WordNet Troponymy and Extraction of "Manner-Result" Relations

Aliaksandr Huminski
Institute of High Performance
Computing, A*STAR, Singapore

`huminskia@iphc.a-star.edu.sg`

Zhang Hao
Nanyang Technological University
Singapore

`hao.zhang@ntu.edu.sg`

Abstract

Commonsense knowledge bases need to have relations that allow to predict the consequences of specific actions (say, if John stabbed Peter, Peter might be killed) and to unfold the possible actions for the specific results (Peter was killed. It could happen because of poisoning, stabbing, shooting, etc.) This kind of causal relations are established between manner verbs and result verbs: manner-result relations.

We offer a procedure on how to extract manner-result relations from WordNet through the analysis of the troponym glosses. The procedure of extraction includes three steps and the results are based on the analysis of the whole set of verbs in WordNet.

1 Introduction

WordNet (WN) as a database is widely used in variety of tasks related with extraction of semantic relations. Verbs in WN are organized hierarchically as troponym-hypernym relations. Meanwhile, the definition of troponym has something in common with the definition of a manner verb suggested by B. Levin and M. Rappaport Hovav.

We consider in more details both types of relations: troponym-hypernym and manner verb-result verb relations.

1.1 Troponym-Hypernym Relation

Verbs in WN are linked through different types of relations – antonym, cause, entailment – but troponym-hypernym relation is a basic and the most frequently found relation among verb synsets (Fellbaum and Miller, 1990). If a hypernym is a verb of a more generalized meaning, a troponym replaces the hypernym by indicating more precisely the manner of doing something. The troponym-hypernym relations are hierarchical (vertical). Therefore, it makes it possible to create a huge verb net with top synsets that represent the highest hypernyms and branches going down to the bottom with corresponding troponyms. The closer to the bottom, the more specific is the verb synset. There are no further clarifications between different types of troponymy in WN.

As a result, the manner relation is polysemic and many different semantic elements are hidden behind the label 'manner' (Fellbaum, 2010). It can be volume as in *talk-whisper*, speed as in *jog-run*, intensity of emotion as in *love-adore-idolize*, etc. The specific manner depends on the semantic field and corresponding dimension.

1.2 Manner Verbs and Result Verbs

The definition of troponym has something in common with the definition of a manner verb suggested by Beth Levin and Malka Rappaport Hovav (2010). They pointed out that a study of the English verb lexicon reveals that within particular semantic domains there can be verbs that describe carrying out activities – manners of doing; and there can be verbs that describe bringing about results. Manner verbs are *walk, jog, stab, scrub, sweep, swim, wipe, yell,* etc. Result verbs are *break, clean, crush, destroy, shatter,* etc.

There are 3 features of manner-result relations that make extraction of them so important for commonsense knowledge bases.

1) Manner verbs and result verbs are in causal relations: *stabbing* causes *killing*; *sweeping* causes *cleaning* and etc.
2) It is an empirical, not a logical causality with probability less than 100%. Actions represented by manner verbs can fail in achievement of desirable results:

I wiped the table, but it's not clean.
John shot Peter, but he survived.

3) It is a common situation when several manner verbs cause the same result verb: *sweeping*, *wiping*, *blowing* cause *cleaning*.

2 Troponym-Hypernym and Manner-Result Relations

In the WN glossary of terms[1], a troponym is defined as a verb expressing a specific manner elaboration of another verb. X is a troponym of Y if to X is to Y in some manner. Having this definition, the obvious question arises: if troponym is defined through the manner, can one state that troponym-hypernym relation equals in manner verb-result verb relation? In other words, is there any correlation between troponym-hypernym relation and manner verb-result verb? The general answer on this question is "no" since there are several types of correspondence that can be unfolded in WN:

1) troponym-hypernym relation can be equal "manner verb-manner verb" relation. For example, the verb *stroll* (walk + slow + relaxed) is a troponym for the verb *walk*. But both of them are manner verbs.
2) troponym-hypernym relation can be equal "manner verb-underspecified verb" relation. For example, the verb *walk* (move + by steps) is a troponym for the verb *move*. The verb *walk* is a manner verb, the verb *move* is underspecified: it is neither a path verb since it doesn't encode direction, nor a manner verb since it doesn't specify any particular manner. So, it is an underspecified verb taking into consideration that manner-result dichotomy does not fully and exhaustively classify verbs.
3) troponym-hypernym relation can be equal "result verb-result verb" relation. For example, the verb *fracture* (break into pieces) is a troponym for the verb *break* (destroy the integrity).
4) troponym-hypernym relation can be equal "manner verb-result verb" relation. For example, the verbs *stone*, *lapidate* (kill by throwing stones at) and *poison* (kill with poison) are troponyms for the verb *kill* (cause to die; put to death).

