~~In constructing my loops, I could exploit order of occurrence in order to render it more efficient, e.g. dobj occurs before nmod.~~

**~~The ‘dep’ relation~~**

**~~Brackets~~**

~~Parsers don’t handle brackets very well. Disrupt the dependencies. Inessential information.~~

~~Split sentences by semicolons (~~*~~Transfer batter to pan; smooth top~~*~~) and make sure to append a period and change to capital letter.~~

~~Issues with slash:~~ *~~Mix in walnuts and 1/4 cup chocolate chips~~* ~~(since~~ *~~/~~* ~~is a meta-character!). Um, actually~~ *~~¼~~* ~~works??? We butter three 9-inch-diameter cake pans with 1 1//2-inch-high sides. Here all of a sudden they need to be double!! I wonder if I have to split by – and remove the first dash sign? Nope.~~ **~~It is the leading 1 that causes the problem!~~**

* ~~If the sentence contains a / and the word just before is a number, then do this simple hack: multiply by base so that it becomes 3/2 and later change back!~~

**~~Until-information~~**

**~~Adverbs~~**

~~We assume adverbs to be relatively meaningless. We also discard speed information.~~

**’We’ tags**

~~Sometimes the first word (that is actually a verb) is tagged as an adjective (JJ). For example:~~ *~~Top~~* ~~in~~ *~~Top with remaining macaroon~~*~~.~~

~~Sometimes NLTK tags it as VB and Stanford as NNP, e.g.~~ *~~Line the bottom and sides of the baking pan~~*~~.~~

~~Sometimes NLTK tags a verb as IN, for example in~~ *~~Add the vanilla~~*~~, yet Stanford tags it as a verb.~~

**Verb tagging**

~~Before discarding VBD, check so that VBD does not equal present terms, e.g.~~ *~~beat~~* ~~is~~ *~~beat~~* ~~for both tenses.~~

~~Stanford~~ *~~and~~* ~~NLTK sometimes make blatant mistakes, e.g. tagging~~ *~~powdered~~* ~~as present tense in~~ *~~Beat cream cheese, powdered sugar, and vanilla in a large bowl~~*~~, possibly as a result of mistaking~~ *~~beat~~* ~~as past tense.~~

~~Sometimes the Stanford POS tagger refuses to identify a verb as such, even when using the “We” hack.~~

~~The “we” hack does not cover internal verbs. E.g.~~ *~~Put oven racks in upper and lower thirds of oven and preheat oven to 300F~~* ~~will tag~~ *~~preheat~~* ~~as a noun, as will NLTK.~~

~~Another example of internal verbs: in~~ *~~In a medium bowl, whisk together the flour, oats, coconut, baking soda, and salt~~*~~, NLTK will tag~~ *~~whisk~~* ~~as a noun, but Stanford will tag it as a verb. This is a fairly common occurrence.~~

* ~~If first word is~~ *~~In~~*~~, and the word after the first comma is a noun (according to Stanford), then include it as a verb.~~

~~Another common pattern is~~ *~~Using…., verb~~* ~~as in~~ *~~Using electric mixer, beat sugar.~~*

* ~~Check for that as well. Also, discard~~ *~~using~~* ~~so that it is not featured as a verb, but use the tool coming next as a noun associated with next-coming verb.~~

~~Another example of an internal is~~ *~~Add the chips and mix until just combined~~*~~. Here,~~ *~~mix~~* ~~is tagged as a noun. Another example is~~ *~~Butter pan, then spray with nonstick spray~~*~~.~~

* ~~Again, the only solution I see is doing like in the~~ *~~bowl~~* ~~situation: check if the noun also exists in the verb ontology, and if so, annotate it as a verb.~~

~~Elsewhere, nouns are tagged as verbs, for example~~ *~~Transfer batter to pan~~*~~, where~~ *~~pan~~* ~~is tagged as a verb.~~

~~Sometimes, true verbs (accurately tagged by NLTK) are interpreted as adverbs by Stanford.~~ *~~Bake cupcakes until toothpick inserted comes out clean~~* ~~gives (('cupcakes', 'VBZ'), 'advmod', ('bake', 'RB')). Another example is~~ *~~Chill uncovered overnight~~* ~~which gives ('uncovered', 'VBD'), 'advmod', ('chill', 'RB')).~~

* ~~Somehow compare tags~~

~~For example,~~ *~~Simmer dried apricots~~* ~~will tag~~ *~~Simmer~~* ~~as a noun, and in~~ *~~We simmer dried apricots~~* ~~it is tagged as an adverbial modifier. The problem is that~~ *~~dried~~* ~~is comprehended as a verb rather than serving an adjectival function.~~

*We heat almonds in a skillet* will tag *heat* as a noun, with *heat almonds* as a compound noun. This will thus be neglected in the first round.

