Structural Sentiment and Entity Annotation Guidelines

Miriam Eckert, Lyndsie Clark, Hanna Lind,
Jason Kessler and Nicolas Nicolov

October 19, 2009

Contents

1 Note

If you wish to cite the JDPA Sentiment Corpus, please cite this paper:

```
Jason S. Kessler, Miriam Eckert, Lyndsie Clark, and Nicolas Nicolov.  The ICWSM 2010 JDPA
Sentiment Corpus for the Automotive Domain. 4th International AAAI Conference on Weblogs
and Social Media Data Challange Workshop (ICWSM-DCW 2010), 2010. Washington, D.C.
```

2 Introduction

Structural sentiment annotation is an attempt to capture the fine-grained details of entities, sentiment-bearing elements and how they are related to each other. The corpus is annotated for Entity Mentions, coreference and meronymy, in the hope that it can be used for building systems that deal with all three of these levels.

The JDPA Corpus consists of user-generated content (blog posts) containing opinions about automobiles and digital cameras. They have been manually annotated for named, nominal, and pronominal mentions of entities. Entities are marked with the aggregate sentiment expressed toward them in the document. Mentions of each entity are marked as co-referential. Mentions are assigned semantic types consisting of the Automatic Content Extraction (ACE) mention types and additional domain-specific types. Meronymy (part-of and feature-of) and instance relations are also annotated. Expressions which convey sentiment toward an entity are annotated with the polarity of their prior and contextual sentiments as well the mentions they target. The following modifiers are annotated. These may target other modifiers or sentiment expressions

- negators (expressions which invert the polarity of a sentiment expression or modifier)

- neutralizers (expressions that do not commit the the speaker to the truth of the target sentiment expression or modifier)

- committers (expressions which shift the commitment of the speaker toward the truth a sentiment expression or modifier)

- intensifiers (expressions which shift the intensity of a sentiment expression or modifier)

Additionally, we have annotated when the opinion holder of a sentiment expression is someone other than the author of the blog by linking the expression to the holder. We also annotate when two entities are compared on a particular dimension.

Where possible, we have attempted to conform to currently existing guidelines, such as the annotation guidelines for the ACE corpus. References in the text indicate passages from other sources.

3 Non-Domain Specific Entity Types

A Mention is any mention of an entity that fits into the definition of one of the Entity Types outlined herein. Mentions can occur in the form of proper names, noun phrases, pronouns (e.g., *I, you, he, she, it*), possessive pronouns (e.g., *my, your, his, mine, hers*) and demonstratives (e.g., *this, that, those, these*) occurring by themselves.

(1)   It has been a bit but on my recent trip to Virginia I rented a new (only 1000 miles on the clock) 2009 Toyota Corolla. Car came with a cloth interior. This was a pretty base model car but it had some nice features.

Mentions can also occur in the form of verbs. This happens mainly when a reference is made to a car feature, e.g.

(2)   I can't believe how smooth they run!

(a) **Mention:** run
EntityType: CarPart
FeatureOf: they

NOTE: Tag only Mentions if they have some sentiment related to them or if they fit into one of the predetermined categories. Don't use the general Mention tag unless it has some sentiment attached to it.

The following sections describe the Entity Types in more detail.

3.1   Person

Label all mentions of persons: this includes generic noun phrases (guy, salesperson etc.), specific names (John Smith, Mary etc) and pronouns (*I, me, you, we, us, he, him, she, her, they, them, their* etc.).

(3)   Prime Minister <u>Kevin Rudd</u> made the announcement here at Toyota's global corporate headquarters today at a joint media conference with Toyota president <u>Katsuaki Watanabe</u>.

   (a) **Mention:** Kevin Rudd
            EntityType: Person

   (b) **Mention:** Katsuaki Watanabe
            EntityType: Person

(4)   <u>We</u> bought a car with our first paycheck.

   (a) **Mention:** We
            EntityType: Person

(5)   A much better system would be if <u>car owners</u> had an easy way of turning into part time <u>cabbies</u>.

   (a) **Mention:** car owners
            EntityType: Person

   (b) **Mention:** cabbies
            EntityType: Person

Also select and label pronouns that refer to the same entity as a preceding Person Mention:

(6)   <u>John Smith</u> walked into the store and <u>he</u> bought a car.

   (a) **Mention:** John Smith
            EntityType: Person

   (b) **Mention:** he
            EntityType: Person
            RefersTo: John Smith

Label first and second person pronouns *I, you , we*. Also, label pronouns that are not coreferring with other noun phrases, such as *he* in Example 8.

(7)   <u>My</u> <u>parents</u> gave <u>me</u> <u>their</u> old car. the 'new' car is an'88 toyota camry <u>i</u> pitched 450 in to get a new engine put in.

   (a) **Mention:** my
            EntityType: Person

   (b) **Mention:** parents
            EntityType: Person

   (c) **Mention:** me
            EntityType: Person
            RefersTo: my

   (c) **Mention:** their
            EntityType: Person
            RefersTo: parents

(8)   Look at the larger sized pic. I think he's confused.

   (a) **Mention:** I
         EntityType: Person

   (b) **Mention:** he
         EntityType: Person

Fictional characters are also labeled as Person, e.g.,

(9)   It was in that year that I first met Lucifer and was transported into Hell.

   (a) **Mention:** I
         EntityType: Person

   (b) **Mention:** Lucifer
         EntityType: Person

(10) Disney's 1937 animated short film, "Modern Inventions," opens with Donald Duck entering a Museum of Modern Marvels.

   (a) **Mention:** Donald Duck
         EntityType: Person

Mark any occurences of *someone, anyone, no-one, somebody, anybody* and *nobody* as Person Mentions:

(11) Someone drove the car into the garage.

   (a) **Mention:** Someone
         EntityType: Person

3.2   Organization

An Organization entity must have some formally established association. Typical examples are businesses, government units, sports teams, and formally organized music groups. Industrial sectors and industries are also treated as Organization entities. (LDC, 2005) For our purposes, certain mentions of websites can also be marked as ORG.

(12) At the Canadian Grand Prix, Timo Glock enjoyed the strongest finish of his Formula One career so far, crossing the line in fourth to clinch five points for Toyota.

   (a) **Mention:** Toyota
         EntityType: Organization

Care should be taken to distinguish the mention of a business from the mention of a product. For example, Toyota can refer to the company as in Example 12 above, or it can refer to one of its products as in Example 13 below, where it should be marked once as VEH and again as ORG.

(13) My Toyota is very reliable.

   (a) **Mention:** Toyota
         EntityType: Organization

   (b) **Mention:** Toyota
         EntityType: Vehicle

Also, a single department of an organization should be annotated as its own organization and then PartOf the larger organization (if mentioned) as illustrated in 14 below.

(14) I am the vice president of Strategic Planning and Product Development at Hyundai Motor America.

    (a) **Mention:** Hyundai Motor America
          EntityType: Organization

    (b) **Mention:** Strategic Planning and Product Development
          EntityType: Organization
          PartOf: Hyundai Motor America

Sometimes the pronouns *they, them* or *their* are used to indirectly refer to a company or organization, as in below. Such examples of *corporate they* should also be labeled as ORG. However, if the context of a mention makes it clear that the referent is a specific individual or group of individuals, label that mention as Person, and then PartOf the organization mention, as illustrated in 17 below.

(15) Good for <u>Toyota</u>. <u>They</u> are producing something people want, and will make a bundle.

    (a) **Mention:** Toyota
          EntityType: Organization

    (b) **Mention:** They
          EntityType: Organization
          RefersTo: Toyota

(16) If you are, you might want to take a look at what <u>All About Luxury Cars</u> has to offer. <u>They</u> are not selling luxury cars but <u>they</u> provide a lot of information in relation to the top luxury cars that <u>many</u> people dream about.

    (a) **Mention:** All About Luxury Cars
          EntityType: Organization

    (b) **Mention:** They
          EntityType: Organization
          RefersTo: All About Luxury Cars

    (c) **Mention:** They
          EntityType: Organization
          RefersTo: All About Luxury Cars

(17) The prices on <u>their</u> cameras were competitive and <u>they</u> were offering to pay the sales tax on all purchases. Plus <u>they</u> threw in all kinds of Sony shwag including t-shirts, hats, 2 flashlights, a couple of caribiners.

    (a) **Mention:** their
          EntityType: Organization

    (b) **Mention:** they
          EntityType: Organization
          RefersTo: their

    (c) **Mention:** they
          EntityType: Person
          PartOf: they

In Example 18, the *Navy Electronic Commerce Online website* is treated like an company in this context (i.e. it is involved in business negotiations) and should be labeled as an Organization.

(18) This contract was competitively negotiated via the <u>Navy Electronic Commerce Online website</u> with six proposals received.

    (a) **Mention:** Navy Electronic Commerce
          EntityType: Organization

3.3 Location

Places defined on a geographical or astronomical basis which are mentioned in a document and do not constitute a political entity give rise to Location entities. These include, for example, the solar system, Mars, the Hudson River, Mt. Everest, and Death Valley. (LDC, 2005) Also tag the generic words 'place' and 'location'.

(19) This email was gushing over the 1.8 trillion barrels of shale oil believed to be underneath the <u>Green River Basin</u>, which stretches across Colorado, Utah and Wyoming.

    (a) **Mention:** Green River Basin
            EntityType: Location

3.4 Facility

A facility is a functional, primarily man-made structure. These include buildings and similar facilities designed for human habitation, such as houses, factories, stadiums, office buildings, gymnasiums, prisons, museums, and space stations; objects of similar size designed for storage, such as barns, parking garages and airplane hangars; elements of transportation infrastructure, including streets, highways, airports, ports, train stations, bridges, and tunnels. Roughly speaking, facilities are artifacts falling under the domains of architecture and civil engineering. For our purposes, mentions of specific websites will also be treated as facilities. Unlike example 18 seen in the Organization section, only mentions ending in '.com' will be annotated as facilities, and possibly double-annotated as an Organization, as seen in 22 (LDC, 2005)

(20) driving along the <u>FDR highway</u>

    (a) **Mention:** FDR highway
            EntityType: Facility

(21) Also had a few drinks in the <u>airport</u> this weekend, which was nice

    (a) **Mention:** airport
            EntityType: Facility

(22) I did not buy it at the store, but purchased it from <u>Amazon.com</u> last night and saved about $200 since it was a little cheaper online and no sales tax.

