

# **Studies in Computational Intelligence**

Volume 910

## **Series Editor**

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland

The series “Studies in Computational Intelligence” (SCI) publishes new developments and advances in the various areas of computational intelligence—quickly and with a high quality. The intent is to cover the theory, applications, and design methods of computational intelligence, as embedded in the fields of engineering, computer science, physics and life sciences, as well as the methodologies behind them. The series contains monographs, lecture notes and edited volumes in computational intelligence spanning the areas of neural networks, connectionist systems, genetic algorithms, evolutionary computation, artificial intelligence, cellular automata, self-organizing systems, soft computing, fuzzy systems, and hybrid intelligent systems. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution, which enable both wide and rapid dissemination of research output.

The books of this series are submitted to indexing to Web of Science, EI-Compendex, DBLP, SCOPUS, Google Scholar and Springerlink.

More information about this series at <http://www.springer.com/series/7092>

Tiansi Dong

# A Geometric Approach to the Unification of Symbolic Structures and Neural Networks

 Springer

Tiansi Dong  
ML2R Competence Center for Machine  
Learning Rhine-Ruhr, MLAI Lab, AI  
Foundations Group, Bonn-Aachen  
International Center for Information  
Technology (b-it)  
University of Bonn  
Bonn, Germany

ISSN 1860-949X                      ISSN 1860-9503 (electronic)  
Studies in Computational Intelligence  
ISBN 978-3-030-56274-8              ISBN 978-3-030-56275-5 (eBook)  
<https://doi.org/10.1007/978-3-030-56275-5>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2021

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG  
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

*The Sixth Law of Cognition: Spatial thinking  
is the foundation of abstract thought.*

—Barbara Tversky (2019) p. 72, p. 142

*to Elias, Sophia, Peiling, and my parents*

# Preface

It takes decades to edit a dictionary, to explain each word in terms of other words, and with examples. Alternatively, the neural approach simply learns a vector for each word from texts. We may encounter a new word that is not in our dictionary, but with the neural approach, we can always have a vector of this new word, search its neighbours in vector space, and guess the meaning (without guarantee). Neural-networks (Deep learning) are robust to noisy inputs, able to learn at a level of approximation from enough high-qualified data, but lack of explainability. In contrast, symbolic systems have a set of manually designed rules. This makes outputs explainable and the results guaranteed, but inputs are not robust to noisy inputs. Do the two kinds of systems talk about the same thing (“human intelligence”)? Neural people hope this, struggle for decades to design elegant neural-networks that can reach symbolic levels of reasoning, and land at a certain level of approximation. Would they change next year? To answer this question, we designed a very simple experiment. Given a manually edited dictionary that only has tree structured category relation among words, and given vectors of these words provided by well-designed neural-networks, what kinds of mechanism can let these vectors precisely encode the given tree structure? The mechanism that we found is to blow these vectors into balloons in higher dimensional space, so that spatial inclusion relations among balloons could precisely encode tree structured category relations. For symbolists, the configuration of these balloons is a spatial semantics of the tree structure. It seems hard for neural people to change next year, if they restrict vectors as processing objects of neural-networks and fix the dimension of outputs. In a survey of the neural-symbol debates in the literature, we found that the mechanism that we found is an open chapter to solve many questions raised in the debates. Without reservation, we collect these observations in this book, hoping that this could activate symbolists and promote the joint research with Deep learning researchers.

Bonn, Germany  
June 2020

Tiansi Dong

# Acknowledgements

First and foremost, I would like to thank Christian Bauckhage, to whom this work owes tremendously. From his reading group at Fraunhofer IAIS to the BMBF P3ML (01/S17064) and ML2R (01/S18038C) projects as well as his encouragement of my leadership role at IPEC (International Program of Excellence) on AI language lab, Christian is a great advocate of my work. I would like to express my deep appreciation to Armin B. Cremers (ABC), who hosted me as a guest scientist at B-IT in 2012 and urged me to pursue a Habilitation, and continues to be an inspiration. ABC has always challenged me to reach for the impossible. I am also indebted to Sören Auer for generously giving me a chance to work with EU H2020 OpenBudgets.eu project in September 2015. A big thank-you to many colleagues, Thomas Bode, Olaf H. Cremers, Martina Dölp, Thomas Fuchs, Bogdan Georgiev, Günter Kniesel, Peter Lachart, Jens Lehmann, Maria Maleshkova, Marie-Luise Liebegut, Sifa Rafet, Alexandra Reitelmann, Daniel Speicher, Thomas Thiel, Susan Tietz, Steffen Wrobel, and Jörg Zimmermann.