[1] https://wordnet.princeton.edu/man/wngloss.7WN.html

Now, we need to find out the way how to extract the 4th type of correspondence which represents exactly what we are looking for.

3 General Procedure to Extract Manner-Result Relations from WordNet

Manner-result relations are hidden in the WN verb hierarchy. We know for sure that this kind of relations is a subset of troponym-hypernym relations. However, there are not any explicit ways to extract them yet.

Our idea is that manner-result relations can be extracted from the set of troponym-hypernym relations if two conditions, applied to troponym-hypernym relation are valid:

1) The hypernym is a result synset;
2) In the glosses of its troponyms one of the two templates can be found: "V + by" or "V + with"; where V = hypernym.

For example, if we consider the result synset *{clean, make clean}* as a hypernym, some its troponyms have glosses that satisfy the patterns:

- *sweep* (clean by sweeping)
- *brush* (clean with a brush)
- *steam, steam clean* (clean by means of steaming)

In this case, it can be stated that *sweep, brush, steam, steam clean* are manner verbs for *clean* and the following causality can be constructed:

sweep, brush, steam, steam clean → *clean*

This idea is the basis of the general procedure for manner-result extraction. It includes 3 steps.

3.1 Extraction of Top Verb Synsets

There are 13789 verb synsets in WN 3.1 ordered by troponym-hypernym hierarchical relation.

At this stage, we need to extract synsets located on the top level of the hierarchy. This kind of synsets will be called further "top verb synsets".

The procedure of extraction is based on the following characteristic of the top verb synsets: they don't have any hypernyms, only troponyms. Using this, all the extracted 13789 verb synsets have been tested whether they have a hypernym. As a result, 564 top verb synsets have been extracted automatically.

3.2 Extraction of Top Result Verb Synsets

Within 564 top synsets we made a manual classification to extract only the result verb synsets. The classification revealed the following 5 classes.

1) one-level top synsets. This type of top synsets has only one level: the top verb synset itself. It is a substantial portion of top synsets: 203. Example: *admit* (give access or entrance to).
2) manner and underspecified verb synsets. Total number: 105. Example of the top manner verb synset: *splash* (strike and dash about in a liquid). Example of the top underspecified verb synset: {*travel, go, move, locomote*}.
3) state verb synsets. Total number: 69. Example of the top state verb: *lie* (be lying, be prostrate; be in a horizontal position).
4) second order predicates. Total number: 60. Second order predicates govern the other predicate. Example: {*begin, start*} (have a beginning, in a temporal, spatial, or evaluative sense).
5) result and change-of-state verb synsets. Total number: 127. We combine these 2 classes of verbs since, as it turned out, change-of-state verbs have manner verbs as troponyms. For instance, the verb *die* has a troponym synset {*suffocate, stifle, asphyxiate*} which obviously contains manner verbs. Example of the result verb synset: {*destroy, ruin*}.

We further analyze the 5th class only. Our assumption was that result verbs as hypernyms can have either result verbs or manner verbs as troponyms. But manner verbs as hypernyms cannot have result verbs as troponyms. They can only have manner verbs as troponyms. Following the assumption, the sequence of troponyms derived from the top result verb hypernym cannot have the subsequence of manner verb as a hypernym and result verb as a troponym. For example, the sequence of 4-level verbs with the top result verb and the bottom manner verb can have the following 3 possible distributions:

- result-result-result-manner
- result-result-manner-manner
- result-manner-manner-manner

The distribution of "result-manner-result-manner" is impossible.

The next step is extraction of manner verbs from the tree with result verb synset on the top.

3.3 Extraction of Manner Verbs through the Patterns in Glosses

At this stage, we look for the manner verbs for each result verb synset through the patterns "*V + with*" or "*V + by*" in the glosses of troponyms. If the synset doesn't contain any patterns we mark it as "NONE". If the synset contains at least one of the patterns we mark it with its gloss.

As a result, we get a sequence of marked synsets from the top verb synset to the bottom verb synset. If the sequence of all synsets or only the tail of it contains "NONE" we exclude the whole sequence or the tail accordingly from the consideration since there is no manner verbs there. The purpose is to extract all lower synsets that contain the patterns. The procedure of the extraction is automatic.