    (a) **Mention:** Amazon.com
            EntityType: Facility

    (b) **Mention:** Amazon.com
            EntityType: Organization

3.5 Geo-Political Entities

Geo-Political Entities are composite entities comprised of a population, a government, a physical location, and a nation (or province, state, county, city, etc.). (LDC, 2005) We also include mentions of continents (e.g., South America) or groups of geo-political entities (e.g., Eastern Europe) in this entity type. Note: In a phrase such as *'We Americans think gas prices are too high.'* where the only person mention is 'Americans', please tag this word as Mention.Person rather than GeoPolitical.Nationality.

    The GEOPOLITICAL subtypes are: CITY, NATIONALITY, COUNTRY, and U.S. STATE. All other types should simply be labeled just as GEOPOLITICAL.

(23) They went to <u>Paris</u> last summer.

    (a) **Mention:** Paris
            EntityType: GeoPolitical.City

(24) This email was gushing over the 1.8 trillion barrels of shale oil believed to be underneath the Green River Basin, which stretches across <u>Colorado</u>, <u>Utah</u> and <u>Wyoming</u>.

(a) **Mention:** Colorado
   EntityType: GeoPolitical.State

(b) **Mention:** Utah
   EntityType: GeoPolitical.State

(c) **Mention:** Wyoming
   EntityType: GeoPolitical.State

(25) I'm slinging from <u>Los Angeles</u> to <u>Eastern Europe</u> and getting about 300-350K on my Slingbox.

(a) **Mention:** Los Angeles
   EntityType: GeoPolitical.City

(a) **Mention:** Eastern Europe
   EntityType: GeoPolitical.Nationality

(26) The <u>U.N.</u> could send more help to war-torn provinces.

(a) **Mention:** U.N.
   EntityType: GeoPolitical

Note that geo-political entities in adjectival form should also be annotated as GEO:

(27) After a long journey Isabel has finally made it to <u>English</u> soil.

(a) **Mention:** English
   EntityType: GeoPolitical.Country

3.6 Time

Label mentions of absolute times and dates, i.e. times and dates that are always interpreted in the same way, regardless of context. Note that in Example 28 below, the possessive pronoun *my* is not included in the span of the Time entity. Also label mentions of relative times and dates, i.e. times and dates that are interpreted relative to the time of speaking/writing (e.g., *yesterday, today*) but do not label the words 'time' or 'date'.

The Time subtypes are DURATION, DAYS OF THE WEEK, MONTH, YEAR, DATE and OCLOCK. All other time entities should be simply labeled as TIME.

(28) My <u>birthday</u> is <u>12.3.1987</u>.

(a) **Mention:** birthday
   EntityType: Time

(b) **Mention:** 12.3.1987
   EntityType: Time.Date
   RefersTo: birthday

(29) Inflation is the talking point this week, and that theme will play again in <u>Wednesday</u>'s stock market when Fed Vice Chairman Donald Kohn weighs in on the topic.

(a) **Mention:** Wednesday
   EntityType: Time.Day

(30) Where will you be <u>12 hours from now</u>?

(a) **Mention:** 12 hours from now
   EntityType: Time

(31) I went to that store <u>last year</u> to get a TV

(a) **Mention:** last year
      EntityType: Time.Year

(32) He doesn't want to sleep <u>now</u> because he's not tired

(a) **Mention:** now
      EntityType: Time

(33) Last I heard they were out there <u>this morning</u> mowing the lawn <u>for two hours</u>.

(a) **Mention:** this morning
      EntityType: Time

(b) **Mention:** for two hours
      EntityType: Time.Duration

(34) They were out there <u>early Saturday morning</u> mowing the lawn

(a) **Mention:** early Saturday morning
      EntityType: Time

Do not label embedded time expressions separately. In the following example, only the date mention as a whole is labeled as a Time Mention. We do not label the month *July* or the year *1774* separately.

(35) <u>July 4, 1776</u> is an important date in history.

(a) **Mention:** July 4, 1774
      EntityType: Time.Date

Label year mentions that are part of a car description, e.g.,

(36) My sweet <u>'85</u> Olds Calais

(a) **Mention:** '85
      EntityType: Time.Year

3.7   Units

Label units that are described in the categories listed below. Do not label basic quantifiers such as '*one* person' or '*twenty* people.'

NOTE: Certain units that are especially relevant to the domains of interest will be labeled Mention.Feature and then FeatureOf the appropriate product mention; for further clarification see 37 below. (More information below in 'Distinguishing Units from Features' section)

(37) The new Mustang has an engine that gives *250 horsepower* at *3,000rpm.*

(a) **Mention:** 250 horsepower
      EntityType: CarFeature

(b) **Mention:** 3,000rpm
      EntityType: CarFeature

(c) **Mention:** horsepower
      EntityType: Unit

3.7.1 Money

Label all mentions of prices, fees and rates.Example 38 shows that the text span should include any peripheral elements that are necessary for identifying the exact amount (i.e. *just under $40000* is not the same as *$40000*).The monetary unit signs, s uch as the dollar or cent sign, should also be included in the span.

(38) you can get into a Corolla for <u>around about $22000</u> whereas a Prius costs you <u>just under $40000</u>.

     (a) **Mention:** around about $22000
           EntityType: Unit.Money

     (b) **Mention:** just under $40000
           EntityType: Unit.Money

3.7.2 Speed

Label all mentions of vehicle speed.

(39) he cruises the highways of my limbic system at <u>90 miles an hour</u>.

     (a) **Mention:** 90 miles an hour
           EntityType: Unit.Speed

3.7.3 Rate

Label all mentions of payment rates.

(40) Way longer than I originally anticipated, but I also originally anticipated not having a <u>$800 a month</u> car loan to repay at the same time.

     (a) **Mention:** $800 a month
           EntityType: Unit

3.7.4 Age

Label all mentions of age of a person or vehicle.

(41) Of course, that works out to a maximum of four full-sized adults plus two kids <u>no older than about age 10</u>.

     (a) **Mention:** no older than about age 10
           EntityType: Unit.Age

3.7.5 Distinguishing Units from Features

- Units and rates that are intrinsically related to the quality of the product should be considered **Mention.Feature**. These are units for which a very high or low instance of that unit or rate could be nearly universally agreed upon as contributing to a product being of very high or low quality. For example, the sensor size in the camera domain or the rate miles per gallon in the car domain are directly related to the quality of the product. Some additional examples of units that are treated as features are: pixel counts ('10 megapixels'), ISO ('200,400,800') and shutter speed ('3 frames per second')in the camera domain, and mileage ('120,000 miles'), horsepower ('178 hp'), and torque ('289lb/ft') in the car domain.

- Units and rates that are not intrinsically related to the quality of the product should be considered **Mention.Unit**. For example, price and dimensions may have an impact on whether or not a particular consumer chooses to buy a certain camera, but there is no intrinsic relation between price/dimensions and whether or not a camera is high or low quality. They should therefore be marked as Mention.Unit.

(42) In contrast, the (amazing) Canon <u>85mm</u> <u>f/1.2</u> is about five-and-a-half-times the price.

(a) **Mention:** 85mm
   EntityType: Mention.Feature
   FeatureOf: Canon 85mm f/1.2

(b) **Mention:** f/1.2
   EntityType: Mention.Feature
   FeatureOf: Canon 85mm f/1.2

(43) The new Hyundai i800 has a fuel efficiency of <u>33.2 mpg</u>.

(a) **Mention:** 33.2 mpg
   EntityType: Mention.Feature
   FeatureOf: fuel efficiency

3.8   Descriptor

Label as 'Descriptor' adjectives preceding CarPart, CarFeature, CameraFeature, CameraPart, CameraAccessory or other head nouns (Entity Mentions). These words give more detailed information about the Mention, further narrowing the sentiment to a specific item in the larger group of Mentions. While they are not technically Mentions, they currently have an EMLevel slot that still needs to be filled. This may change in the future. In example 44, '8-way', 'power', and 'heated' describe the characteristics of the seat, distinguishing them from other car seats that do not have these characteristics. Sentiment Bearing Expressions, however, should still target the head noun as it is the overall receiver of the sentiment.

   The 'Describes' slot should be filled with Mentions targed by the Descriptor.

(44) The Hyundai Santa Fe has wonderful <u>8-way</u>, <u>power</u>, <u>heated</u> seats.

(a) **Mention:** 8-way
   EntityType: Descriptor
   Describes: seats

(b) **Mention:** power
   EntityType: Descriptor
   Describes: seats

(c) **Mention:** heated
   EntityType: Descriptor
   Describes: seats

(d) **SentimentBearingExpression:** wonderful
   PriorPolarity: Positive
   Target: seats

3.8.1   Distinguishing Descriptors from Entity Mentions

If a word evokes a discourse entity (for example, a pronoun) or has the ability to co-refer with another nominal or named entity, then it should be marked as a *part* or *feature*, as in example 45. However, if it modifies a Mention (with the exception of a Descriptor which cannot be modified), it should be marked as a Descriptor. This can be seen in example 46 as 'they' cannot refer to '4-door' since '4-door' is describing the type of sedan.

(45) My sedan has 4 <u>doors</u>. <u>They</u> close quietly.

(a) **Mention:** doors
   EntityType: CarPart
   PartOf: sedan

(b) **Mention:** They
   EntityType: CarPart
   RefersTo: doors

(46) ?I have a <u>4-door</u> sedan. <u>They</u> close quietly.

    (a) **Mention:** 4-door
           EntityType: Descriptor
           Describes: sedan

    (b) **Mention:** They
           EntityType: CarPart
           RefersTo: ?

Following are some diagnostic tests that help in determining if a compound noun should be 1 span therefore annotated as an Entity Mention (CarPart, CarFeature etc.) or 2+ spans where the head noun (usually the final word) would be an Entity Mention and the preceding ones Descriptors.

- To determine if a Mention is a Descriptor or CarPart, put the word into the sentence: *'I removed/bought/installed/broke the <u>windshield</u>.'* If the word fits into this construction, mark it as a CarPart.

- The 'made of' relationship occurs when the head noun is made of the describing noun(s). These cases would be marked as Descriptor(s) + Entity Mention: <u>leather</u> <u>seats</u> - the seats are *made of leather.*

- The 'for' relationship occurs when the head noun is used for the describing noun(s). These cases would be marked as 1 Entity Mention span. Ex). <u>camera bag</u>, <u>steering wheel</u> - the former is a bag *for* the camera and the latter is a wheel *for* steering.

