co-possessors than singular nouns.

Examples. Table 3 presents annotation examples on top of possessions identified in text.

In Example (1), the possessor is the author of the blog and the possessee is the *ice cream*. The author is describing a meal, and it is clear that the possession lasted for a short period of time. There is no indication that the author shared the *ice cream* thus annotators chose no for co-possession.

Example (2) belongs to a document describing a war zone were *bombs* (the possessee) were dropped. Annotators interpreted that the speaker uses *we* to refer to his nation, and annotated medium duration as bombs are not stored for long periods of time during war. They also decided that there is no co-possession since the possessor *we* is a collective noun referring to an entire nation. Example (3) is from a weblog. The possessor is the author and the possessee is a *phone*. It is reasonable to infer from context that the possessee is a cell phone (landline phones do not have cameras) and that the author is the owner. Thus, annotators chose long duration and no co-possession.

In Example (4), the possessor *we* is the client of a taxi driver, and the possessee is the *taxi*. While not explicitly stated, annotators inferred that (a) the possession lasted for a short period of time and (b) there are concurrent co-possessors (e.g., the taxi driver). Note that the possession duration between the taxi driver and the same possessee is likely to be medium or long, but we only annotate the duration between *we* and *taxi*.

Example (5) illustrates a rare phenomenon: an explicit temporal interval (i.e., *two months*) indicating the possession duration. Thus, annotators chose medium duration. Regarding co-possession, the company loaning the car was clearly a co-possessor of the *loaner car* while the author of the blog borrowed the car, so annotators chose yes.

Finally, Example (6) exemplifies a long pos-

| a) Scotty ✿ @FC2020 @S · Apr 17 | b) larry0711@sbcglobal.n... · Apr 17 | c) AJ @ldkblue · Apr 17 |

Possessee: *bowl*		Possessee: *pen*		Possessee: *computer*	
Duration	Co-Poss.	Duration	Co-Poss.	Duration	Co-Poss.
`short`	`no`	`medium`	`no`	`long`	`no`

| d) TB76 Returns @TB76_... · Apr 18 | e) P.B. @sweetnuggetTL · Apr 16 | f) вuттяσ¢кſαη69 @SSR_John · Apr 15 |

Possessee: *hats*		Possessee: *jackets*		Possessee: *shirts*	
Duration	Co-Poss.	Duration	Co-Poss.	Duration	Co-Poss.
`short`	`yes`	`medium`	`yes`	`long`	`yes`

Table 4: Annotation examples on selected possessions identified in tweets consisting of text and an image. The possessors are the authors of the tweets, and the possessees are concrete objects in their tweets.

session with co-possession. The context is a law enforcement operation in which *They* (the police) kept the possessee (car). The duration of the possession is explicit (*a year*), and during that time *my father* was still the owner. Thus, annotators chose `long` and `yes` for duration and co-possession.

Table 4 presents annotation examples using possession relations identified in tweets consisting of text and images. We do not describe these examples in detail as they are self-explanatory.

5 Experiments and Results

In order to predict possession duration and co-possession, we experiment with Logistic Regression and a neural network ensemble including a text component and two image components. Each possession relation becomes an instance, and we create stratified training (80%) and test (20%) sets. We also reserve 20% of the training as validation set. More specifically, we build two classifiers: one for possession duration (`short`, `medium`, or `long`) and one for co-possession (`yes` or `no`).

Logistic Regression. We use the implementation by scikit-learn (Pedregosa et al., 2011), and use

bag-of-words features for the sentence at hand. Specifically, we use binary flags indicating word presence, and additional flags to indicate the word corresponding to the possessor and possessee.

Neural Network. The network architecture is similar to the one in our previous work (Chinnappa et al., 2019). It includes a text component and an image component (Table 4). The latter component is disabled if no image is available.

The text component is an LSTM that takes as input the sentence (or tweet) containing the possessee. Words are represented with the concatenation of their 300-dimensional GloVe embedding (Pennington et al., 2014) and an additional embedding indicating whether a token is the possessor, possessee, or neither. We train the additional embeddings from scratch with the rest of the network.

The image component uses two pretrained neural networks. First, we concatenate to the softmax output layer the weights from the average pooling layer (second to last layer) of Inception-Net (Szegedy et al., 2015). Second, we obtain the top 5 tags from the Google Cloud Vision API and incorporate them as an additional textual input.
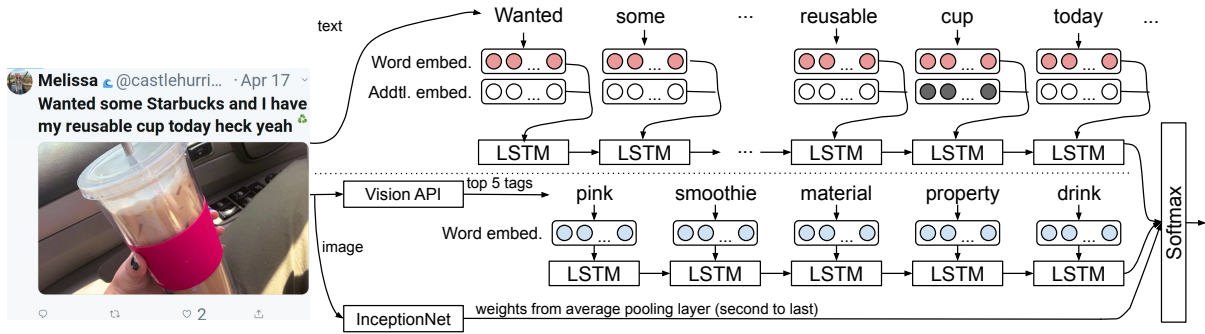
Figure 4: Neural network architecture to predict possession duration and co-possession. We include a text component (above dotted line) and two image components (below dotted line). Note that the top 5 tags from the Vision API become a textual input, and we use pretrained word embeddings and an LSTM for them.