For the support of my early work in Spatial Calculus and of organizing Sino-German Symposium GZ1607 “Integrating Symbolic Representation with Numeric Representation for Commonsense Reasoning”, I am greatly indebted to Xiaoxing Ma (Nanjing University), and Juanzi Li (Tsinghua University), who also encourages me to challenge hard tasks, and generously shares her research experiences. I am also indebted to Achim Rettinger (University of Trier), Alexander Mehler (Goethe-University of Frankfurt am Main), Barbara Tversky (Stanford and Columbia Universities), Bo Zhang (Tsinghua University), Christian F. Hempelmann (Texas A&M University), Erhard Hinrichs (University of Tübingen), Jie Tang (Tsinghua University), Jun Zhao (Chinese Academy of Sciences), Ron Sun (RPI), Steffen Staab (University of Stuttgart), Xiaoming Fu (Georg-August-University of Göttingen), and Volker Tresp (LMU) for their fruitful discussions during and after the Symposium.

I am also very thankful to Yixin Yang (Northwestern Polytechnical University) for the invitation to Xi’an and the delicious Chinese meals and the delightful conversations in 2018, and to Charles Dee, Xiaoru Meng, Jun Pang, Junfeng Qin, Jianqiu Xu, and Cham Zhong for their friendship.



I am indebted to my parents for their sacrifice and for the happiness they bring whenever they come to visit from China, to members of EFG (Evangelische freie Gemeinde) Bonn, in particular, Margit Jehle, Matthias Jehle, Heidi Raschke, Ebi Raschke, Blanckarts Jürgen, and Ulf Beiderbeck, also to Shih-Kung Cheng and Rui Chen, for the joyful company.

Most importantly, I thank Peiling Cui, for her love and dedicated support and understanding. As a humor researcher, she shows me a third way of facing failure and frustration. Finally, my love to Sophia and Elias for our fun times together.

Bonn, Germany  
January 2020

Tiansi Dong

# Contents

<b>1</b>	<b>Introduction</b>	1
1.1	Two Paradigms in Artificial Intelligence	1
1.1.1	Symbolic Approach	1
1.1.2	Connectionist Approach	2
1.1.3	Complementary Strength	3
1.2	Recent Debates on Structure and Learning	4
1.3	Logic and Its Two Perspectives	4
1.4	Geometric Connectionist Machine	6
1.4.1	An Approach of Geometric Construction	6
1.4.2	Two Kinds of Geometric Connectionist Machines	9
1.5	Towards the Two-System Model of the Mind	10
	References	12
<b>2</b>	<b>The Gap Between Symbolic and Connectionist Approaches</b>	17
2.1	Differences Between Symbolic and Connectionist Approaches	17
2.1.1	Different Styles of Learning	17
2.1.2	Combinatorial Syntax and Semantics or Not?	18
2.1.3	Structured Processing or Distributed Representation?	19
2.2	The Shore of the Gap	20
2.2.1	Connectionists are <i>Representationalists</i> , Too	20
2.2.2	No Significant Difference in Computational Architectures	21
2.2.3	Connectionists are Functionalists	21
2.3	Philosophies for the Gap	21
2.4	Connectionists' Attempts to Represent Symbolic Structures	22
2.4.1	Representing Propositions	22
2.4.2	Representing Part-Whole Relations	23
2.4.3	Representing Tree Structures	24
2.5	Hybrid Systems	27