- If the word is in pre-nominal position but is not part of a noun-noun compound then it is a Descriptor.

- If the word is a size/length/weight descripton then it is a Descriptor, not a Unit. Ex). *The <u>19-inch</u> wheels.*

- Some words can be Descriptors or CarFeature/CameraFeature depending where they are in the sentence. If they are pre-nominal, then mark as Descriptor, but if they follow a copula or verb like *be, has, get, etc.* then they would be marked as Feature. Ex). *The <u>250HP</u> Mustang* vs. *My Mustang gets <u>250-horespower</u>.* The former would be Descriptor where the latter

3.9 Entity Mention Text Spans

Please see the section 'Text Spans' below for other rules on how to determine the span of an Entity Mention.

4 Car Domain Entity Types

The following entity types are a part of the car domain and have been selected for their relevance in this domain. The Mention types included here are Mention.Vehicle (with sublevels of Car, SUV and Truck), Mention.CarPart and Mention.CarFeature.

4.1 Vehicle

Label mentions of any vehicles that fall under the categories car, SUV or truck as Entities of their respective types. Also label pronouns that are coreferring with such Entities as well as the generic words 'car', 'truck' etc. If the subcategory of the vehicle entity is not clear, either because it is underspecified (e.g., 'van') or because the annotator does not know the exact category and cannot attain this information through further research, then it should be labeled simply as Vehicle.

Note that with mentions of car models the organization EntityType should be labeled separately if part of the car model name. For example, in *Toyota Camry*, *Toyota* should also be labeled as EntityType Organization. Also, label any year mentions both as part of the car name but also separately as Time.Year (e.g., '89). Be sure to include the apostrophe when labeling these.

Also note: If the vehicle being discussed is referred to by the author as a different type than it is (i.e. a Jeep Grand Cherokee being called a 'car' even though it's actually an SUV) mark the contradictory term as the actual type of vehicle (i.e. SUV), not the referred-to type (car).

4.1.1   Cars

Label as 'car' all sedans, hatchbacks, station wagons, coupes, and convertibles as well as crossover vehicles. Some examples of crossover vehicles are Toyota Matrix, Pontiac Vibe and Dodge Caliber, however some crossovers are considered SUVs so please see the Kelley Blue Book SUVs (R:/SCIENCE/SENTIMENT/JDPA-corpus/doc/Kelley Blue Book SUV List.xls list if you are unsure.

(47)   Now, mind you, I do not keep my <u>car</u> spotlessly clean - not by any stretch of the imagination.

    (a) **Mention:** car
           EntityType: Veh.Car

(48)   I've been driving <u>pre-1986 Mercedes turbo diesels</u>, converted to WVO using Lovecraft systems, for about 3 years now. While <u>they</u>'ve been great and it's been a total success, I'm not really all that willing to buy another <u>one</u>.

    (a) **Mention:** pre-1986 Mercedes turbo diesels
           EntityType: Veh.Car

    (b) **Mention:** Mercedes
           EntityType: Organization

    (c) **Mention:** they
           EntityType: Veh.Car
           RefersTo: pre-1986 Mercedes turbo diesels

    (d) **Mention:** one
           EntityType: Veh.Car
           MentionLevel: Pronominal
           RefersTo: pre-1986 Mercedes turbo diesels

    (e) **Mention:** pre-1986
           EntityType: Time

(49)   She didn't seem proud that she drove a <u>Lexus</u> (after I had told her about my sick <u>'89 gold camry wagon</u>) .

    (a) **Mention:** Lexus
           EntityType: Veh.Car

    (b) **Mention:** Lexus
           EntityType: Organization

    (c) **Mention:** '89 gold camry wagon
           EntityType: Veh.Car

    (d) **Mention:** '89
           EntityType: Time.Year

(50)   my tiny gladiator ego also drive a <u>porsche</u> and has a ponytail and a triangle shaped goatee

    (a) **Mention:** porsche
           EntityType: Veh.Car

    (b) **Mention:** porsche
           EntityType: Organization

(51)   Imagine that the perfect <u>car</u> - always ready to go.

    (a) **Mention:** car
           EntityType: Veh.Car

NOTE: If a specific year of a car is mentioned, include the year in the Mention.Car span but also tag it separately as the appropriate Mention.Time:

(52) I'm looking to buy a 1994-1998 Suzuki Swift or a 1994 Toyota Corolla.

     (a) **Mention:** 1994-1998 Suzuki Swift
            EntityType: Veh.Car

     (b) **Mention:** 1994-1998
            EntityType: Time.Duration

     (c) **Mention:** 1994 Toyota Corolla
            EntityType: Veh.Car

     (d) **Mention:** 1994
            EntityType: Time.Year

4.1.2 SUVs

Label as 'SUV' any large vehicle that is not a truck or car. Please see the Kelley Blue Book SUVs list if you are unsure. Some examples of SUVs are Dodge Durango, Jeep Compass, and Ford Explorer.

(53) Wired's Autopia blog documents what we all knew was coming: rising gas prices have killed off the SUV. Auto industry watchers had predicted that the gas guzzlers in the 'light truck' category would lose the ascendancy by 2010; no one expected their reign to end in a month,

     (a) **Mention:** SUV
            EntityType: Veh.SUV

     (b) **Mention:** gas guzzlers in the light truck category
            EntityType: Veh.SUV
            RefersTo: SUV

     (c) **Mention:** their
            EntityType: Veh.SUV
            RefersTo: SUV

(54) The Chevy Tahoe hybrid may have won the Green Car of the Year award, and the Chicago Police is considering a switch to the vehicle for policing

     (a) **Mention:** Chevy Tahoe hybrid
            EntityType: Veh.SUV

     (b) **Mention:** Chevy
            EntityType: Organization

     (c) **Mention:** vehicle
            EntityType: Veh.SUV
            RefersTo: Chevy Tahoe hybrid

4.1.3 Trucks

Label as 'truck' all pick-ups and heavy lifting trucks. Some examples of these are: Ford F150, Dodge Ram and Toyota Tacoma.

(55) How do you replace the glass on a 2004 Ford F150 driver side mirror?

     (a) **Mention:** 2004 Ford F150
            EntityType: Veh.Truck

     (b) **Mention:** Ford
            EntityType: Organization

4.2   CarPart

Mentions of concrete car parts should be labeled separately. This refers to entities that a person can touch, hold or otherwise physically or visualy perceive.

(56) It has that hard, nasty, cheap plastic they've chosen for the <u>steering wheel</u> and <u>dashboard</u>.

    (a) **Mention:** steering wheel
        EntityType: CarPart
        PartOf: it

    (b) **Mention:** dashboard
        EntityType: CarPart
        PartOf: it

    (c) **Mention:** it
        EntityType: Car

Also, if a car, truck or SUV is referred to by a part of its construction, tag it as both CarPart and Vehicle (car, truck, SUV if known) and then make the CarPart instance a PartOf the Veh.car instance as in Example 57.

(57) My <u>4-cylinder</u> gets good gas mileage.

    (a) **Mention:** 4-cylinder
        EntityType: Car

    (b) **Mention:** 4-cylinder
        EntityType: CarPart
        PartOf: 4-cylinder

4.3   CarFeature

Features of cars should also be labled separately. These are entities that are not concrete or physical such as handling, horsepower, space (in the cab, trunk, etc.), miles per gallon, and torque.

(58) The Toyota Corolla XRS puts out <u>158 horsepower</u> and <u>162 lb-ft of torque</u>.

    (a) **Mention:** Toyota Corolla XRS
        EntityType: Car

    (b) **Mention:** 158 horsepower
        EntityType: CarFeature
        FeatureOf: Toyota Corolla XRS

    (c) **Mention:** 162 lb-ft of torque
        EntityType: CarFeature
        FeatureOf: Toyota Corolla XRS

Expressions that are not nouns can also be marked as CarFeature if they refer to abstract features of the car that people frequently express opinions about. Often verbs are used to for such features, e.g., *runs, works, handles, sounds*:

(59) I can't believe how smooth they <u>run</u>!

    (a) **Mention:** run
        EntityType: CarFeature
        FeatureOf: they

(60) The car <u>handled</u> well and felt neutral.

(a) **Mention:** handled
        EntityType: CarFeature
        FeatureOf: car

(b) **Sentiment Expression:** well
        Target: handled
        Polarity: positive

(61) The car <u>sounds</u> even more noisier.

  (a) **Mention:** sounds
        EntityType: CarFeature
        FeatureOf: car

  (b) **Sentiment Expression:** noisier
        Target: sounds
        Polarity: negative

Unlike for noun CarFeatures, these verbs should only be annotated if some sentiment is being expressed around them. Do NOT annotate expressions like *I <u>drove</u> my car down the highway.* where a factual statement is being made with no sentiment expressed.

5 Camera Domain Entity Types

This domain refers specifically to the camera market and includes relevant aspects of cameras including Mention.Camera, Mention.CameraPart, Mention.CameraFeature and Mention.CameraAccessory.

5.1 Camera

Label as 'camera' all digital, film, SLR, DSLR, and point-and-shoot cameras. Be sure to mark the whole span be it a camera name like 'Canon EOS 40D camera' or a type of camera like 'digital SLR camera'. Mark the word 'camera' everywhere it occurs even in: *camera bag, camera case and camera store.*

(62) The <u>Nikon Coolpix L11</u>, a 6-megapixel <u>point-and-shoot camera</u>, produces very nice pictures.

  (a) **Mention:** Nikon Coolpix L11
        EntityType: Camera
        MentionLevel: Named
        InstanceOf: point-and-shoot camera

  (b) **Mention:** point-and-shoot camera
        EntityType: Camera
        MentionLevel: Nominal

The span 'camera kit' should be marked as Mention.Camera. This is used to refer to both the camera body and an included lens, usually of DSLR or SLR cameras. If there is another mention of the camera itself that is included in the kit, mark the camera mention as Mention.Camera, and PartOf the 'camera kit' mention. In example 63 they are not coreferring.

IMPORTANT: Do not make 'camera kit' and a specific camera mention ('Nikon D70') co-referring unless you're sure they're identical (i.e. you're sure that 'Nikon D70' refers to the entire kit, instead of just the body.

(63) I got a <u>camera kit</u> for under 1000 dollars. It included a <u>bag</u> and a <u>lens</u>. The <u>camera</u> itself was high quality.