* ~~Disregard verbs in VBD (i.e. verb past tense).~~

**~~Problem:~~** *~~Just before serving~~* ~~would tag~~ *~~serving~~* ~~as a verb to pursue.~~

* ~~Solution: only consider VBG verbs if it isn’t the first verb. OR: if the first verb is VBG, disregard it.~~

~~However, NLTK’s own part-of-speech-tagger correctly identifies~~ *~~simmer~~* ~~as a verb and~~ *~~dried~~* ~~as an adjective.~~

~~This suggests a possible solution:~~

* ~~Begin by scanning using NLTK for verbs~~

~~It is not always the case that NLTK is more verb-sensitive than Stanford.~~ *~~Add confectioners sugar~~* ~~will in NLTK tag~~ *~~Add~~* ~~as a noun, but in Stanford as a verb. However, NLTK picks it up after the~~ *~~We~~* ~~hack.~~

~~Sometimes the misunderstandings are subtle:~~ *~~Beat together mascarpone and cream until thick and smooth~~*~~, where~~ *~~smooth~~* ~~is tagged as a verb. In this particular case it is harmless.~~

* ~~Is it always harmless?~~

Sometimes there are verbs we don’t want. In *Continue to beat until whites hold stiff, glossy peaks, about 3 minutes*, the verb *hold* will be accurately tagged as a verb.

* However, we could exclude it if it figures in a **nsubj** relation possibly (*nsubj(hold, whites)* ) UNLESS it is part of a complex verb (*Let stand*, for example). Also, unless the thing it is “nsubj” with is ‘We’ because then it is due to the ‘We’ hack! That’s because then the verb is performed by an external entity rather than the recipe reader.

Another example is *sift in remaining sugar*, where *remaining* is tagged as a verb. However, it has a relation with *sift* as adverbial clause modifier ('sift', 'VB'), 'advcl', ('remaining', 'VBG')) and with *sugar* as (('remaining', 'VBG'), 'xcomp', ('sugar', 'NN')).

* Ignore verbs that are part of a ‘**xcomp**’ where the other one is a NOUN (open clausal complements are when words don’t have a subject), however, include the NOUN (since we want the sugar).
* Except sometimes it figures as an adjectival modifier: *top with remaining macaroon* (('compote', 'NN'), 'amod', ('remaining', 'VBG')). These must also be ignored somehow.

*Divide batter among cups, filling 1/3 full* gives (('filling', 'VBG'), 'xcomp', ('full', 'JJ')) where xcomp joins a verb with an adjective. *Whisking to dissolve sugar.* ('whisking', 'VBG'), 'xcomp', ('dissolve', 'VB')), (('dissolve', 'VB'), 'mark', ('to', 'TO'))

* Should I check for marker “to”?
  + Not if the verb figures in a **dobj:** (('dissolve', 'VB'), 'dobj', ('sugar', 'NN'))

*Divide batter among pans* gives (('divide', 'VBP'), 'dep', ('batter', 'VB')), (('batter', 'VB'), 'nmod', ('pans', 'NNS')). NLTK interprets *batter* as an adverb so that doesn’t help either.

* Check if a verb is a noun *BUT DON’T NEGLECT THE NOUNS CAUGHT VIA NMOD.*

Sometimes verbs that act as adverbial modifiers are caught. For example *Transfer the brownies to a cutting board*, tags *cutting* as a VBG, but it features in ('board', 'NN'), 'amod', ('cutting', 'VBG')).

* Discard verbs that figure in ‘amod’ relations.