	Majority Baseline			Log. Regression			LSTM$_{word\ embeddings}$			LSTM$_{+addtl.\ embeds.}$		
	P	R	F	P	R	F	P	R	F	P	R	F
short	0.00	0.00	0.00	1.00	0.35	0.52	0.88	0.35	0.50	0.75	0.60	0.67
medium	0.00	0.00	0.00	0.68	0.35	0.46	0.64	0.21	0.32	0.68	0.49	0.57
long	0.73	1.00	0.84	0.81	0.97	0.88	0.78	0.96	0.86	0.94	0.90	0.82
W. Avg.	0.53	0.73	0.61	0.80	0.80	0.77	0.76	0.73	0.77	0.82	0.83	0.82
yes	0.00	0.00	0.00	0.62	0.58	0.60	0.75	0.58	0.65	0.72	0.62	0.67
no	0.56	1.00	0.72	0.69	0.72	0.70	0.72	0.85	0.78	0.73	0.82	0.77
W. Avg	0.31	0.56	0.40	0.66	0.66	0.66	0.73	0.73	0.72	0.73	0.73	0.73

Table 5: Results obtained with possession relations identified from text (OntoNotes and weblogs). Addtl. embeddings refers to the embeddings indicating whether a token is the possessor, the possessee, or neither one.

More specifically, we use GloVe embeddings and an LSTM to process the additional textual input. Note that individual tags identified in the image are sometimes multiple tokens (e.g., coffee mug), so an LSTM is a good choice.

We use the implementation by Keras (Chollet et al., 2015) with TensorFlow backend (Abadi et al., 2015). More specifically, we use the Adam optimizer (Kingma and Ba, 2014) and categorical cross entropy as a loss function. We use batch size 32 for up to 200 epochs, but stop earlier if there is no improvements in the validation for 5 epochs.

5.1 Results

Table 5 presents the results with instances including only text. Regarding possession duration, the majority baseline (always long) obtains 0.61 F-measure. The second baseline, Logistic Regression, obtains 0.77 F-measure. These results are strong, however, Logistic Regression is biased towards the most common label (long, Table 1), and performs poorly with the other labels (short and medium). In fact, Logistic Regression outperforms LSTM$_{+addtl.\ embeds.}$ with long, but the weighted F-measure is lower (0.77 vs. 0.82). Re-

garding co-possession, we observe a similar trend, but the LSTM performs similar with both labels.

LSTM and Additional Embeddings. Table 5 presents results obtained with the LSTM using (a) only the word embeddings and (b) incorporating the additional embeddings for the possessor and possessee. The LSTM with only word embeddings obtains worse results predicting possession durations (0.77 vs. 0.82 weighted F-measure), and virtually the same results predicting co-possessions (0.72 vs. 0.73 weighted F-measure). These results lead to the conclusion that the specific possessor and possessee along with context are important to determine how long a possession holds true. On the other hand, determining whether there are concurrent co-possessors does not benefit from the specific possessor and possessee (i.e., events and other information contained in the sentence are sufficient).

Table 6 presents results with the tweets (all of them include both text and images). The results indicate that the text is vital to determine possession duration and co-possession, and that the image components do not bring any improvements. Logistic Regression obtains best results for both pos-

	Majority Baseline			Log. Regression			LSTM$_{+addtl.\ embeds}$			only image comp.			text + image		
	P	R	F	P	R	F	P	R	F	P	R	F	P	R	F
short	.00	.00	.00	.00	.00	.00	.50	.11	.18	.00	.00	.00	.12	.10	.11
medium	.00	.00	.00	.61	.57	.59	.58	.45	.50	.39	.29	.33	.52	.30	.38
long	.58	1.00	.73	.70	.78	.74	.66	.80	.73	.57	.71	.63	.64	.83	.72
W. Avg.	.33	.58	.42	.64	.67	.65	.62	.64	.62	.48	.52	.49	.57	.59	.56
yes	.00	.00	.00	.41	.29	.34	.37	.32	.34	.30	.14	.19	.33	.15	.21
no	.73	1.00	.84	.76	.85	.80	.76	.79	.78	.73	.88	.80	.74	.89	.81
W.Avg.	.53	.73	.62	.67	.70	.68	.65	.67	.66	.61	.68	.63	.63	.69	.64

Table 6: Results obtained with possession relations identified from text and image (tweets).

session duration and co-possession, and obtains similar results than the text component of the neural network (LSTM$_{+addtl\_embeddings}$): 0.65 vs. 0.62 F-measure (duration) and 0.68 vs 0.66 F-measure (co-possession). While including the image component slightly decreases the results predicting co-possession (0.66 vs. 0.64 F-measure), it heavily decreases results predicting possession duration (0.62 vs. 0.56 F-measure). We attribute these unexpected results to the nature of the tasks. Image tags provide high-level information about the possessee (e.g., cup), and determining possession durations and co-possessions require fine-grained information about the possessee (e.g., reusable, disposable) as well as knowledge about the events that connect the possessor and possessee.