  (a) **Mention:** camera kit
        EntityType: Camera

  (b) **Mention:** bag
        EntityType: CameraAccessory

     (c) **Mention:** lens
          EntityType: CameraPart
          PartOf: camera kit

     (d) **Mention:** camera
          EntityType: Camera
          PartOf: camera kit

5.2 CameraPart

Mentions of concrete camera parts should be labeled separately. This refers to entities that a person can touch, hold or otherwise physically or visually perceive including body, lens, power button, and LCD screen.

(64) The Sony V1 has no flip out <u>LCD screen</u>.

     (a) **Mention:** LCD screen
          EntityType: CameraPart
          MentionLevel: Nominal
          PartOf: Canon Powershot A590IS

A CameraPart is different from an CameraAccessory in that a part is something that the camera requires to be able to function and perform the way the manufacturer intended. Thus, even if a post lists several different types of lenses, they should all be marked as CameraParts because the camera needs at least one lens to be able to take pictures. However, if it is clear that a part was purchased separately from the camera, do not annotate the mention of that part as PartOf the camera mention.

5.3 CameraFeature

Features of cameras should also be labeled separately. These are entities that are not concrete or physical such as photo/image, software, ISO, image stabilization, aperture priority mode and autoflash.

(65) My Canon Powershot A590IS has <u>image stabilization</u> and <u>'Face Smile'</u> whatever that is.

     (a) **Mention:** Image Stabilization
          EntityType: CameraFeature
          MentionLevel: Nominal
          FeatureOf: Canon Powershot A590IS

     (b) **Mention:** Face Smile
          EntityType: CameraFeature
          MentionLevel: Named
          FeatureOf: Canon Powershot A590IS

5.4 CameraAccessory

An accessory is anything non-essential to the function of the camera. Sentiment expressed toward accessories has a smaller effect on the overall sentiment of the camera than sentiment expressed toward parts or features. Examples of this would be battery grip, camera bag, UV filter, and external computer software that interacts with camera software (like Adobe or iPhoto).

(66) Adding a <u>battery grip</u> resolves the problem of the camera body being short.

     (a) **Mention:** battery grip
          EntityType: CameraAccessory
          MentionLevel: Nominal
          AccessoryOf: body

6   Consumer Packaged Goods Domain Entity Types

The Consumer Packaged Goods (CPG) Domain encompasses all commercially packaged and marketed food products and their respective features. Entity Mention categories include Mention.Food, Mention.Beverage, Mention.FoodPart, Mention.FoodFeature, Mention.Nutrient, Mention.CookingMehod, Mention.CookingTool, Mention.Marketing and Mention.Meal.

6.1   Food

Label as 'food' all food, ingredients, prepared dishes and spices or anything bought at a grocery store with the exception of vitamins and minerals. The latter two will belong to an Entity Mention group entitled 'nutrient'. It does not matter whether the item is raw or has been cooked, mixed or otherwise transformed. Please see example 67 below for examples of Mention.Food.

(67)   Combine the <u>lettuce</u>, <u>parmesan cheese</u> and <u>salad dressing</u> to make a <u>Caesar salad</u>.

6.2   Beverage

Label as 'beverage' all drinks as in example 68.

(68)   For my party I bought <u>Pepsi</u> and <u>cranberry juice</u> and I stopped on the way home to get a <u>smoothie</u>.

6.3   Nutrient

Label as 'nutrient' all vitamins and minerals (*Vitamin D, Vitamin C, Potassium, Calcium*), additives (*MSG, artificial flavor, BHA, yellow 5*) and other ingredients in food that are not directly foods themselves. Nutrients will be marked as PartOf the food mention. An example of this can be seen in 69.

(69)   This yogurt contains <u>Omega 3s</u>, <u>calcium</u> and <u>red 40</u>.

   (a) **Mention:** Omega 3s
        PartOf: yogurt

   (b) **Mention:** calcium
        PartOf: yogurt

   (c) **Mention:** red 40
        PartOf: yogurt

6.4   FoodFeature

As with both the camera and car domains, the food domain contains both a 'FoodFeature' and 'FoodPart' class. The same conditions for this domain apply: concrete objects are labeled as 'part' whereas abstract objects are labeled as 'feature'. FoodFeatures would be any characteristic of the food such as its texture, flavor, color etc.

(70)   Lay's potato chips are delightfully <u>crunchy</u> and come in many <u>flavors</u> including <u>Bar-b-Que</u> and
       <u>Sour Cream 'n' Chive</u>.

   (a) **Mention:** crunchy
        EntityType: FoodFeature
        FeatureOf: Lay's potato chips

   (b) **Mention:** flavors
        EntityType: FoodFeature
        FeatureOf: Lay's potato chips

   (c) **Mention:** flavors
        EntityType: Food
        InstanceOf: Lay's potato chips

    (d) **Mention:** Bar-b-Que
           EntityType: Food

    (e) **Mention:** Bar-b-Que
           EntityType: FoodFeature
           InstanceOf: flavors
           FeatureOf: Bar-b-Que (Mention.Food)

    (f) **Mention:** Sour Cream 'n' Chive
           EntityType: Food

    (g) **Mention:** Sour Cream 'n' Chive
           EntityType: FoodFeature
           InstanceOf: flavors
           FeatureOF: Sour Cream 'n' Chive (Mention.Food)

Note: As illustrated in example 70, *Bar-b-Que* and *Sour Cream 'n' Chive* would be annotated twice, once as FoodFeature and once as Food. This is because the Mention 'potato chips' is not mentioned in the second part of the sentence and it can be assumed that the flavor of the potato chip is being used to discuss the potato chip itself. Also, if there are later mentions of that type of potato chip in the article, this convention allows them to corefer. This assumption is consistent with the car and camera domains, wherein mentions of parts are often used to refer to the product as a whole, as seen in example 57.

6.5   FoodPart

(Tentative) Label all chunks, slices and physical parts of an individual food item as FoodPart. This Mention class is still undergoing deliberation, please use at your discretion and send feedback.

6.6   Meal

Use this section to mark entities like *breakfast, lunch, dinner, meal, snack* etc. You can target the food items that make up a meal using the 'Food' slot which is discussed in the Entity Slots section below.

(71)   Pizza bagels are good for both <u>breakfast</u> and <u>lunch</u>!

    (a) **Mention:** breakfast
           EntityType: Meal
           Food: Pizza bagels

    (b) **Mention:** lunch
           EntityType: Meal
           Food: Pizza bagels

6.7   CookingMethod and CookingTool

The former Mention class refers to the method by which a meal is cooked like *steaming, boiling, toasting* etc. This is closely related with the latter Mention type CookingTool which consists of items used to aid in the cooking process like *blender, toaster, steaming bag* etc. Both CookingMethod and CookingTool contain the 'Food' slot which should be filled with the end product of the mentioned method or tool. CookingMethod also has the 'CookingTool' slot which can be filled by the CookingTool.

(72)   I put the broccoli in the <u>Ziploc Steam Bag</u> and <u>microwaved</u> it for 10 minutes and the broccoli was cooked perfectly.

    (a) **Mention:** Ziploc Steam Bag
           EntityType: CookingTool
           Food: broccoli

    (b) **Mention:** microwaved
        EntityType: CookingMethod
        Food: broccoli
        CookingTool: Ziploc Steam Bag

6.8 Marketing

Use this Mention type when annotating advertisements, contests, giveaways, promotions and packaging of a specific food brand. All Mentions labeled as marketing must fill the Produces slot of the Mention.Org mention so that it is clear what organization produces this marketing material. This is displayed in examples 73 and 74.

    The Marketing class also contains the AccessoryOf slot, which should be used for all types of marketing except for packaging. In cases where the marketing mention is not packaging, (see example 73) the Mention labeled as 'marketing' would fill its AccessoryOf slot with the marketed food item. In the case of packaging, it would me marked as PartOf the food instead of AccessoryOf, as shown in example 74.

(73) The new <u>ad</u> for Kraft cheese singles is really funny.

    (a) **Mention:** Kraft cheese singles
        EntityType: Food

    (b) **Mention:** Kraft
        EntityType: Organization
        Produces: ad

    (c) **Mention:** ad
        EntityType: Marketing
        AccessoryOf: Kraft cheese singles

(74) Pom Wonderful's new <u>bottle</u> is a great <u>shape</u>, and you can get a <u>coupon</u> for 50 cents off in this week's paper!

    (a) **Mention:** bottle
        EntityType: Marketing
        PartOf: Pom Wonderful

    (b) **Mention:** shape
        EntityType: Marketing
        FeatureOf: bottle

    (c) **Mention:** coupon
        EntityType: Marketing
        AccessoryOf: Pom Wonderful

    (d) **Mention:** 50 cents off
        EntityType: Unit.Money
        FeatureOf: coupon
    Mention: Pom Wonderful

    (e)   **EntityType:** Organization
        Produces: coupon, bottle, shape

7 General rules for all Entity types

7.1 Entity subtypes

For all entity types, choose the appropriate subtype where possible. If there is no appropriate subtype or if you are not sure, then label the entity as the highest level type. For example, label a vehicle as Mention.SUV if you are sure that it is an SUV. Otherwise label it as Mention.Vehicle.

7.2   Text spans

When determining the text span of a Entity Mention, only select the text that is directly relevant to the classification of the Mention. See the sections on Time Entities and Descriptors for the span selection rules that are specific to those mentions.

- Leave out definite and indefinite articles (e.g., *the <u>car</u>, a <u>car</u>, some <u>car</u>*);

- Leave out quantifiers, i.e. words indicating the quantity (e.g., *three <u>cars</u>, no <u>cars</u>, all <u>cars</u>*);

- Leave out adjectives, both relevant and not relevant to the definition of the entity (e.g., *the new <u>car</u>*). Mark relative adjectives as Descriptors such as 'hybrid' in *<u>hybrid</u> Honda Civic*;

- Label possessive pronouns (*my, your, his, her, their, our*) as a separate Mention (e.g., in *<u>My</u> new <u>Toyota</u>* tag 'Toyota' as one Mention and 'My' as another);

- Leave out demonstrative pronouns (*this, that, these, those*) that are used as determiners before another noun (e.g., *this <u>car</u>, these <u>cars</u>, that <u>car</u>*);

- Label demonstrative pronouns that occur independently as Mentions and are not relative pronouns. (e.g., *I like <u>that</u>; <u>Those</u> look great!*);

- Label any compounds formed with *some-, any-* and *no-* if they conform to an entity type (e.g., *<u>Anybody</u> can drive that car.*).