~~Sometimes both Stanford and NLTK tags a verb as a noun.~~ *~~Cover the bowl with plastic wrap and refrigerate for at least 30 minutes or overnight~~*~~. Here~~ *~~refrigerate~~* ~~is tagged as a noun.~~

* ~~We must search through the verb ontology and annotate it as a verb before proceeding to make sure~~ *~~30 minutes~~* ~~is attached to it.~~

~~If the verb is “bring” it has a high likelihood of being in~~ *~~Bring to simmer~~* ~~or~~ *~~Bring to boil.~~* ~~In the sentence~~ *~~Bring cream and sugar to simmer in saucepan~~* ~~it gives (('bring', 'VBP'),~~ **~~'advcl'~~**~~, ('simmer', 'VB')), (('simmer', 'VB'),~~ **~~'mark', ('to', 'TO')).~~**

* ~~Check this pattern specifically.~~

**Properly discarding inessential information**

~~Recipe sentences can contain clauses at risk of being interpreted as high-information action verbs. For example:~~ *~~We stir from time to time until the butter is melted and the mixture is smooth and hot enough that you want to remove your finger fairly quickly after dipping it in to test.~~* ~~Here we have~~ *~~want, remove~~* ~~and~~ *~~dipping~~* ~~as verbs (~~*~~test~~* ~~inaccurately tagged as~~

* ~~If ’you’ is the subject, ignore it. ('want', 'VBP'),~~ **~~'nsubj'~~**~~, ('you', 'PRP')). Then ignore any verbs in a relation with ‘want’ or any of its extensions.~~

~~Another example is~~ *~~We stir in the nuts, if using~~*~~, where~~ *~~using~~* ~~is tagged as a VBG and thus included. However, it is an adjective according to NLTK.~~

* ~~If something is a verb in Stanford but adjective in NLTK, discard it.~~

**~~Verbs not acting on concrete objects~~**

~~Even if the~~ **~~dobj~~** ~~object is neither an ingredient nor aggregate include it and annotate using text. For example,~~ *~~Reduce the speed to medium-low~~* ~~has~~ *~~speed~~* ~~as object and~~ *~~medium-low~~* ~~as nmod.~~

**Noun tagging**

~~Sometimes nouns are tagged incorrectly. For example, Stanford tags~~ *~~preserves~~* ~~as an adverb in~~ *~~We stir in preserves, then cool completely~~*~~. However, NLTK tags it as a noun, so the “unfetched noun” technique would pick it up.~~

**Subject and Object**

*~~Turn off oven~~* ~~has (('Turn', 'VB'), 'dobj', ('oven', 'NN')).~~ *~~Let macaroons stand~~* ~~has (('stand', 'VB'), 'nsubj', ('macaroons', 'NNS')). I think this is when the verb has a clausal complement (‘~~**~~ccomp’~~**~~).~~

~~Actually, exclude nsubj – I think it will lead to errors, as in~~ *~~until whites hold stiff peaks~~* ~~becomes (('hold', 'VBP'),~~ **~~'nsubj'~~**~~, ('whites', 'NNS')) and more.~~

~~HOWEVER:~~ *~~Beat cream cheese~~* ~~gives (('beat', 'VBD'), 'nsubj', ('cheese', 'NN')). But ‘beat’ is VBD, which won’t get captured!~~

* ~~Let cheese be picked up by “neglected nouns”~~

~~It is important to make sure that one of the words in~~ *~~dobj~~* ~~is a verb, because it is the case that~~ *~~dobj(almonds,minutes)~~*~~. Also make sure that the other is a noun and not for example a cardinal number, as ('spread', 'VB'), 'dobj', ('one', 'CD').~~

~~One problem is that units are sometimes caught as objects, e.g.~~ *~~dobj(Add,minutes)~~*~~.~~

* ~~Check if the noun object figures as part of a~~ **~~‘nummod’~~** ~~relationship.~~

~~Also investigate~~ **~~iobj~~** ~~relationships:~~ *~~We cool cakes 5 minutes~~* ~~gives (('cool', 'VBP'), 'iobj', ('cakes', 'NNS')), (('cool', 'VBP'), 'dobj', ('minutes', 'NNS')).~~

* ~~We have already dealt with the “measure being caught in dobj” issue, but now we should also give preference to~~ **~~iobj~~**~~.~~

