Transformation Patterns

May 2019

1 Clausal Disembedding

1.1 Coordinate Clauses

Rule	Tregex pattern	Extracted sentence
CoordinationExtractor	ROOT <<: (S < (S $.. [S]))	[S].
Example: The Assistant Attorney in Orlando investigated the modeling company, *and decided that they were not doing anything wrong, and after Pearlman's bankruptcy, the company emerged unscathed and was sold to a Canadian company.*		

Table 1: The transformation rule pattern for splitting coordinate clauses. The framed pattern represents the part of a sentence that is extracted from the source and turned into a stand-alone sentence.

1.2 Adverbial Clauses

1.3 Relative Clauses

1.4 Reported Speech

2 Phrasal Disembedding

2.1 Coordinate Verb Phrases

2.2 Coordinate Noun Phrases

2.3 Appositions

2.4 Prepositional Phrases

2.5 Adjectival and Adverbial Phrases

2.6 Lead Noun Phrases

Rule	Tregex pattern	Extracted sentence
Subordination-PostExtractor	ROOT <<: (S < (NP $.. (VP < +(VP) (**SBAR** < (S < (NP $.. VP))))))	S < (NP $.. VP).
Example: It is also used by those who commute into the city using Metrorail, *as the two systems are connected at the Government Center and Brickell stations.*		
Subordination-PreExtractor	ROOT <<: (S < (**SBAR** < (S < (NP $.. VP))) $.. (NP $.. VP)))	S < (NP $.. VP).
Example: *Though shorts are an option for many casual occasions,* they may also be inappropriate for more formal occasions, and a skirt or dress, such as a sundress, may be the best choice.		
PurposePost-Extractor	ROOT <<: (S < (NP $.. (VP < +(VP) (NP|PP $.. (S <<, (VP <<, /(T|t)o/))))))	*This* + Be + S <<, (VP <<, /(T|t)o/).
Example: In 1923 he joined Ireland's new Department of External Affairs and was sent to Geneva in 1929 *to act as Ireland's representative to the League of Nations.*		
PurposePre-Extractor	ROOT <<: (S < (S <<, (VP <<, /(T|t)o/)) $.. (NP $.. VP)))	*This* + Be + S <<, (VP <<, /(T|t)o/).
Example: *To support this claim,* he points out the actor is wearing a very large fake moustache and a hat, and when his moustache detached before the end of the film, he hid his face by an arm until the moustache was reattached.		
Subordination-PostPurpose-Extractor	ROOT <<: (S < (NP $.. (VP < +(VP) (**SBAR** < (S <<, (VP <<, /(T|t)o/))))))	*This* + Be + S <<, (VP <<, /(T|t)o/).
Example: A graduate student, Ian Gray, is researching the evolution of human eyes with Karen and Kenny, *in order to discredit creationists by proving that eyes have evolved.*		
Subordination-PrePurpose-Extractor	ROOT <<: (S < (**SBAR** < (S <<, (VP <<, /(T|t)o/)) $.. (NP $.. VP)))	*This* + Be + S <<, (VP <<, /(T|t)o/).
Example: *In order to cater for large items and fast loading,* the entire tail section was hinged, and could be opened using hydraulic actuators.		

Table 2: A selection of transformation rule patterns. A framed pattern represents the extracted part of a sentence. An underlined pattern refers to its referent. A pattern in bold will be excluded in the remaining sentence.