- Label wh-words (*who, what, where, how, when*) if they correspond to one of the entity types and are not used as relative pronouns (e.g., *I don't know <u>when</u> he arrived.*)

- Leave out clauses that modify the noun (e.g., *the <u>car</u> that was parked on the street; the <u>man</u> who walked into the store; the <u>car</u> parking on the street; the <u>seat</u> in the back; a <u>tire</u> of a large SUV*).

7.3   Dummy Pronouns

Only pronouns that clearly refer to an entity should be labeled as EntityMentions. Pronouns and demonstratives that do not refer to a specific entity but are there for purely grammatical reasons or are part of an idiomatic phrase should not be annotated.

(75) **Dummy pronouns:**
It's raining.
It's tough when you don't have the money to get it repaired.
At highway speeds *it* was hard to keep the car on center.

(76) **Relative pronouns:**
I saw the car *that* was on the road.

(77) **Idiomatic expressions:**
He was so mad, he just lost *it*.
The car just couldn't make *it* up the hill.

7.4   Compounds

When annotating entity mentions that are compounds (noun+noun, adj+noun, verb+noun, etc), the following heuristics can be useful in determining the correct span of the entity mention.

- Compounds that are one word or hyphenated must be treated as one span

- For all others, determine whether the mention is a noun+noun compound, an adj+noun compound, or a verb+noun compound

1. **noun+noun compound:** determine if the entity described by the head noun is made of the entity described by the first noun. For example, *salt pile - a pile made of salt; cloth bag - a bag made of cloth.*

   – If so, then the head noun should be labelled separately. The first noun can be annotated as a separate Entity Mention, if relevant to the domain, or not at all. For example, in *leather seat* 'leather' is what the seat is made of, thus it should not be part of the Entity Mention 'seat'.

   – Conversely, if the entity described by the first noun holds some other relation to the entity described by the head noun, then the first noun should be included in one span with the head noun. For example, in *camera bag* the bag is not made of cameras, it is a bag for cameras; thus the modifying noun 'camera' is part of the span of the whole Entity Mention span *camera bag*).

2. **adj+noun compound:** determine if the adjective preceding the head noun can be graded with suffixes or adverbs (very shiny, shiniest), without drastically altering the meaning of the compound.

   – If the adjective can be graded, then it should not be included in the span of the head noun. The adjective is a relevant feature to the domain it can be annotated separately. For example, in *professional camera*, 'professional' can be graded so it should not be included in the Mention.Camera span. It is relevant to the camera domain so it can be annotated separately as a Mention.Feature and then FeatureOf the camera mention).

   – If the adjective cannot be graded without altering the meaning of the compound, then it should be included in one span with the head noun (e.g *automatic transmission* and *digital camera*: 'automatic' and 'digital' cannot be logically graded with suffixes and adverbs, rather these are either/or properties of the head noun; thus these adjectives should be included in the span of the head noun.

3. **verb+noun compound:** these should be treated as one span (e.g., *point-and-shoot cameras*).

- Where the modifying noun/adj/verb is included in one span with the head noun, the annotator may optionally double-annotate the modifying term, often as Mention.Feature and featureOf the larger, compound mention (e.g., *point-and-shoot cameras*: 'point-and-shoot' would be annotated separately as Mention.Feature and then annotated as FeatureOf 'point-and-shoot' cameras).

7.5   Some tricky cases

7.5.1   Entity type ambiguity

There are many instances where the same text string can be used to refer to different entity types. In the examples below, Alcatraz can be used to refer to a location (the island) or to a facility (the prison on the island). If possible, use the context to determine which type is being referred to. In Examples 78 and 79, the preposition 'on' implies that the reference is to the island. In Example 80, the descriptions 'aging' and 'crumbling' imply that the reference is to the prison building.

(78) Growing up on <u>Alcatraz</u>, Jean Comerford, daughter of a prison guard, was used to spectacular views of the Golden Gate Bridge and the East Bay from her bedroom window

    (a) **Mention:** Alcatraz
          EntityType: Location
          Level: Name

(79) The idea to turn one of the historic buildings on <u>Alcatraz</u> into a hotel is among a variety of proposals by the National Park Service.

    (a) **Mention:** Alcatraz
          EntityType: Location
          Level: Name

(80) <u>Alcatraz</u> as portrayed in the movie is an aging, crumbling shadow of <u>its</u> former self. (Indeed, the <u>prison</u> was closed in 1963.)

(a) **Mention:** Alcatraz
   EntityType: Facility
   Level: Name

(b) **Mention:** it
   EntityType: Facility
   RefersTo: Alcatraz
   Level: Pronominal

(c) **Mention:** prison
   EntityType: Facility
   RefersTo: Alcatraz
   Level: Nominal

7.5.2 Misspellings

If words that refer to entities are misspelled, they should still be annotated as the entity they are intended to be referring to. In example 81 below, 'residents' is misspelled as 'residence' but should still be annotated as a PERS entity. In Example 82, there is no space between 'website' and 'with'. 'Website' can still be included in the span for the ORG entity (using the arrows in the top right box of the screen) 'Navy Electronic Commerce Online website' by excluding 'with' from the selection.

(81) The ever expanding Moreton Hall estate, is to have a recording studio built. This will be on the road that leads down to the Abbey gardens. <u>Local residence</u> are concerned about noise issues, as is the church reverend

 (a) **Mention:** Local residence
   EntityType: Person
   Level: Nominal

(82) This contract was competitively negotiated via the <u>Navy Electronic Commerce Online website</u>with six proposals received.

 (a) **Mention:** Navy Electronic Commerce Online
   EntityType: Organization
   Level: Name

Similarly, ignore incorrect apostrophes in plurals, e.g.

(83) I sold two <u>car's</u>.

 (a) **Mention:** car's
   EntityType: Vehicle.Car
   Level: Nominal

7.5.3 Punctuation

When selecting a text span for an annotation class, do not include surrounding punctuation, such as parentheses, quotation marks or commas. In Example 84 below, the quotation marks are not included in the annotation of the Sentiment Expression *spongy*.

(84) The brake pedal has a "<u>spongy</u>" feel to it.

 (a) **SentimentBearingExpression:** spongy

Do include apostrophes that are used as abbreviations, for example in year mentions:

(85) The <u>'89</u> Toyota Camry

   (a) **Mention:** '89

Do not include end-of-sentence punctuation, such as periods, question marks or exclamation marks, even if these can be interpreted as contributing to the sentiment of the expression, e.g.

(86) It sure did <u>run</u>!

   (a) **Mention:** run
       EntityType: CarPart
       FeatureOf: It

   (b) **SentimentBearingExpression:** sure
       Polarity: positive

8   Entity Slots

The following subsections describe the slots that can be filled when annotating a Mention.

8.1   Levels

When marking a Mention you must also determine and label its entity level, i.e. whether the entity is a pronominal, nominal or named. This applies to all Mention types *except for Time Mentions which are not marked in this way.*

8.1.1   Pronominal

This level includes the following:

- Pronouns: *I, me, you, he, him, she, her, it, we, us, they, them. anyone.*

- Possessive pronouns: *mine, your, yours, his, her, hers, its, our, ours, their, theirs.*

- Demonstrative pronouns: *this, that, these, those*

- Other pronouns: *one, something, anything, nothing*

All of these should be marked as *Pronominal* if they are Mentions.

8.1.2   Name

A named entity is an entity that is being referred to with a proper name. These should be marked as *Name* if they are Mentions. In English, proper names are often capitalized. Examples are *John, Mary, Mrs Smith, Subaru Forester, Subaru WRX, (the) White House, Paris, England, USA, (the) US.*

Note that in blogs, named entities are not always capitalized because blog authors do not always pay attention to upper and lower case distinctions in their spelling. Even if named entities are written in lower case they should still be marked as *Named*, e.g.

(87) i love my *subaru*!!!

8.1.3   Nominal

Entities at the nominal level are all Mentions that do not fall in either the pronominal or the named entity category. These are usually ordinary nouns that cannot be considered names, e.g., *car, country, city, person, man, woman, guy.*

When annotating the level of a Mention, take into account only the level of the Mention you are currently annotating, not any other Mentions it may be coreferring with. In the example below, *Subaru* and *It* are coreferring, and both are marked as EntityType Vehicle.Car. However, *Subaru* is marked as Named and *It* is marked as Pronominal.

(88) I love my <u>Subaru</u>!!! <u>It</u>'s a great <u>car</u>.

    (a) **Mention:** Subaru
           EntityType: Vehicle.Car
           Level: Name

    (b) **Mention:** It
           EntityType: Vehicle.Car
           RefersTo: Subaru
           InstanceOf: car
           Level: Pronominal

    (c) **Mention:** car
           EntityType: Vehicle.Car
           Level: Nominal

Entities where the head noun is a nominal should be labeled as nominal, even if there is a proper name in a modifying position, e.g.

(89) I went to the <u>Pontiac dealer</u>.

    (a) **Mention:** Pontiac dealer
           EntityType: Mention
           Level: Nominal

8.2   RefersTo

Coreference occurs when noun phrases in a text refer to the same entity. For example, in the sentence *Mary went to the store and she bought some ice cream* the noun phrases *Mary* and *she* refer to the same entity. They are said to be coreferring.

The term coreference chain is used to describe all the noun phrases in a text that refer to the same entity.

If two EntityMentions are coreferring they should be classed as the same EntityType and the first Mention should appear in the RefersTo slot of the second. In Example 90, Toyota is given the EntityType Organization and the first *it*, which is coreferring, gets the same EntityType. The second *it* is coreferring with the *hybrid Camry* and therefore gets the EntityType Veh.Car.

(90) Mr Watanabe said <u>Toyota</u> had decided to add <u>Australia</u> as a location for producing the <u>hybrid Camry</u> because <u>it</u> was confident <u>it</u> would be well-received

    (a) **Mention:** Toyota
           EntityType: Organization
           Level: Name

    (b) **Mention:** hybrid Camry
           EntityType: Veh.Car
           Level: Name

    (b) **Mention:** it
           EntityType: Organization
           RefersTo: Toyota
           Level: Pronominal

    (b) **Mention:** it
           EntityType: Veh.Car
           RefersTo: hybrid Camry
           Level: Pronominal

In Example (91) the EntityMentions *2009 Toyota Corolla, Car, This* and *it* are all coreferring. In the annotation, the RefersTo slots of these entities should be filled with the immediately preceding entity from the coreference chain:

(91) It has been a bit but on my recent trip to Virginia I rented a new (only 1000 miles on the clock) <u>2009 Toyota Corolla</u>. <u>Car</u> came with a cloth interior. <u>This</u> was a pretty base model car but <u>it</u> had some nice features.

