Paranymy: Enriching Ontological Knowledge in Wordnets

Chu-Ren Huang, I-Li Su, Pei-Yi Hsiao, Xiu-Ling Ke
Institute of Linguistics, Academia Sinica
http://cwn.ling.sinica.edu.tw/huang/huang.htm
http://cwn.ling.sinica.edu.tw/
Research Questions

Coordinate terms are not all equal

• How to capture conceptual salient groupings among coordinate terms?
• How to solve the ‘ISA overload’ problem (Guarino 1999) while maintaining the original WordNet structure?
Outline

- Hyponymy, Taxonomy, and Paronymy
- Sister terms in wordnets
- Definitions and types of paronyms
- Practical examples in Chinese
- Conclusion
Hyponymy, Taxonomy, and Paronymy

• Sister terms in Princeton WordNet
  – Sister terms are coordinate words with same hypernym (also called “superordinate”)
• Lexical Semantic Theory (Cruse 1991)
  – ‘simple’ hyponymy: An X is a Y
  – taxonomy: An X is a kind/type of Y

Wordnet hyponymy contains both types of relations, though more dominated by taxonomy.
• Simple hyponymy
  – A house is one’s castle.
  – A child is a blessing.
  – A moose is a big deer. (from ShuoWenJieZi, AD 100-121) 麋，鹿之大者
Hyponymy, Taxonomy, and Paranymy

- Intuitively, paranymy is the relation among a conceptual cluster among coordinate terms.
- They may correspond to a semantic field
- Or a the exhaustive list of immediate sisters in a specific taxonomy.
  - Taxonomy not in Cruse’s definition but the conventional information science sense of a hierarchical classification system governed by a set of criteria (such as the taxonomy of animals).

Possible explicit solution to the ontological problem of ‘ISA’ overload, where ISA encodes many conceptually different relations (Guarino 1999a)
To Solve the ISA Overload

- We need to be able to describe WN coordinate terms with more precise and richer semantic representation of lexical conceptual structure and ontology
  - The classificatory criteria for those elements in order to define X as a C
  - The salient relations among those different elements in X
Sister terms in wordnets

• Coordinate terms sharing the same hypernym,
  – Is there an direct semantic relation connecting these terms?

• However, not all coordinate terms are equal
  – Seasons in a year: four or two?
  – Four seasons: spring, summer, fall, and winter
  – Two seasons: dry season and rainy/monsoon season

• WordNet
  – Season > Spring
  – Season > rainy season > monsoon
Other example

• Four cardinal compass points
  – East
  – West
  – South
  – North

• Conventional collocation
  – North/South pair or East/ West pair
• Northeast in English, vs.
• East-north in Chinese (dong1bei2 東北)
Definitions of Paronyms

• X is a *paronym* of Y iff
  – X and Y are coordinate terms, and
  – X stands in minimal semantic contrast with Y
  – [or] X and Y are the shared the same ‘is a kind of’ relation to their common hypernym
  – “seasons in a year”
    • Chun1/ xia4/ qiu1/ dong1
      “spring/ summer/ fall(autumn)/ winter”
    • Gan1 ji4 vs. yu3 ji4
      “dry season vs. rainy season’
Paranyms wi

WN structure

With overlaying classes added by paranymy
The Nature of Paranyms

- Paranyms are paradigmatic relations
  - Among and/or between coordinate terms
  - The ‘class’ name of a set of paranyms are virtual and not realized in wordnets (although they may exist in some specific taxonomy)

- Paranyms are complimentary to and not diagonal with hyponym/isa relation relations in WN.
  - It is crucial that no intermediate levels will be introduced in WN unless they are linguistically motivated
WN structure

With overlaying classes added by paranymy
One or Two Paranymic Classes

• Directions in Chinese
  – Si4mian4 “four directions”
    • Dong1/ xi1/ nan2/ bei3 “East/ West/ South/ North”
  – Ba1fang1 ‘eight directions’
    • Dong1/ xi1/ nan2/ bei3/
      Dong1nan2/ xi1bei3/ dong1bei3/ xi1nan2
      “East/ West/ South/ North/ SouthEast/ NorthWest/ NorthEast/ SouthWest”
Types of Paronym