Following the assumption from 3.2 one can get different types of result-manner sequences. For example, for the top synset {*change, alter, modify*} we will get the following 3 sequences among many others:

{*change, alter, modify*}-NONE	{*damage*}-NONE	{*frost*}-damage by frost

The causality *frost → damage* can be made from this sequence, where *frost* is manner verb and *damage* is result verb.

{change, alter, modify}-NONE	{*damage*}-NONE	{*burn*}-damage by burning with heat, fire, or radiation	{*scald*}-burn with a hot liquid or steam

The causality is *scald, burn → damage*

{*change, alter, modify*}-NONE	{*indispose*}-NONE	{*hurt*}-NONE	{*injure, wound*}-NONE	{*trample*}-injure by trampling or as if by trampling

The causality is *trample → injure, wound, hurt*.

For each *n*-level synset one can get a restricted number of the valid sequences. For example, for each 3-level synset we can get only two valid sequences: "result-result-manner" as in

{*change, alter, modify*}-NONE	{*sharpen*}-NONE	{*whet*}-sharpen by rubbing, as on a whetstone

and "result-manner-manner" as in

{*damage*}-NONE	{*burn*}-damage by burning with heat, fire, or radiation	{*scald*}-burn with a hot liquid or steam

As a whole, for the different sublevels of the same top result synset, one can get full variety of valid n-level sequences:

{change, alter, modify}-NONE	{shape, form}-NONE	{tabulate}-Manner verb		
{change, alter, modify}-NONE	{shape, form}-NONE	{roll}-Manner verb		
{change, alter, modify}-NONE	{shape, form}-NONE	{draw}-Manner verb		
{change, alter, modify}-NONE	{shape, form}-NONE	{fit}-NONE	{dovetail}-Manner verb	
{change, alter, modify}-NONE	{shape, form}-NONE	{flatten}-NONE	{steamroll, steamroller}-Manner verb	
{change, alter, modify}-NONE	{shape, form}-NONE	{flatten}-NONE	{roll_out, roll}-Manner verb	
{change, alter, modify}-NONE	{shape, form}-NONE	{flatten}-NONE	{roll_out, roll}-Manner verb	{mill}-Manner verb

Table 1. Part of valid n-level sequences from {*change, alter, modify*} result synset.

To make the table more compact we replaced the glosses that match the patterns to the phrase "Manner verb". Figure 1. shows the Table 1. in the structural graphical form with glosses.
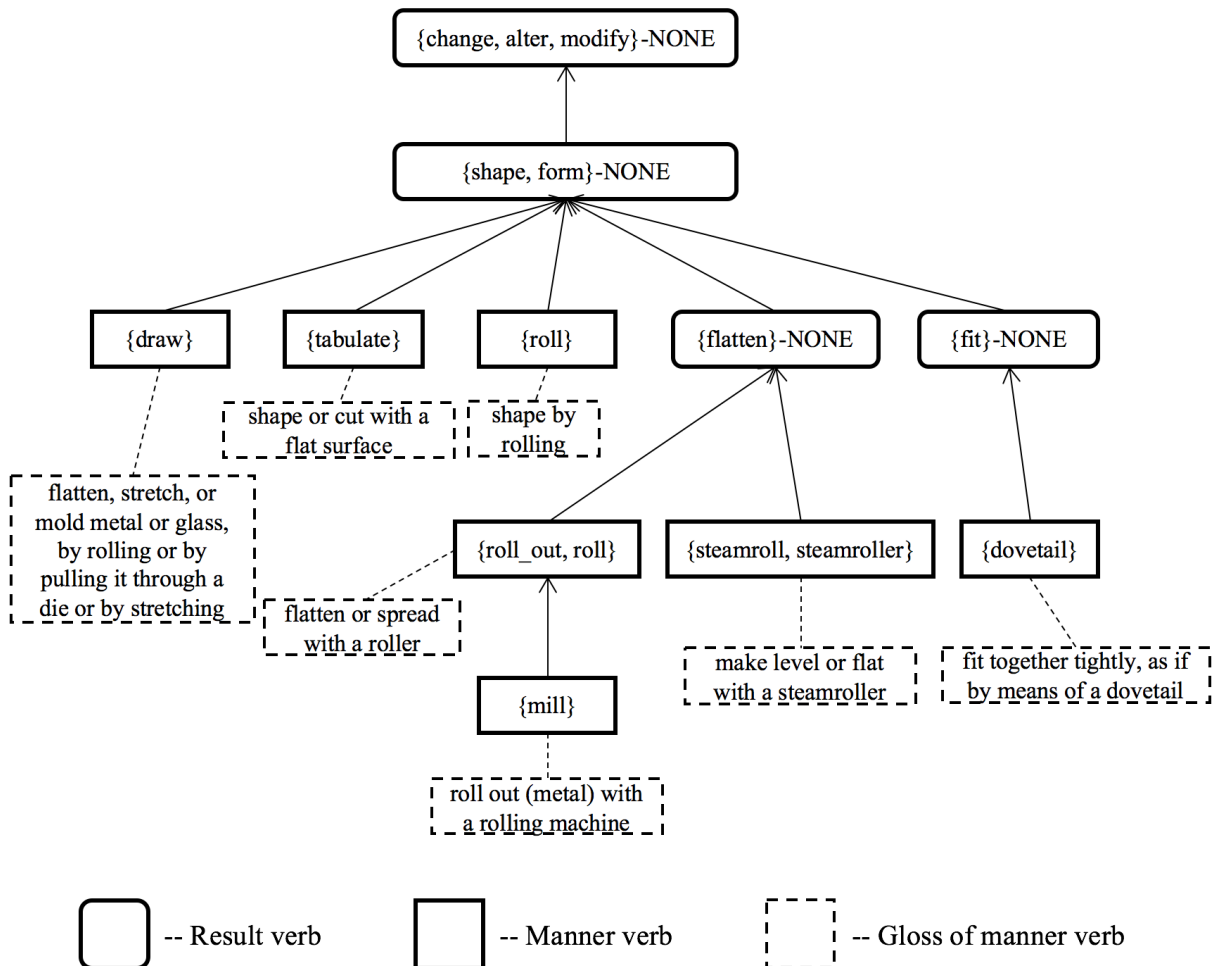