    (a) **Mention:** 2009 Toyota Corolla
        EntityType: Veh.Car

    (b) **Mention:** Toyota
        EntityType: Organization

    (c) **Mention:** 2009
        EntityType: Time.Year

    (d) **Mention:** Car
        EntityType: Veh.Car
        RefersTo: 2009 Toyota Corolla

    (e) **Mention:** This
        EntityType: Veh.Car
        RefersTo: Car

    (f) **Mention:** it
        EntityType: Veh.Car
        RefersTo: This

In rare cases, a pronoun will appear before its coreferring expression in a text. In Example 92, below, the proper name *R8 V10* occurs after the pronoun *she* that refers to it. In these cases, the RefersTo slot of the pronoun should be filled with the more descriptive Mention instance (for example, a proper noun).

(92) So here <u>she</u> is, the much talked about <u>R8 V10</u>, without camouflage or tape to cover the details.

    (a) **Mention:** she
        RefersTo: R8 V10

The RefersTo slot should be filled for all co-referring Mentions *except* for Time Mentions. If the same year, time or day is mentioned more than once in the text they should be labeled individually as Time Mentions but the RefersTo slots should be empty.

8.3 PartOf

The PartOf slot is used to annotate whether a given Mention can be described as being a part of another. In Example 1, *a cloth interior* is part of the Toyota Corolla. Its PartOf slot should therefore be filled with the most recent mention of this entity, in this case the mention *Car* (in a previous sentence).

(93) Don't care too much about the <u>cloth interior</u>
    Mention: cloth interior
    EntityType: CarPart
        PartOf: Car

8.4 MemberOf

The MemberOf slot is used when a plural mention includes more than one individual Mention as in example 94. This also works when more than one individual Mention is subsequently referred to using a plural Mention. In example 95 *we* includes both *John* and *I*. The MemberOf slot of the two singular nouns indicates that they are each a member of *we*.

(94) This car has a lot of good <u>safety features</u> like: <u>airbags</u>, <u>ABS</u>, and a <u>roll bar</u>.

(a) **Mention:** airbags
   EntityType: CarPart
   MemberOf: safety features

(b) **Mention:** ABS
   EntityType: CarPart
   MemberOf: safety features

(c) **Mention:** roll bar
   EntityType: CarPart
   MemberOf: safety features

(d) **Mention:** safety features
   EntityType: CarPart

(95) <u>John</u> worked late and <u>I</u> stayed up to watch TV. The next day <u>we</u> were both tired.

(a) **Mention:** John
   EntityType: Person
   MemberOf: we

(b) **Mention:** I
   EntityType: Person
   MemberOf: we

(c) **Mention:** we
   EntityType: Person

In cases where the group word is a pronoun, the MemberOf slot should still be filled by the individual Mentions. In example 96 below, *interior* and *exterior* are MemberOf *both*.

(96) <u>Both</u> the <u>interior</u> and <u>exterior</u> look nice.

(a) **Mention:** Both
   EntityType: Vehicle.CarPart

(b) **Mention:** interior
   EntityType: Vehicle.CarPart
   MemberOf: both

(c) **Mention:** exterior
   EntityType: Vehicle.CarPart
   MemberOf: both

Note that we do NOT fill the RefersTo slot of *we* because it is not strictly coreferring with either of the preceding Mentions but only to both of them together as a group.

8.5   FeatureOf

FeatureOf is similar to PartOf except that it is used for abstract features rather than concrete parts of an entity. For the domain cars, feature examples would be *gas mileage* or *performance*.

Example 97 has two EntityMentions (*about 8 miles to the gallon* and *gas milage*) that describe the feature of a car. In both of these cases the FeatureOf slot should be filled with *car*. Note that, the RefersTo slot of *gas milage* is filled with *about 8 miles to the gallon*, as the two are co-referring.

(97) The car ha been sitting many years before I got it. Here is my problem I am getting
   <u>about 8 miles to the gallon</u>. Where should I start to try and improve the <u>gas milage</u>?

(a) **Mention:** about 8 miles to the gallon
   EntityType: CarPart
   FeatureOf: car

(b) **Mention:** gas milage
>> **RefersTo:** about 8 miles to the gallon
>> **EntityType:** CarPart
>> **FeatureOf:** car

Some other car features can be phrases like *goes 0-60 in 8 seconds* and *does a 1/4 mile in 14.6s at 97mph.*

For some words, like *looks* or *feels* it can often be difficult to decide whether the author is actually using them to refer to a feature or a committer. Because of frequent ambiguities our default will be to *mark these verbs as committers.*

(98) This car <u>feels</u> safe.

(a) **Mention:** car
>> **EntityType:** Vehicle.Car

(b) **Committer:** feel
>> **Direction:** downward
>> **target:** safe

(c) **SentimentBearingExpresseion:** safe
>> **Polarity:** positive
>> **Target:** car

Only if the word is clearly used to express the author experiencing a tactile or visual stimulus should be annotated as a CarPart/FeatureOf, e.g.

(99) These seats <u>feel</u> soft.

(a) **Mention:** seats
>> **EntityType:** CarPart

(b) **Mention:** feel
>> **EntityType:** CarPart
>> **FeatureOf:** seats

(c) **SentimentBearingExpresseion:** soft
>> **Polarity:** positive
>> **Target:** feel

8.6 InstanceOf

The InstanceOf slot is used for cases where a text refers to a generic type and then subsequently to an instance of that type. In Example 100 below, the Mention *R8 V10* refers to a particular model of car. The subsequent Mention *car*, however, refers to a specific instance of the model, namely the particular car in the photo. The context makes clear that the first mention *R8 V10* is not referring to a particular instance (i.e. the car in the photo is not *much talked about*, instead it is that new model in general.)

(100) So here <u>she</u> is, the much talked about <u>R8 V10</u>, without camouflage or tape to cover the details.The peculiarity of this spy shot is that the <u>car</u> has been snapped in a dealership.

(a) **Mention:** she
>> **EntityType:** Car
>> **RefersTo:** R8 V10

(b) **Mention:** R8V10

(c) **Mention:** car
>> **EntityType:** Car
>> **InstanceOf:** R8 V10

There is some ambiguity around the use of the pronoun *she* in the above example. It seems to be used to refer to the specific instance of the car in the photo but also to the generic model. The annotation in this example assumes that it is referring to the generic model and so its InstanceOf slot is empty. In difficult cases like these, the annotators should use their own judgement to make the final decision.

Copula constructions are sentences with the verb *be* that equate two entities. In Example 101 the Mention *This* is one entity and the Mention *car* is the other. The second entity is often an attribute or a description of the first. In these constructions we will fill the InstanceOf slot of the first entity with the second one. In the example below, *This* is an InstanceOf *car*.

(101) This is a nice car!

    (a) **Mention:** This
          EntityType: Veh.Car
          InstanceOf: car

    (b) **Mention:** car
          EntityType: Veh.Car
          RefersTo: This

    (c) **SentimentBearingExpression:** nice
          Target: car
          Polarity: positive

(102) The VW Rabbit is <u>one</u> of the more powerful cars of the group.

    (a) **Mention:** VW Rabbit
          EntityType: Veh.Car

    (b) **Mention:** one
          EntityType: Veh.Car
          InstanceOf: cars
          RefersTo: VW Rabbit

    (c) **SentimentBearingExpression:** powerful
          Target: one
          Polarity: positive

8.7 AccessoryOf

This slot is used for the Mentions CameraAccessory and Food.Marketing. This slot is used to link the accessory or marketing Mention with the camera or food Mention, respectively. An accessory for the camera domain is anything non-essential to the function of the camera or peripheral objects.

An accessory for the food domain would be any type of marketing (not including packaging) like coupons, promotions, giveaways etc. This is to ensure that Mentions like giveaway prizes and coupons do not appear to be part of the actual food item.

Sentiment expressed toward Accessories has a smaller effect on the overall sentiment of the camera than sentiment expressed toward parts.

(103) This <u>camera bag</u> fit my camera perfectly!

    (a) **Mention:** camera bag
          EntityType: CameraAccessory
          AccessoryOf: camera

Note: Also see examples 73 and 74 in the marketing section above for examples of how this slot is used in the Consumer Packaged Goods section.

8.8 Produces

Organization Mentions have this slot that can be filled with the product they create. This applies to all the domains: car, camera and CPG, as displayed in example 104 The Produces slot can also be filled with the Mention.Marketing in regards to ads, promotions and other related marketing strategies for food items in the CPG Domain. For an example of this please see 73.

(104) Hyndai has finally released their new Genesis Coupe.

    (a) **Mention:** Hyundai
           EntityType: Organization
           Produces: Genesis Coupe

    (b) **Mention:** Genesis Coupe
           EntityType: Vehicle.Car

8.9 Food

This slot is to be used with CookingMethod, CookingTool and Meal. The slot in the CookingMethod and CookingTool Mentions should only be filled with the Mention.food that represents the final product, not the ingredients. In example 105 below only the Mention.food 'bread' should be used for the 'food' slot in Mention.CookingTool and Mention.CookingMethod.

(105) I put some Harvest Wheat slices in my Black and Decker toaster and the toast was perfect.

    (a) **Mention:** Black and Decker toaster
           EntityType: CookingTool
           Food: toast

Likewise, in example 106 only 'milkshake' should fill the Food slot instead of all the ingredients.

(106) You can combine bananas, apples, cherries and vanilla ice cream then blend for a few minutes to make a cool summer milkshake.

    (a) **Mention:** blend
           EntityType: CookingMethod
           Food: milkshake

Mention.Food can also fill the 'food' slot for Mention.meal when the food is discussed in terms of meals.

(107) Yoplait Kids Yogurt is a great snack.

    (a) **Mention:** snack
           EntityType: Meal
           Food: Yoplait Kids Yogurt

8.10 CookingTool

The cooking tool slot should be filled with the relevant Mention.CookingTool when both a cooking tool and a cooking method are mentioned. The 'Glad Steamer Bag' would fill the 'cooking tool' slot for the Mention.CookingMethod 'microwaved'.