- Contrary paronym
- Conventional paronym
- Overlapping paronym
Contrary Paranymy

• Contrary paranymy conforms to a condition
  – Each of a set of terms is related to all the others by the relation of incompatibility (Cruse 2004)
• Perceptonal and conventional paradigm
  – Perceptonal paradigm: based on human perception or senses
  e.g. fast/slow
• Antonym
**Conventional paranamy**

- **Parent addressing**
  - Registers are conventionally for the use of these terms in Chinese
  - Further classified into different groups, instead of directly placing them all under the same superordinate parent.
Conventional paranamy

- Spouse addressing: concepts Re-clustered based on the collocation
  - The contrary paronym of xian1 sheng1 “husband” is tai4 tai4 “wife” rather than qi1 zi5 “wife”
How to Identify Paranyms

- Identification of paranyms is a process of re-clustering sister terms
- This re-clustering can be applied to augment wordnets with both collocational and paradigmatic relations
A process of re-clustering sister terms

Sister terms

Re-clustered by using the same classificatory criterion or the collocation

New contrary paronyms
Overlapping Paranymy

• Note that paranym requires semantic contrast, but not mutual exclusiveness in semantic content

• Ex: Expressions of greeting, “good afternoon” and “good evening”
  – Both expressions are alternative in a certain time period: say the late afternoon, which indicates the overlap between the time periods denoted by these two sister terms
  — “afternoon” and “evening” (and night)
  -from “midnight”, “dawn”, to “morning”,
  in Chinese: ling2chen2, qing1chen2, qing1zao3, zao3shang4, shang4wu3
Overlapping Paranyms II

- A (usually rectangular) container:
  \texttt{xiang1 zi5} vs. \texttt{he2 zi5} “box”
  - Both \texttt{xiang1 zi5} and \texttt{he2 zi5} can be used to refer to “box”, but when we see a container for a diamond ring, we may call it \texttt{he2 zi5} rather than \texttt{xiang1 zi5}
  - Conversely, we call a container for a TV set \texttt{xiang1 zi5} rather than \texttt{he2 zi5}
Implementing Paranyms: Practical examples in Chinese

• The concept of collateral relatives by blood
  – Ge1 ge1, jie3 jie3, di4 di4 , mei4 mei4
    “elder brother, elder sister, younger brother, younger sister”
  – Same gender but different birth order (older or younger):
    • Ge1 ge1, di4 di4 “elder brother younger brother ”
    • Jie3 jie3, mei4 mei4 “elder sister younger sister ”
  – Different genders but having the same birth order (older or younger)
    • Ge1 ge1, jie3 jie3 “elder brother elder sister ”
    • Di4 di4, mei4 mei4 “younger brother younger sister ”
Solution: Indexing

• Hypernym: Siblings
  – ge1ge5 elder brother <1,2,3>
  – jie3jie5 elder sister <1,2,4>
  – di4di5 younger brother <1,3,5>
  – mei4mei5 younger sister <1,4,5>

• In addition, there is a mirror image sets of by conventional paranym
  – xiong1, jie3, di4, mei4
Conclusion

• There is more than one possible paradigmatic relations among coordinate terms
• Paronym allows explicit representation of all possible paradigmatic relations
• Hence overcomes the ISA overload problem without compromising WN’s hierarchical structure
  – Or adding ad hoc classificatory criteria (since only the grouping of paronyms, not why they are grouped together, is represented)
• Allows different (and possibly contradicting) taxonomies to be encoded
• Indirectly encodes linguistic information such as collocational and paradigmatic substitution information
• May be crucial in facilitating WN to ontology interface
Acknowledgements

• Thank all members of the Chinese Wordnet group, Academia Sinica
• Comments from GWC reviewers and CLSW participants (on an earlier paper)
Questions ?