Figure 1. Visualization of the valid n-level sequences.

It is necessary to stress that each line in result-manner causal relation can contain both direct (*frost → damage*) and indirect (*scald → damage*) causality. Regardless of that, each line is considered as one specific type of causal relations.

After running all the top 127 result verb synsets and counting the lines we got the total number of 1541 lines. It means, 1541 manner-result causal relations have been extracted from WN.

4 Scope of the Results

To evaluate what is the scope of the results we compare them with another type of causal relations that is explicitly presented in WN 3.1: *cause*-relation.

Cause-relation refers to the relation between two verbs V_1 and V_2 where V_1 logically causes V_2 (Fellbaum, 1998). For example, the verb *kill* causes the verb *die*.

Running through 13789 verb synsets in WN 3.1 we automatically extracted 219 verb synsets that contain *cause*-relation. Among them there are 63 verb synsets that cause the same synset. In other words, there are 63 causal relations with absolutely identical left and right sides:

> *{dry, dry_out}* causes *{dry, dry_out}*
> *{lengthen}* causes *{lengthen}*, etc.

It happened because of polysemy in verb meaning. Synsets here are formally identical but represent different meanings of verbs. Since it is hard to use such kind of causality in applications, the real number of the verb synsets that contain *cause*-relation can be reduced to 156.

Comparison of 156 verb synsets containing logical *cause*-relation with 1541 non-logical (empirical) causal relations shows that the scope of the latter relations is significant.

5 Conclusions and Future Work

In this paper, we have described how to extract manner verb-result verb causal relations from WN. The procedure of extraction includes 3 steps: a) extraction of the top verb synsets (total 564), b) extraction of the result synsets and the change-of-state synsets among them manually (total 127), c) running automatically the algorithm "*V + by*" and "*V + with*" on 127 top synsets and getting 1541 types of manner-result causal relations. The results are considered as preliminary ones.

As future work, the algorithm can be elaborated by adding new patterns and tuning the original ones. For example, the change-of-state verb *die* has a troponym synset *suffocate*, *stifle*, *asphyxiate* (be asphyxiated; die from lack of oxygen) which clearly indicates the manner of dying but the gloss doesn't contain the patterns we are working with.

These types of extracted relations can be widely used in commonsense knowledge bases for the prediction of action consequences and unfolding the possible reasons for the results. Commonsense knowledge bases enriched by using this approach can be exploited in dialog systems and the other specific technologies and applications.

References

Fellbaum C. and G. Miller. 1990. *Folk psychology or semantic entailment? A reply to Rips and Conrad*. The Psychological Review. 97, 565-570.

Fellbaum, C. 2000. Autotroponymy. In Yael Ravin (ed). *Polysemy: theoretical and computational approaches*. New York, Oxford University Press.

Levin, B. and M. Rappaport Hovav. 2010. *Lexicalized Meaning and Manner/Result Complementarity*. Ms. Stanford University and The Hebrew University of Jerusalem.

Fellbaum C. 1998. *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.