(108) I microwaved the butternut squash in the Glad Steamer Bag and it was yummy!

    (a) **Mention:** Microwaved
           EntityType: CookingMethod
           CookingTool: Glad Steamer Bag

    (b) **Mention:** Glad Steamer Bag
           EntityType: CookingTool
           Food: butternut squash

NOTE: In example 108 above, the Food Mentions 'butternut squash' and 'it' would corefer.

8.11   EntitySentiment

EntitySentiment should be used to describe the overall sentiment of an entity as expressed in the entire passage. EntitySentiment should be marked on the *anchor mention* or the first mention of an entity that is not a pronoun. (This is not necessarily the first mention of the entity. See example 109 below) Also EntitySentiment needs only to be expressed once per entity.

Every new Mention should have its EntitySentiment marked if there is sentiment expressed toward it in some way. This means that *I* and *last July* in the example below do not need EntitySentiment as there isn't any sentiment expressed.

(109) I had heard about <u>it</u> for years but it was only last July that I was able to drive the <u>Honda</u> Civic and it was awesome! Way better than my old <u>Ford Taurus</u>.

    (a) **EntityMention1:** it
           EntityType: Vehicle.Car

    (b) **Mention:** Honda Civic
           EntityType: Vehicle.Car (*anchor mention*)
           EntitySentiment: Positive

    (c) **EntityMention2:** Ford Taurus
           EntityType: Vehicle.Car (*anchor mention*)
           EntitySentiment: Negative

8.12   ContextualSentiment

ContextualSentiment should be marked on a Mention if two conflicting sentiments are applied to a single instance of the mention. The sentiment should be marked on each SentimentBearingExpression that links to the Mention and the ContextualSentiment should be marked as the stronger sentiment in the sentence. See 110 below.

(110) The Jeep Liberty is <u>attractive</u>, but <u>poorly built</u>.

    (a) **SentimentBearingExpression:** attractive
           PriorPolarity: Positive
           Target: Jeep Liberty

    (b) **SentimentBearingExpression:** poorly built
           PriorPolarity: Negative
           Target: Jeep Liberty

    (c) **Mention:** Jeep Liberty
           EntityType: Vehicle.Car
           ContextualSentiment: Negative

Also mark ContextualSentiment when there is convoluted Sentiment Expression chain where a single descriptive Sentiment Expression is not apparent, illustrated in example 111

(111) Without a doubt the EOS 450D will have to do more today to prove itself against the popular Nikon D40.

    (a) **Mention:** EOS 450D
           ContextualSentiment: Positive (based on other content in article)

    (b) **Mention:** Nikon D40
           ContextualSentiment: Positive (based on content of sentence)

8.13 MentionPriorPolarity

The MentionPriorPolarity slot has two possible values: positive and negative. It is used for EntityMentions whose name alone indicates a certain sentiment. In Example 112 below, the Mention *lemon* carries a negative sentiment so its MentionPriorPolarity slot is set to *negative*.

(112) That's what the brand new facade is about. They figured a quick surface fix and they unload the <u>lemon</u> on some unsuspecting chump and walk away. Quickly.

    (a) **Mention:** lemon
            MentionPriorPolarity: Negative

(113) Is it worth trading in your <u>gas guzzler</u>?

    (a) **Mention:** gas guzzler
            MentionPriorPolarity: Negative

Authors can be very creative in their use of MentionPriorPolarity in Entity Mentions. In Example 114 below the author cites the term *'Saabaru'* - a mix of *Subaru* and *Saab* - as a derogatory description of a new Saab model. This newly-coined term can be labeled as an EntityMention.Car with a negative MentionPriorPolarity.

(114) Five grand. Depending on options, incentives and fire sales, that's the difference between the cost of a Saab 9-2X Aero and a Subaru WRX Sport Wagon. Underneath, there's not much in it: same platform, same bag of tricks.No wonder auto industry wags have taken to calling the Saab 9-2X Aero the <u>'Saabaru'</u>.

    (a) **Entity Mention:** Saabaru
            EntityType: Car
            RefersTo: Saab 9-2X
            MentionPriorPolarity: negative

9 SentimentBearingExpression

Sentiment bearing expressions (Sentiment Expressions) are expressions that indicate that the author has some sentiment toward an entity (or sometimes toward another sentiment bearing expression). Sentiment Expressions are varied in their form and can be adjectives, verbs, nouns or more complex phrases.

Sentiment Expressions have two slots that must be filled (i.e.they are not optional). The *Target* is the entity that the sentiment is expressed about. The *PriorPolarity* is the polarity of the sentiment, and it has 4 potential values: *Positive, Negative, Neutral* and *Mixed*.

In Example 115 below, the adjective *great* is a positive Sentiment Expression with *car* as its target.

(115) This MGB has been re-listed here on eBay. It remains a <u>great</u> car and we wish it would sell instead of languishing at the dealer.

    (a) **SentimentBearingExpression:** great
            Target: car
            PriorPolarity: positive

Note that Sentiment Expressions should not be marked in sentences of reported fact, only mark them if an emotion or value judgement is conveyed. In Example 116 below, the adjective *wrecked* is not necessarily a complaint, however it carries a negative connotation and so should be marked as a negative Sentiment Expression. In contrast, Example 117 has the adjective *old*, which is neither a direct complaint nor does it necessarily have a negative connotation in this context. It should therefore be left unmarked.

(116) I drove my <u>wrecked</u> Ford all the way to Los Angeles.

    (a) **SentimentBearingExpression:** wrecked
            Target: Ford
            PriorPolarity: negative

(117) I drove my old Ford all the way to Los Angeles.

Examples 118 and 119 show that Sentiment Expressions can also be more complex. The descriptions *within reach* and *easy to read* as a whole are classed as a positive Sentiment Expressions.

(118) Seats are comfortable, all the controls are well within reach.

    (a) **SentimentBearingExpression:** comfortable
        Target: Seats
        PriorPolarity: positive

    (b) **SentimentBearingExpression:** within reach
        Target: controls
        PriorPolarity: positive

    (c) **Intensifier:** well
        Target: within reach
        IntensifierDirection: strengthen

(119) Gauges where all easy to read.

    (a) **SentimentBearingExpression:** easy to read
        Target: Gauges
        PriorPolarity: positive

In Example 120, the author is expressing an opinion on the stereo but it is neither positive nor negative. The PriorPolarity of *ok* is marked as *neutral*.

(120) Stereo was ok

    (a) **SentimentBearingExpression:** ok
        Target: Stereo
        PriorPolarity: neutral

Verbs by themselves can also be Sentiment Expressions. These must bear sentiment toward either the speaker and Mention (example 121) or just the Mention (example 122). Words that can bear sentiment toward the the speaker and Mention are: *brag, bitch, whine, complain.*

(121) Joe whined about his 800 dollar car payment.

    (a) **SentimentBearingExpression:** whine
        Target: Joe, 800 dollar car payment
        PriorPolarity: negative

(122) I liked the exterior.

    (a) **SentimentBearingExpression:** liked
        Target: exterior
        PriorPolarity: positive

It is also possible to link one Sentiment Expression to two or more EntityMentions as in Example 123 below.

(123) I liked the exterior and the power steering.

    (a) **SentimentBearingExpression:** liked
        Target 1: exterior
        Target 2: power steering
        PriorPolarity: positive

Adjectival words and phrases that have inherent sentiment should always be marked as a Sentiment Expression. These words include: *ugly/pretty, good/bad, wonderful/terrible/horrible, dirty/clean.* There is also another type of adjective that doesn't have inherent sentiment but rather sentiment based on the context of the sentence. This means that these adjectives can take either positive or negative sentiment depending on the Mention that they are targeting and also other Sentiment Expressions in the sentence. For example, a *large salary* is positive whereas a *large phone bill* is negative. These adjectives should only be marked as Sentiment Expressions if the sentiment they are conveying is stated clearly in the surrounding context. In other cases these adjectives merely specify a Mention further instead of changing its sentiment.

Nouns can be marked as Sentiment Expressions if they convey strong sentiment, portray an abstract concept and don't fit into one of the Mention subcategories. These are nouns such as doom and gloom, idiot, piece of junk, etc. Any nouns that don't have all three qualities need to be marked as a Mention with either ContextualSentiment, EntitySentiment, MentionPriorPolarity or a combination of the three.

9.1   What not to mark

- Do not annotate Sentiment Expression's that do not have a target.

- Do not annotate events that can be interpreted as being positive or negative but do not have a clear Sentiment Expression and/or target. Example 124, is clearly negative in sentiment towards a car. However, there is no obvious Sentiment Expression or Sentiment Expression-Target, so no sentiment is annotated around this sentence.

(124) I couldn't get the seats installed.

10   Intensifiers

Intensifiers serve to increase or decrease the intensity of an Sentiment Expression. Intensifiers have a Target slot, which is filled by the instance of the Sentiment Expression or another intensifier (see example 126). The Direction slot has two values: *strengthen* or *weaken*, depending on whether the Sentiment Expression PriorPolarity is being increased or decreased by the Intensifier.

In (125) the negative polarity of the Sentiment Expression *annoying* is increased by the intensifier *extremely*.

(125) This was extremely annoying when cruising the interstates.

  (a) **SentimentBearingExpression:** annoying
       Target: This
       PriorPolarity: negative

  (b) **Intensifier:** extremely
       IntensifierTarget: annoying
       IntensifierDirection: strengthen

(126) The seats in this car are way more comfortable.

  (a) **SentimentBearingExpression:** comfortable
       Target: seats
       PriorPolarity: positive

  (b) **Intensifier:** way
       IntensifierTarget: more
       IntensifierDirection: strengthen

NOTE: Some words can act as Intensifiers or Sentiment Expressions depending on where they are in the sentence. An example of this is *well*. When following a verb, well is always a Sentiment Expression: *This car drives well.* When preceding an adjective, usually in partially idiomatic usages, it then becomes an intensifire. In example 127 'well' targets 'above'.

(127) This car's preformance is <u>well</u> above most other's.

    (a) **SentimentBearingExpression:** above
         Target: preformance
         PriorPolarity: positive

    (b) **Intensifier:** well
         IntensifierTarget: above
         IntensifierDirection: strengthen

11   Negator

Negators are terms that can flip the sentiment of a Sentiment Expression, other Negator or an Intensifier, e.g *not, no, never.*

In Example 128, there is a positive Sentiment Expression *amazing* that has *Stereo* as its target. However, there is also a negator *not* that has that Sentiment Expression as its target. Although the sentiment expressed here is negative, this is not explicitly annotated anywhere. Instead, we label a positive Sentiment Expression and assume that its sentiment is flipped by the Negator.