~~In long-range relations, objects get lost, e.g.~~ *~~Trace 2 circles on 1 sheet and a third circle on a second sheet~~* ~~will have to be picked up on “neglected nouns” because it has no traceable connection to the verb in the Stanford output due to the conjunction being interpreted incorrectly.~~

~~If e.g.~~ *~~dobj(VB, PRP)~~* ~~where PRP is~~ *~~it~~* ~~or~~ *~~they~~*~~, for example, then it’s almost certain that you can substitute this for the last object of the previous sentence (if singular: NN or NNP, if plural: NNS or NNPS, depending on whether it is a proper noun or not).~~

~~If there is an ‘xcomp’ verb relation, then we need to use ‘nsubj’ for that, though this may have undesired consequences. See:~~ *~~Let the cookies cool on the baking sheet for 5 minutes, then transfer them to a wire rack to cool completely~~*~~. (('Let', 'VB'), 'xcomp', ('cool', 'JJ')), (('cool', 'JJ'), 'nsubj', ('cookies', 'NNS')). There won’t be a~~ **~~dobj~~**~~.~~

~~Beware that a caught object (via conjuncts) can be a verb falsely tagged as a noun. For example~~ *~~We add egg, sugar, salt, and vanilla and beat until almost smooth~~*~~, will give (('egg', 'NN'), 'conj', ('beat', 'NN')).~~

~~Sometimes the sentence is not very self-contained and inappropriate for sentence-by-sentence analysis. For example:~~ *~~Whisk first 5 ingredients in small bowl~~*~~. Discard nouns that are not in aggregate/ingredient ontology~~ *~~or~~* ~~recipe-specific ingredient-list. For example, the above sentence would only retain~~ *~~bowl~~*~~.~~

* ~~Discard nouns that are not in aggregate/ingredient ontology~~ *~~or~~* ~~recipe-specific ingredient-list. For example, the above sentence would only retain~~ *~~bowl~~*~~.~~

**~~Noun conjuncts~~**

~~Take conjuncts for nouns even though none is a~~ **~~dobj~~** ~~with a verb – instead attach them to the most recent verb. Basically look for~~ *~~conj(NN,NN)~~* ~~and then check for compounds and other conjuncts of the left noun. I.e. also for those that only figure in~~ **~~nmod~~** ~~relations.~~

~~BUUUT: NEGLECTED VERBS?~~

~~Example sentence where this is relevant is~~ *~~Add the vanilla and the eggs, one at a time, beating after each addition for 30 seconds~~*~~, where~~ *~~vanilla~~* ~~has no~~ **~~dobj~~** ~~with~~ *~~Add~~*~~.~~

**Compound words**

~~In order to differentiate between terms like~~ *~~cream~~* ~~and~~ *~~sour cream~~*~~, the parser must check whether a noun is part of a compound word. The Stanford Dependency Parser tends to accurately annotate these as ‘~~**~~amod’~~**~~, i.e. adjectival modifier.~~ *~~Baking~~* ~~functions as an adjectival phrase that alters the meaning of the noun phrase~~ *~~sheets~~*~~.~~

* ~~For every noun (NN or NNS), traverse the tree to check if it is part of an ‘~~**~~amod’~~** ~~relation. If so, join by a blank.~~

~~Sometimes, looking for compounds will cause measures to be included, e.g.~~ *~~compound(almonds,cup)~~* ~~because the text has~~ *~~reserve ¼ cup almonds~~* ~~and (('almonds', 'NNS'), 'nummod', ('1/4', 'CD')).~~

* ~~An ugly hack would be to check that the word it is a compound to isn’t among the most common unit measures.~~

~~(('cream', 'NN'), 'amod', ('fourth', 'JJ')), (('fourth', 'JJ'), 'nmod', ('mascarpone', 'NN'))~~

~~Compound nouns raise the issue of concatenation order: generally adjectival modifiers should be before the compound. E.g. in~~ *~~Reserved praline almonds~~* ~~we have~~ *~~amod(reserved,almonds)~~* ~~and~~ *~~compound(almonds,praline)~~*~~.~~