Rule	Tregex pattern	Extracted sentence
Non-Restrictive-Relative-ClausePrep-WhichWho-Extractor	ROOT <<: (S << (NP <, NP & < (/,/ $+ **(SBAR** <, **(WHPP** $+ S & <, IN=in & < − WHNP=whnp) & ?$+ /,/))))	*in + attachment phrase + whnp.*
Example: His family had been exiled to, and enrolled into Harvard University, *at which he studied archaeology.*		
Non-Restrictive-Relative-ClauseWhere-Extractor	ROOT <<: (S << (/.*/ < (NP|PP $+ (/,/ $+ **(SBAR** <, **(WHADVP** $+ S & <<: WRB) & ?$+ /,/)))))	*Extraction.*
Example: The noted communist, Sakhavu Kurumpakara Thankappan, was born and raised in Kurumpakara, *where a memorial dedicated to him is situated at the Udayonmuttam Junction.*		
Non-Restrictive-Relative-ClauseWhom-Extractor	ROOT <<: (S << (NP <, NP & < (/,/ $+ **(SBAR** <, **(WHNP** $+ (S <, NP & < − (VP ? < +(VP) PP=pp)) & <<: (WP <: whom)) & ?$+ /,/))))	*np + vp + attachment phrase + pp.*
Example: He is best known for his work with The Byrds, *whom he joined in September 1968 as a replacement for the band's original bass player Chris Hillman.*		
Non-Restrictive-Relative-ClauseWhose-Extractor	ROOT <<: (S << (NP < (NP $+ (/,/ $+ **(SBAR** <, **(WHNP** $+ S & <, (/WP \\$/ $+ /.*/)) & ?$+ /,/)))))	*Attachment phrase + "'s" + extraction.*
Example: Lacking a close relationship with his own estranged daughter, Dunn comes to establish a strong bond with Maggie, *whose own family cares little for her well-being.*		
Non-Restrictive-Relative-Clause-WhoWhich-Extractor	ROOT <<: (S << (NP <, NP & < (/,/ $+ **(SBAR** <, **(WHNP** $+ S & <<: WP\|WDT) & ?$+ /,/))))	*Attachment phrase + extraction.*
Example: However, while there, she discovered she could live on her income as a novelist and also met her husband, *who was completing his doctorate in physics.*		

Table 3: A selection of transformation rule patterns. A framed pattern represents the extracted part of a sentence. An underlined pattern refers to its referent. A pattern in bold will be excluded in the remaining sentence.

Rule	Tregex pattern	Extracted sentence
Participial-Middle-Extractor	ROOT <<: (S < (VP< +(VP) /VB.*/) & << (NP|PP <, (NP ?$+ PP & $ + + (/,/ $+ (VP [<, (ADVP|PP $+ VBG\|VBN) \| <, VBG\|VBN] & ?$+ /,/)))))	*Attachment phrase* + Be + *extraction.*
Example: The Metox, *named after its manufacturer*, was a high frequency very sensitive radar receiver manufactured by a small French company in occupied Paris, which could detect ASV transmissions from patrolling Allied aircraft.		
Shared-NPPost-Participial-Extractor	String participalNode = "(__ [== S=s | == (PP|ADVP < +(PP|ADVP) S=s)]) : (=s <: (VP <<, VBG|VBN))"; ROOT <<: (S < (NP $.. (VP < +(VP) (NP|PP $.. " + participalNode + ")))))	*Attachment phrase* + (Have) + Be + *extraction.*
Example: He served as chief judge from 1987 to 1994, *assuming senior status on November 2, 1995.*		
Shared-NPPre-Participial-Extractor	String participalNode = "(__=node [== S=s | == (PP|ADVP < +(PP|ADVP) S=s)]) : (=s <: (VP <<, VBG|VBN))"; ROOT <<: (S < " + participalNode + ") : (=node $.. (NP $.. VP))	*Attachment phrase* + (Have) + Be + *extraction.*
Example: *Founded in 1973,* the campus is situated on a hillside below the neighborhood.		

Table 4: A selection of transformation rule patterns. A framed pattern represents the extracted part of a sentence. An underlined pattern refers to its referent. A pattern in bold will be excluded in the remaining sentence.