(128) Stereo was ok, but <u>not</u> amazing.

    (a) **SentimentBearingExpression:** amazing
         Target: Stereo
         PriorPolarity: positive

    (b) **Negator:** not
         NegatorTarget: amazing

This example also shows that for contractions involving negations (e.g., *didn't, wouldn't, can't, hasn't* etc) we do not separate the negation (*n't*)from the verb (*did*) and instead select the whole construct as the text span of the negator.

Negators can also be used to express the non-existence of entities, as in Example 129, below.

(129) Seats are comfortable, all the controls are well within reach. Glad to see <u>no</u> leatherette.

    (a) **Negator:** no
         NegatorTarget: leatherette

    (b) **Mention:** leatherette
         PartOf: seat

12   Neutralizer

A Neutralizer is an expression that is used to denote a hypothetical case or a case that is assumed, desired or required but not definitely true.*If* is a common neutralizer because the subsequent phrase is hypothetical. The verbs *want, hope, would, could, should* and *expect* are also examples of neutralizers because they express the author's wishes rather than an actual state in the world.

Neutralizers have a NeutralizerTarget slot that should be filled with a Sentiment Expression, a Negator a Mention or another Neutralizer that is hypothetical or non-factual.

(130) <u>If</u> the car is fast I will be happy.

    (a) **Neutralizer:** If
         NeutralizerTarget: fast

(131) These are the first photos I made from the new Audi R8 V10 that is coming to us in 2009. I <u>hope</u> you like it.

    (a) **Neutralizer:** hope
         NeutralizerTarget: like

13   Other Person's Opinion (OPO)

Quite often the sentiments related in a text are not the author's own. We will mark these cases as an Other Person's Opinion (OPO). The text span of the OPO is whatever indicates that an opinion is being expressed. Often OPOs are verbs of saying and thinking where the subject is not the author. These are words such as *say, think, believe, encourage, recognize, tell* etc. The OPOTarget slot is filled by a Sentiment Expression or a Mention that is the other person's expression of sentiment. The OPOSource slot is filled by a Mention which corresponds to the person expressing the sentiment.

(132) My brother <u>says</u> this car sucks.

    (a) **OPO:** says
           OPOSource: My brother
           OPOTarget: sucks

    (b) **Sentiment Expression:** sucks
           Target: car
           PriorPolarity: Negative

(133) Five grand. Depending on options, incentives and fire sales, that's the difference between the cost of a Saab 9-2X Aero and a Subaru WRX Sport Wagon. Underneath, there's not much in it: same platform, same bag of tricks. No wonder <u>auto industry wags</u> have taken to <u>calling</u> the Saab 9-2X Aero the '<u>Saabaru</u>'.

    (a) **OPO:** calling
           OPOTarget: 'Saabaru'
           OPOSource: auto industry wags

    (b) **Sentiment Expression:** 'Saabaru'
           OPOTarget: Saab 9-2X Aero
           PriorPolarity: negative

In some cases the OPOSource slot can be left empty when there is no mention of the person expressing the opinion:

(134) For Audi, however, this engine would be producing upwards of 350 horsepower despite a <u>rumor</u> that it is even <u>smaller</u> in displacement at 2.3 liters.

    (a) **OPO:** rumor
           OPOTarget: smaller

    (b) **Sentiment Expression:** smaller
           Target: displacement
           PriorPolarity: negative

There is also a category of OPO that can be a Sentiment Expression as well, where the Sentiment Expression instance would be the Dimension of the OPO instance (example 135). Words that fit into this category would be: *recognize, encourage, commend, disparage, diss.* If there is no obvious OPO present then a Sentiment Expression can function as an OPO as well if the Mention is someone other than the author as in example 136. These would would be: *like, hate, despise.*

(135) The 2008 Hyundai Sonata was <u>recognized</u> by the Greener Choices Society.

    (a) **OPO:** recognized
           OPOTarget: recognized (Sentiment Expression instance)
           OPOSource: Greener Choices Society

    (b) **Sentiment Expression:** recognized
           Target: 2008 Hyundai Sonata
           PriorPolarity: positive

(136) My brother <u>likes</u> this camera a lot.

    (a) **OPO:** likes
            OPOTarget: likes (Sentiment Expression instance)
            OPOSource: brother

    (b) **Sentiment Expression:** likes
            Target: camera
            PriorPolarity: positive

14   Comparison

The Comparison class is used for labeling comparative adjectives and comparative quantifiers. The former are single-word adjectives with the suffix -er (*prettier, bigger, faster*) as well as irregulars such as *worse*. The latter are two-word phrases with the quantifier (*more, less, most*) plus an adjective or adverb (*more beautiful, less efficient*).

    Comparisons have four slots that are filled in each instance. The Less and More slots are filled with the Entity Mentions that are being compared. The Dimension must be marked by the comparative word or phrase. The Same slot is marked 'true' when two Entity Mentions are equal in the trait in question and 'false' when they are not equal.

    In the adjective+er comparisons, the comparison word is also marked as a Sentiment Expression (either positive or negative) and the Sentiment Expression instance of that word acts as the Dimension of the Comparison. This is illustrated in example 137 below.

(137) Witness the rise of the SUV and the recent explosion of crossovers, all of which are basically five-door wagons with <u>higher</u> seating positions, <u>greater</u> (but often unused) towing and payload capacities, and <u>cruddier</u> fuel economy.

    (a) **Comparison:** higher
            Less: wagons
            More:crossovers
            Dimension: higher (Sentiment Expression instance)

    (b) **Comparison:** greater
            Less: wagons
            More: crossovers
            Dimension: greater (Sentiment Expression instance)

    (b) **Comparison:** cruddier
            Less: wagons
            More: crossovers
            Dimension: cruddier (Sentiment Expression instance)

    In sentences with a two or more word comparison, the adjective(s) is marked as the Sentiment Expression and the adverbs *more* or *less* are marked as the comparison word. The Dimension of 'more' or 'less' would be the Sentiment Expression adjective. See 138 below. If it is a noun being compared, then that noun is marked as an Entity Mention instead of Sentiment expression, but will still be the Dimension of the Comparison. This can be seen in 139.

(138) [Instead of the Scion xB] I would prefer a <u>more</u>modest looking station wagon from Subaru or Volkswagen, but I would miss the attention the Toyota [Scion] brings.

    (a) **Sentiment Expression:** modest
            Target: station wagon

    (b) **Comparison:** more
            Less: Toyota   **More:** station wagon
            Dimension: modest

(139) In terms of pricing, the A4 Avant [convertible] is about 3K <u>more</u> than a comparable A4 sedan.

    (a) **Sentiment Expression:** 3K
           EntityMentionType: EntityMention.Unit.Money

    (b) **Comparison:** more
           Less: A4 sedan
           More: A4 Avant
           Dimension: 3K

The last type of comparison is a 'same' comparison. These comparisons are expressed with words such as *like* or *both*. Whereas the other comparisons make one of the Mentions more of something and the other Mention less, these comparisons express both Mentions as similar or alike. In these cases, the 'same' box will be marked true (it defaults to false in all other comparisons). The Dimension of comparisons like these is not needed but can be included if it appears in the sentence. It is also not important which Mention goes into the More or Less spaces, since they are equivalent. See example 140. For *as ... as* constructions the first *as* is marked as the Comparison word and the second *as* is left alone., shown in Example 141. If there is only one *as* then that one is marked as the Comparison.

(140) The bad news for Suzuki is that the Kizashi competes against the Volkswagen Jetta, which offers a <u>similarly</u> upscale feel.

    (a) **Sentiment Expression:** upscale

    (b) **Comparison:** similarly
           Less: Volkswagen Jetta
           More: Kizashi
           Dimension: upscale
           Same: *True*

(141) While it [the Ford Flex] may not be <u>as</u> tall as full-size SUVs, it surely provides the same amount of room, if not more.

    (a) **Sentiment Expression:** tall
           PriorPolarity: positive

    (b) **Comparison:** as
           Less: it
           More: SUVs
           Dimension: tall
           Same: *True*

(142) My Honda Civic is good <u>as</u> your car.

    (a) **Sentiment Expression:** good
           PriorPolarity: positive

    (b) **Comparison:** as
           Less: My Honda Civic
           More: your car
           Dimension: good
           Same: *True*

15   Committer

Committers increase or decrease the author's committment to an opinion. In other words, a commiter solidifies a statement's sentiment. Committers have a CommitterTarget slot which is filled by an Sentiment Expression, a Neutralizer or an Intensifier, and a CommitterDirection slot, which can be set to Upward (for increased committment) or Downward (for decreased committment).

In the example below, the author gives a positive sentiment about the car but hedges this opinion by saying *in my book*.

(143) In my book, this is a great car.

    (a) **Committer:** In my book
        CommitterTarget: great
        CommitterDirection: downward

Other examples of committers are: *definitely, certainly, kind of, sort of, appears to, might be.*

NOTE: Words like *looks, feels, sounds* can have two different meanings: one which means to appeal to the senses (visual/tactile/aural etc.) as in example 144 and the other which means to have a general conviction of something, as in example 145. In the latter example, the word *feel* can be replaced by *seem* and still retain the same meaning. When this is the case, mark the word as a Committer with a downward direction but in examples like the former annotate it as a Mention.CarPart/FeatureOf.

(144) I like the <u>look</u> of this car.

    (a) **Mention.CarPart:** look
        FeatureOf: car

    (b) **Sentiment Expression:** like
        PriorPolarity: positive
        Target: look

(145) These cars <u>look</u> nice, let's see how they drive.

    (a) **Committer:** look
        CommitterDirection: Downward
        CommitterTarget: nice

    (b) **Sentiment Expression:** nice
        PriorPolarity: positive
        Target: cars

Acknowledgements

16   Bibliography and Related Work

LDC. (2005). ACE (Automatic Content Extraction) English Annotation Guidelines for Entities. http://projects.ldc.upenn.edu/ac
EntitiesGuidelines_v5.6.1.pdf: Linguistic Data Consortium.

Janyce Wiebe, Theresa Wilson & Claire Cardie (2005). Annotating expressions of opinions and emotions in language. Language Resources and Evaluation, volume 39, issue 2-3, pp. 165-210.