~~Compounds are generally accurate, but not always. For example~~ *~~Add the eggs one at a time~~* ~~gives~~ *~~compound(one,eggs)~~*~~. Another example is~~ *~~Beat in flour mixture alternately with milk~~* ~~which gives~~ *~~compound(alternately,flour)~~* ~~and~~ *~~compound(alternately,mixture).~~* ~~Another example is~~ *~~Beat butter and peanut butter in large bowl~~* ~~which gives~~ *~~dobj(beat,butter), compound(butter,butter), conj(butter,peanut),~~* ~~i.e. it is interpreted as a conjunction rather than compound.~~

* ~~When analysing fetched nouns, decompose into component words, and discard any part that is not in verb or ingredient list/repository. Basically “clean them up”.~~

~~There are instances where the parser seems completely thrown off by decidedly simple instances. For example,~~ *~~Mix powdered sugar and sour cream in bowl~~* ~~returns~~ *~~compound(cream, sugar)~~* ~~etc. There is no~~ **~~dobj~~** ~~or~~ **~~nsubj,~~**

*~~Whisk whipping cream and ¾ cup powdered sugar~~* ~~gives~~ *~~dobj(whipping,sugar)~~* ~~and~~ *~~compound(sugar,cream)~~*~~.~~

**Complex verbs**

~~Recipes are rife with phrases like~~ *~~Turn off~~* ~~and~~ *~~Let stand~~*~~. It is important that these are treated as a single entry. Usually only one of the words that works as “key” to other triplets.~~

* ~~When you first identify a verb, find which “complex verb” relations it is member of first, save the concatenation somewhere separate, and remember to ignore these as you continue to look for verbs.~~

*~~Turn off~~* ~~is tagged as a phrasal verb relation between a verb and particle (‘~~**~~compound:prt’~~**~~). In the event where one is VB/VBG and the other is RP.~~

*~~Let stand~~* ~~is tagged as a clausal complement (~~**~~‘ccomp’~~**~~). In the event when both are VB/VBG. However this also caught~~ *~~stir~~* ~~and~~ *~~cool~~* ~~in~~ *~~We stir in preserves, then cool completely~~*~~. It also caught~~ *~~Continue cooking~~*~~.~~

* ~~For every verb (VB or VBG), check if it is part of a compound or complement relation. If so, join by a blank.~~

~~This is not guaranteed to be correct:~~ *~~Line standard muffin pan with 12 liners~~* ~~gives (('line', 'VBP'), 'ccomp', ('pan', 'VB')). Here ‘pan’ is tagged as NN by NLTK, however.~~

* ~~Solution?~~

*~~Continue to beat~~* ~~gives (('continue', 'VB'),~~ **~~'xcomp'~~**~~, ('beat', 'VB')).~~ *~~Let cool completely on a rack~~* ~~gives ('Let', 'VB'),~~ **~~'xcomp'~~**~~, ('cool', 'VB')) and ('cool', 'VB'),~~ **~~'nmod'~~**~~, ('rack', 'NN')).~~

* ~~Thus make sure to pursue the second verb.~~

~~Problem with~~ **~~xcomp~~** ~~is that the two words may be insufficient to form a coherent phrasal verb. For example~~ *~~xcomp(repeat,make)~~* ~~does not capture~~ *~~Repeat to make the rest of the cake~~* ~~unless we include~~ *~~mark(make,to)~~*~~.~~ *~~Beat to blend:~~* ~~(('beat', 'VBD'), 'xcomp', ('blend', 'VB')), (('blend', 'VB'), 'mark', ('to', 'TO'))~~

* ~~Check if there is a mark. (“marker” = word introducing a finite clause subordinate to another clause).~~

~~(('stir', 'VBP'), 'xcomp', ('beginning', 'VBG')) is another example (~~*~~Stir in flour mixture, beginning and ending with flour mixture~~*~~). Another example: (('whisk', 'VBP'), 'xcomp', ('whipping', 'VBG'))~~

*~~Beat in eggs, 1 at a time, then vanilla:~~* ~~(('beat', 'VBD'), 'xcomp', ('vanilla', 'NN'))~~

~~Informative prepositions may go uncaptured, e.g.~~ *~~put on 2 baking sheets~~* ~~returns ('put', 'VBD'), 'nmod', ('sheets', 'NNS')), (('sheets', 'NNS'), 'case', ('on', 'IN')). We neglect these for now, but it is worth exploring the consequences of include the ‘case’ relation in the analysis.~~