6 Conclusions

Standard relation extraction does not provide information about for how long relations hold true or whether relations are one-to-one or one-to-many. In this paper, we tackle both problems and determine possession durations and co-possessions. Possessions are ubiquitous yet understudied from a computational perspective. From a theoretical perspective, they include having control over something (e.g. flying a plane, impounding a vehicle, eating ice cream) thus most objects are actually possessees of one or more possessors. Additionally, as just exemplified, many possessions can be extracted even if prototypical possession verbs (e.g., have, buy, acquire) are missing.

We have presented new annotations on top of existing corpora. Regarding durations, we collect lower and upper bounds in order to derive sound duration intervals. The resulting three intervals obtain substantial agreement (0.63 Cohen's κ). Regarding co-possessions, we obtain slightly better agreement (0.65 Cohen's κ). We have also presented baseline models and a neural network

architecture to solve both tasks. Beyond word embeddings, the LSTM benefits from additional embeddings indicating the tokens that are the possessor and possessee. Information extracted from the image, however, is not helpful.

While the work presented here targets possession relations, we believe that a similar approach could be used to to determine for how long any semantic relation holds true.

References

Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. 2015. TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org.

Omri Abend and Ari Rappoport. 2017. The state of the art in semantic representation. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 77–89, Vancouver, Canada. Association for Computational Linguistics.

James F Allen and George Ferguson. 1994. Actions and events in interval temporal logic. *Journal of logic and computation*, 4(5):531–579.

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 and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 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.

Adriana Badulescu and Dan Moldovan. 2009. A semantic scattering model for the automatic interpretation of english genitives. *NLE*.

Carmen Banea and Rada Mihalcea. 2018. Possession identification in text. *Natural Language Engineering*, 24(4):589610.

Austin Blodgett and Nathan Schneider. 2018. Semantic Supersenses for English Possessives. In *LREC*.

Dhivya Chinnappa and Eduardo Blanco. 2018. Mining possessions: Existence, type and temporal anchors. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 496–505, New Orleans, Louisiana, USA. Association for Computational Linguistics.

Dhivya Chinnappa, Srikala Murugan, and Eduardo Blanco. 2019. Extracting possessions from social media: Images complement language. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 663–672, Hong Kong, China. Association for Computational Linguistics.

François Chollet et al. 2015. Keras. https://github.com/fchollet/keras.

Heng Ji, Ralph Grishman, and Hoa Dang. 2011. Overview of the TAC2011 knowledge base population track. In *TAC 2011 Proceedings Papers*.

Diederik P. Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980.

Parisa Kordjamshidi, Archna Bhatia, James Pustejovsky, and Marie-Francine Moens. 2018. Proceedings of the first international workshop on spatial language understanding. New Orleans. Association for Computational Linguistics.

David McClosky and Christopher D. Manning. 2012. Learning constraints for consistent timeline extraction. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 873–882, Jeju Island, Korea. Association for Computational Linguistics.

Bill McDowell, Nathanael Chambers, Alexander Ororbia II, and David Reitter. 2017. Event ordering with a generalized model for sieve prediction ranking. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1:*

Long Papers), pages 843–853, Taipei, Taiwan. Asian Federation of Natural Language Processing.

Preslav I. Nakov and Marti A. Hearst. 2013. Semantic interpretation of noun compounds using verbal and other paraphrases. *ACM Trans. Speech Lang. Process.*, 10(3):13:1–13:51.

Feng Pan, Rutu Mulkar-Mehta, and Jerry R. Hobbs. 2011. Annotating and learning event durations in text. *Computational Linguistics*, 37(4):727–752.

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.

Jeffrey Pennington, Richard Socher, and Christopher D. Manning. 2014. Glove: Global vectors for word representation. In *Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543.

Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. 2015. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9.

Shiao Wei Tham. 2004. *Representing Possessive Predication: Semantic Dimensions and Pragmatic Bases*. Ph.D. thesis, Stanford University.

Stephen Tratz and Eduard Hovy. 2010. A taxonomy, dataset, and classifier for automatic noun compound interpretation. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, ACL '10, pages 678–687, Stroudsburg, PA, USA. Association for Computational Linguistics.

Stephen Tratz and Eduard H. Hovy. 2013. Automatic interpretation of the english possessive. In *ACL (1)*, pages 372–381. The Association for Computer Linguistics.