Rule	Tregex pattern	Extracted sentence
Restrictive-Relative-ClauseWhom-Extractor	ROOT <<: (S << (NP <, NP & < (**SBAR** <, **(WHNP** $+ (S <, NP=np & < − (VP=vp ? < +(VP) PP=pp)) & <<: (WP <: whom)))))	*np + vp + attachment phrase + pp.*
Example:		
Restrictive-Relative-ClauseWhose-Extractor	ROOT <<: (S << (NP < (NP $+ (**SBAR** <, **(WHNP** $+ S & <, (/ WP\\$/ $+ /.*/))))))	*Attachment phrase + "'s" + extraction.*
Example: The rescue operation to reach Flight 608 was carried out by the Canadian Forces *whose plane spotted the downed aircraft and parachuted in nine search-rescue workers.*		
Restrictive-Relative-Clause-WhoWhich-Extractor	ROOT <<: (S << (NP <, (NP $ + + (**SBAR** <, **(WHNP** $+ S & <<: WP\|WDT) & ?$+ /,/))))	*Attachment phrase + extraction.*
Example: Ishak Belfodil is a Franco-Algerian football player *who currently plays for French club Olympique Lyonnais in Ligue 1.*		
Restrictive-Relative-Clause-Without-Relative-Pronoun-Extractor	ROOT <<: (S << (NP <, (NP $ + + (**SBAR** <: (S < (VP ? < (PP ? <: IN)))))))	*Attachment phrase + extraction.*
Example: In 1894 Verga moved back to the house *he was born in*, where he died of a cerebral thrombosis in 1922.		
Restrictive-Participial-Extractor	ROOT <<: (S < (VP < +(VP) /.*/=mainverb) & << (NP|PP <, (NP $+ (VP [<, (ADVP|PP $+ VBG|VBN) | <, VBG|VBN])) & [> (PP ! > S)| > (VP > S)]))	*Attachment phrase + Be + extraction.*
Example: The Muppets at Walt Disney World is a film *starring Jim Henson's Muppets at Walt Disney World,* and this was the last Muppets project that Jim Henson worked on before his death in 1990.		

Table 5: A selection of transformation rule patterns. A framed pattern represents the extracted part of a sentence. An underlined pattern refers to its referent. A pattern in bold will be excluded in the remaining sentence.

Rule	Tregex pattern	Extracted sentence
PreAttribution-Extractor	ROOT <<: (S < (S|SBAR|SBARQ $.. (NP=np [$,, VP=vp | $.. VP=vp]))); *vb* is an attribution-verb; Extraction ← text until *np* and text after *vp*	*This* + Be + *what extraction.*
Example: Witness memories don't get better with time, *she said in an interview with the International Herald Tribune.*		
Quoted-Attribution-PostExtractor	ROOT <<: (S < (NP $.. (VP < +(VP) (SBAR [,, /"/=start | <<, /"/=start] [.. /"/=end | << − /"/=end])))); Extraction ← text until *start* and text after *end*	*This* + Be + *what extraction.*
Example: *Pauli remarked sadly* "It is not even wrong".		
Quoted-Attribution-PreExtractor	ROOT <<: (S < (S|SBAR|SBARQ [,, /"/=start | <<, /"/=start] [.. /"/=end | << − /"/=end] $.. (NP [$,, VP | $.. VP]))); Extraction ← text until *start* and text after *end*	*This* + Be + *what extraction.*
Example:		
Subordination-PostAttribution-Extractor	ROOT <<: (S < (NP $.. (VP=vp < +(VP) SBAR=sbar))); *vb* is an attribution-verb; Extraction ← text until *sbar* and text after *sbar*	*This* + Be + *what extraction.*
Example: *Later, Ellis claimed that* the character was not in fact based on his father.		

Table 6: A selection of transformation rule patterns. A framed pattern represents the extracted part of a sentence. An underlined pattern refers to its referent. A pattern in bold will be excluded in the remaining sentence.

Rule	Tregex pattern	Extracted sentence
SharedNPPost-Coordination-Extractor	ROOT <<: (S < (NP $.. (VP < +(VP) (VP > VP $.. VP))))	NP + VP.
Example: After Pearlman's bankruptcy, the company emerged unscathed *and was sold to a Canadian company.*		

Table 7: The transformation rule pattern for disembedding coordinate verb phrases. A framed pattern represents the part of the input that is extracted and rephrased into a self-contained sentence. The underlined pattern refers to its referent.