~~Some errors are really difficult to explain and remedy, where the parser fails almost completely, for example~~ *~~We combine the butter, sugar, cocoa, and salt in a medium heatproof bowl and set the bowl in a wide skillet of barely simmering water~~* ~~has~~ *~~bowl~~* ~~interpreted as a verb and~~ *~~ccomp(combine, bowl)~~*~~, which throws off the functionality of the parser. The objects then figure as~~ *~~nsubj(bowl,butter)~~*~~.~~

~~Similar issue arose with~~ *~~Lift up the ends of the parchment or foil~~*~~, where~~ *~~foil~~* ~~was interpreted as a verb by Stanford but noun by NLTK.~~

* ~~The only ugly hack I can think of is excluding clausal complements that are tagged as noun by NLTK, e.g. ‘bowl’ is tagged as noun. Then, the other nouns will be fetched as neglected nouns.~~

**~~Verb conjunctions~~**

~~Verbs sometimes figure in conjunctions, for example:~~ *~~Continue cooking, stirring and tossing~~*~~. Stanford tags it as follows: (('continue', 'VB'), 'ccomp', ('cooking', 'VBG')), (('cooking', 'VBG'), 'conj', ('stirring', 'VBG')), (('cooking', 'VBG'), 'cc', ('and', 'CC')), (('cooking', 'VBG'), 'conj', ('tossing', 'VBG')). Thus, the right verb in the ccomp (if there is indeed a complement or particles!) is what I need to check conjunctions for. I should be able to ignore the~~ **~~‘cc’~~**~~. But actually it is better to make separate entries.~~

* ~~Don’t be concerned with conjunctions. Make separate entries for conjuncts.~~
* ~~Yeah, I still think this is the wisest idea but conjunctions are sometimes very long-range.~~

**~~Nominal modifiers~~**

~~Often, there is a lot of valuable information from prepositional complements or genitives, which by the Stanford Tagger are tagged as nominal modifiers (‘~~**~~nmod’~~** ~~- nouns that change the meaning of other nouns or clausal predicates). For example:~~ *~~sheets~~* ~~modifies~~ *~~cool~~* ~~and~~ *~~rack~~* ~~modifies~~ *~~sheets~~* ~~in the sentence~~ *~~Cool on baking sheets on racks~~*~~.~~

~~Another example is~~ *~~Beat together cream and Amaretto with cleaned beaters~~*~~, where~~ *~~beaters~~* ~~is a nominal modifier.~~

~~For diagram-purposes, we could disregard prepositions and the like, and focus on content words.~~

~~Here we would ask for the parser to do the following:~~

1. ~~Have you identified an action verb (and its dependencies)?~~
2. ~~If so, scan for ‘nmod’ with this verb. Save the modifier and append to noun-list.~~
3. ~~Scan for ‘nmod’ with the modifier until none left. Append all to noun-list.~~

~~Do not look for conjuncts to nominal modifiers, as these will probably be misleading!~~

~~This will lead to mistakes, e.g. (('beaters', 'NNS'),~~ **~~'nmod'~~**~~, ('speed', 'NN')) or (('macaroon', 'NN'),~~ **~~'nmod'~~**~~, ('manner', 'NN')). (('put', 'VBD'),~~ **~~'nmod'~~**~~, ('upper', 'JJ')).~~

* ~~Discard modifiers not found in dictionary (in this case~~ *~~speed~~*~~).~~

*~~Beat for 40 strokes with the wooden spoon or a rubber spatula~~*~~. Returns (('beat', 'VB'), 'nmod', ('strokes', 'NNS')) which provides the “link” to spoon (('strokes', 'NNS'), 'nmod', ('spoon', 'NN')). The important thing is that we later disregard~~ *~~strokes~~* ~~even if there is a diagram for it, because it figures in a ‘nummod’, and then we let ‘neglected noun’ mechanism pick up~~ *~~wooden spoon~~* ~~and~~ *~~rubber spatula.~~*

* ~~Discard if it is elsewhere as a unit~~

~~One question is whether we should check for ‘~~**~~nummod’~~** ~~along with ‘~~**~~nmod’~~**~~, because~~ *~~Add the flour mixture in three parts~~* ~~gives (('mixture', 'NN'), 'nmod', ('parts', 'NNS')) and (('parts', 'NNS'), 'nummod', ('three', 'CD')).~~

* ~~Don’t! Let “three parts” be annotated as numerical information, so later discard~~ *~~parts~~* ~~since it figures in a “nummod”.~~

~~Sometimes a~~ *~~nmod~~* ~~has already been used by a verb, even though it also refers to another verb. For example,~~ *~~Stir in flour mixture with sour cream in 3 additions, beginning and ending with flour mixture~~*~~. It gives (('beginning', 'VBG'), 'conj', ('ending', 'VBG')), (('beginning', 'VBG'), 'nmod', ('mixture', 'NN')) but since we have decided to disregard verb conjunctions, it will be disregarded that~~ *~~mixture~~* ~~also modifies~~ *~~ending~~*~~.~~

* ~~If there is a fetched ingredient afterwards, even though it has “picked up” by another verb, attach it anyway to the last verb.~~

LET’S DO CARDINALS OVER THE WEEKEND

**~~Cardinal numbers~~**

~~Should they map onto verbs or ingredients) Verbs! But one-to-many or one-to-one?~~

* ~~If several cardinal, check for closest ingredient.~~

~~Fairly straightforward. Numbers are tagged as ‘~~**~~CD’~~** ~~and ‘~~**~~nummod~~**~~’, e.g. nummod(minutes,10) or nummod(hour,1). But how to associate it with a verb?~~

*~~Cool completely on baking sheets on racks, about 1 hour~~*~~.~~

~~(('rack', 'NN'), 'appos', ('hour', 'NN'))~~

*~~Let macaroons stand in oven 10 minutes.~~*

~~(('oven', 'NN'), 'dep', ('minutes', 'NNS')). (('compote', 'NN'), 'dep', ('cup', 'CD')). However, there is no general way for finding referent from Stanford output.~~

~~The Stanford parser does not catch numerical information in verbal form. For example,~~ *~~One third of compote~~* ~~will have~~ *~~third~~* ~~as a noun rather than cardinal. However: (('third', 'NN'), 'nummod', ('one', 'CD')), so as long as we look for~~ **~~‘nummod’~~** ~~we should be fine. It then figures as ('third', 'NN'),~~ **~~'nmod'~~**~~, ('compote', 'NN')). (('cream', 'NN'), 'amod', ('fourth', 'JJ')). Risk that cardinals are tagged as adjectives!~~

~~(('one', 'CD'),~~ **~~'advmod'~~**~~, ('fourth', 'JJ')) (('one', 'CD'),~~ **~~'nmod'~~**~~, ('cream', 'NN')), (('third', 'NN'), 'nummod', ('one', 'CD'))~~

* ~~In finding cardinals, check for whether~~

~~We also have to check for compounds: (('cup', 'CD'), 'compound', ('1/2', 'CD')). Is there a compound of two CD? In that case, concatenate.~~

~~And sometimes simply a nmod with a verb: ('preheat', 'VBP'), 'nmod', ('350F', 'CD')) for~~ *~~Preheat oven to 350F.~~*

~~Sometimes, there are subtleties like~~ *~~Beat in eggs 1 at a time~~* ~~that gives~~ *~~nummod(eggs,1)~~* ~~but does not mean “Beat in 1 egg”.~~

* ~~Check that the noun isn’t in plural while the CD is “one” or “1”.~~

~~Sometimes cardinal numbers are objects:~~ *~~Divide batter among cups, filling 1/3 full~~* ~~gives ('filling', 'VBG'), 'dobj', ('1/3', 'CD')), (('filling', 'VBG'), 'xcomp', ('full', 'JJ')).~~

* ~~Not sure which action to take? There is no nummod! Include CD even if no measurement unit.~~

~~Sometimes cardinal nummods are directly attached to the ingredient, but the unit is connected via amod. For example:~~ *~~Place 1 heaping tablespoonful of cream cheese filling in each center~~* ~~gives (('cheese', 'NN'), 'nummod', ('1', 'CD')), (('cheese', 'NN'), 'amod', ('heaping', 'JJ')), (('cheese', 'NN'), 'amod', ('tablespoonful', 'JJ'))~~