Rule	Tregex pattern	Extracted sentence
PostListNP-Extractor	ROOT <<: (S < (NP $.. (VP << (NP=np1 < (NP $.. NP=np2))))); Extraction ← replace *np1* with *np2*	*Extraction.*
Example: Demola Aladekomo is *a computer engineer, technology pioneer, entrepreneur and philanthropist.*		
PreListNP-Extractor	ROOT <<: (S < (NP=np1 < (NP $.. NP=np2) $.. VP)); Extraction ← replace *np1* with *np2*	*Extraction.*
Example: At the end of the progression, 47 seconds into the song, the intensity of the synthesizer rises before *an organ, bass guitar, and piano* subsequently enter.		

Table 8: A selection of transformation rule patterns. A framed pattern represents the extracted part of a sentence. An underlined pattern refers to its referent. A pattern in bold will be excluded in the remaining sentence.

Rule	Tregex pattern	Extracted sentence
Non-Restrictive-Apposition-Extractor	ROOT <<: (S < (VP < +(VP) /.*/) & << (NP $+ (/,/ $+ (NP !$ CC & ?$+ /,/))))	*Attachment phrase* + Be + *extraction.*
Example: The president of Lithuania, *Antanas Smetona*, proposes armed resistance.		
Restrictive-Apposition-Extractor	(((PRP\\$|DT)\\ s)*(JJ\\ s)*((NN|NNS|NNP|NNPS)\\ s))+((CC\\ s)((PRP\\$|DT)\\ s)*(JJ\\ s)*((NN|NNS|NNP|NNPS)\\ s)) followed by a named entity	*Attachment phrase* + Be + *extraction.*
Example:		

Table 9: A selection of transformation rule patterns. A framed pattern represents the extracted part of a sentence. An underlined pattern refers to its referent. A pattern in bold will be excluded in the remaining sentence.

Rule	Tregex pattern	Extracted sentence
Prepositional-Attachedto-VPExtractor	ROOT <<: (S < +(S|VP) (VP < (PP $− NP|PP)) & < VP)	*This* + Be + PP.
Example:		
Prepositional-Initial-Extractor	ROOT <<: (S <, (PP ?$+ /,/ & $ + + VP))	*This* + Be + PP.
Example: *After his retirement,* he took charge as director of the Indian Institute of Science, holding this position from 1915 to 1921.		
Prepositional-MiddleFinal-Extractor	ROOT <<: (S < VP & << (/,/ $+ (PP ?$+ /,/)))	*This* + Be + PP.
Example: Founded as a Numidian colony in 4th to 5th century BC, it later became a Roman town in the province of Africa, *before its eventual abandonment around 9th to 10th century.*		

Table 10: A selection of transformation rule patterns. A framed pattern represents the extracted part of a sentence. An underlined pattern refers to its referent. A pattern in bold will be excluded in the remaining sentence.

Rule	Tregex pattern	Extracted sentence
Adjectival-Adverbial-Initial-Extractor	ROOT <<: (S <, (ADJP|ADVP $+ (/,/ $ + + VP)))	*This* + Be + ADJP|ADVP.
Example: *Meanwhile,* unemployment in France threw skilled workers down to the level of the proletariat.		
Adjectival-Adverbial-Middle-Final-Extractor	ROOT <<: (S < VP & << (/,/ $+ (ADJP\|ADVP ?$+ /,/)))	*This* + Be + ADJP|ADVP.
Example: Gustafsson lived a normal life until 2004, *almost 44 years after the event.*		

Table 11: A selection of transformation rule patterns. A framed pattern represents the extracted part of a sentence. An underlined pattern refers to its referent. A pattern in bold will be excluded in the remaining sentence.

Rule	Tregex pattern	Extracted sentence
LeadNPExtractor	ROOT <<: (S <, (NP $+ **(/,/** $+ NP & $++ VP)))	*This* + Be + NP.
Example: *Six days later,* NATO took over leadership of the effort.		

Table 12: The transformation rule pattern for splitting and rephrasing leading noun phrases. A framed pattern represents the part of the input that is extracted and turned into a stand-alone sentence. A pattern in bold will be deleted from the remaining sentence.