* ~~“Tablespoonful” and “heaping” will be rejected as a result of failing to match. Yet I am not quite sure what to do.~~

~~Sometimes, attaching a cardinal number to the most recent verb will be misleading. For example~~ *~~Bake macaroon layers, switching position of baking sheets halfway through cooking, until macaroons are crisp and edges are just barely pale golden, about 25 minutes~~*~~, we would wish for~~ *~~25 minutes~~* ~~to be attached to~~ *~~Bake~~*~~, not~~ *~~Switching~~*~~. There seems to be no easy way to “backtrack” and find verb referent.~~

* ~~Attach it to the most recent VBP-verb instead of VBG.~~

~~Difficulties are introduced by hyphenate descriptions, such as~~ *~~½-inch-thick rounds~~*~~. Here~~ *~~½-inch-thick~~* ~~is tagged as simply an adjective.~~

* ~~In order to annotate it as quantitative, we could check whether any character is a number.~~

~~Disjunctions could potentially be dealth with if we have~~ *~~conj(CD,CD)~~*~~. (('squares', 'NNS'), 'nummod', ('16', 'CD')), (('16', 'CD'), 'cc', ('or', 'CC')), (('16', 'CD'), 'conj', ('25', 'CD'))~~

~~Sometimes there were disjunctions or options, like~~ *~~2 to 3 minutes~~*~~. In the Stanford output, this will translate to a single~~ **~~‘nummod’~~**~~, so 3 minutes should be fine.~~

~~Not all numbers are tagged as numerical modifiers. For example, in~~ *~~Preheat the oven to 325F~~*~~, we have~~ *~~nmod(preheat VB, 325F CD)~~*~~. However, it will go captured, but won’t be annotated as CD.~~

* ~~Whenever something is tagged as a CD, even though it is an nmod, annotate it as numerical.~~

~~Sometimes there are several numerical quantities associated with a single action verb, as a result of the verb having several associated ingredients.~~ *~~Stir in flour mixture in 4 additions alternately with sour cream in 3 additions~~* ~~has both 3 and 4 as CDs related to~~ *~~additions~~*~~.~~

* ~~One solution would be to check the latest ‘nmod’ that includes the unit, e.g.~~ *~~(('mixture', 'NN'), 'nmod', ('additions', 'NNS')), (('additions', 'NNS'), 'case', ('in', 'IN')), (('additions', 'NNS'), 'nummod', ('4', 'CD'))~~* ~~and~~ *~~(('cream', 'NN'), 'nmod', ('additions', 'NNS')), (('additions', 'NNS'), 'case', ('in', 'IN')), (('additions', 'NNS'), 'nummod', ('3', 'CD'))~~*~~.~~

~~Sometimes the measure becomes part of the compound, for example~~ *~~Stir in 1 cup chocolate chips~~* ~~gives (('chips', 'NNS'), 'nummod', ('1', 'CD')), (('chips', 'NNS'), 'compound', ('cup', 'NN')), (('chips', 'NNS'), 'compound', ('chocolate', 'NN')). The fact that~~ *~~nummod(noun in plural, 1)~~* ~~means that something is dodgy.~~

* ~~Check if noun is an ingredient. If so, choose instead the next-coming noun.~~

~~Sometimes it is a compound instead of a nummod, for example~~ *~~Bake cake until tester inserted near center comes out clean, about 1 hour~~*~~. This gives ('hour', 'NN'), 'compound', ('1', 'CD')).~~

* ~~Check for CD relations more generally.~~

~~Problems~~

~~Multiple numbers: certain numerical specifications will be neglected as a result of this.~~

~~Consider the sentence~~ *~~Butter three 9-inch-diameter cake pans with 1 ½-inch-high sides~~*~~.~~

**~~Fetching neglected verbs and nouns~~**

~~In analysing words tagged as nouns, there is a likelihood that it is actually a verb. Therefore, check if it is contained in the actions-list. If so, let this be the one you MIGHT attach the cardinal number to!~~

~~We need to find compounds for these as well.~~

**~~Lemmatizing… mapping to ontology~~**

~~Lemmatizing is a relatively expensive operation, compared with stemming and naively removing “s”.~~

~~Check that some measure hasn’t been concatenated with the noun.~~