- 2.6 Summary . . . . . 27
- References . . . . . 27
- 3 Spatializing Symbolic Structures for the Gap . . . . . 31**
  - 3.1 System 1 of the Mind . . . . . 31
  - 3.2 System 2 of the Mind . . . . . 33
  - 3.3 Interaction Between System 1 and System 2 . . . . . 33
  - 3.4 Representation and Reasoning . . . . . 34
  - 3.5 Precisely Spatializing Symbolic Structures onto Vector Space . . . . . 37
  - 3.6 Summary . . . . . 40
  - References . . . . . 40
- 4 The Criteria, Challenges, and the Back-Propagation Method . . . . . 43**
  - 4.1 Spatializing Symbolic Tree Structures Onto Vector Embeddings . . . . . 43
  - 4.2 The Challenge of Zero Energy Cost . . . . . 44
  - 4.3 The Back-Propagation Method for Way-Finding . . . . . 44
    - 4.3.1 Matrix Computing . . . . . 47
    - 4.3.2 Forward Computing . . . . . 47
    - 4.3.3 Backward Updating . . . . . 48
  - 4.4 Back-Propagation for Symbol Spatialization . . . . . 51
    - 4.4.1 A Toy System to Test Back-Propagation Method . . . . . 52
    - 4.4.2 Updating Configuration Using Back-Propagation Method . . . . . 52
    - 4.4.3 Experiment Results and Analysis . . . . . 57
  - 4.5 Summary . . . . . 58
  - References . . . . . 59
- 5 Design Principles of Geometric Connectionist Machines . . . . . 61**
  - 5.1 Principle of Family Action (PFA) . . . . . 61
  - 5.2 Principle of Depth First (PDF) . . . . . 61
  - 5.3 Principle of Large Sibling Family First (PLSFF) . . . . . 62
  - 5.4 Principle of Homothetic Transformation First (PHTF) . . . . . 62
  - 5.5 Principle of Shifting and Rotation Transformation (PSRT) . . . . . 65
  - 5.6 Principle of Increasing Dimensions (POID) . . . . . 66
  - 5.7 Geometric Approach to Spatializing a Tree Structure . . . . . 67
    - 5.7.1 Separating Sibling Balls Apart . . . . . 68
    - 5.7.2 The Construction of the Parent Ball . . . . . 69
  - 5.8 Summary . . . . . 71
  - References . . . . . 71
- 6 A Geometric Connectionist Machine for Word-Senses . . . . . 73**
  - 6.1 Tree Structured Hypernym Relations of Word-Senses . . . . . 73
  - 6.2 A Geometric Connectionist Machine to Spatialize Unlabeled Tree Structures . . . . . 75

6.3	The Experiment with Word-Senses . . . . .	77
6.3.1	Input Datasets . . . . .	77
6.3.2	Output Datasets . . . . .	78
6.4	Evaluation . . . . .	78
6.4.1	Tree Structure Is Precisely Spatialized . . . . .	78
6.4.2	Pre-trained Word-Embeddings Are Well Preserved . . . . .	78
6.4.3	Consistency to Benchmark Tests . . . . .	78
6.5	Experiments with Geometric Connectionist Machines . . . . .	80
6.5.1	Similarity Measurement and Comparison . . . . .	80
6.5.2	Nearest Neighbors . . . . .	80
6.5.3	Representing Structures Beyond Word-Embeddings . . . . .	82
6.5.4	Word-Sense Validation as Way-Finding . . . . .	83
6.6	Summary . . . . .	88
	References . . . . .	88
<b>7</b>	<b>Geometric Connectionist Machines for Triple Classification . . . . .</b>	<b>89</b>
7.1	Introduction . . . . .	89
7.1.1	Subspaces for Multiple Relations . . . . .	90
7.1.2	Enriching Knowledge Graph Embedding with Text Information . . . . .	90
7.2	Geometric Connectionist Machines for Triple Classification . . . . .	91
7.2.1	Structure of the Orientation of the Central Point Vector . . . . .	92
7.2.2	Geometric Connectionist Machine to Spatialize Labeled Tree Structures . . . . .	94
7.3	The Setting of Experiments . . . . .	95
7.3.1	Datasets . . . . .	95
7.3.2	The Design of Experiment and the Evaluation Protocol . . . . .	97
7.4	Experiment 6: Triple Classification for FB13 Dataset . . . . .	98
7.5	Experiment 7: Triple Classification for WN11 Dataset . . . . .	98
7.6	Experiment 8: Triple Classification for WN18 Dataset . . . . .	101
7.7	Summary . . . . .	102
	References . . . . .	102
<b>8</b>	<b>Resolving the Symbol-Subsymbol Debates . . . . .</b>	<b>105</b>
8.1	Philosophies Behind the Symbol-Subsymbol Relation . . . . .	105
8.2	Resolving a Variety of Symbol-Subsymbol Debates . . . . .	105
8.2.1	The Necessity of Precisely Imposing External Knowledge onto Connectionist Networks . . . . .	106
8.2.2	Connectionism and Symbolicism are Compatible in $\mathcal{N}$ -Ball Configurations . . . . .	108
8.2.3	$\mathcal{N}$ -Ball Configuration as a Continuum between Connectionist Networks and Symbolic Structures . . . . .	108

8.2.4 Resolving Epistemological Challenges . . . . . 110

8.2.5 Psychological Appeal for Continuous Representations . . . 112

8.2.6  $\mathcal{N}$ -Ball Configuration Shapes Semantics . . . . . 112

8.2.7 GCM for Instantly Updating Symbolic Knowledge . . . . . 113

8.2.8 GCM Refuting the Metaphor of the Symbol-Subsymbol  
Relation to the Macro-Micro Relation in Physics . . . . . 113

8.2.9 GCM as a Content-Addressable Memory . . . . . 114

8.2.10 GCM Saves Two Birds Through a Marriage . . . . . 114

8.3 Summary . . . . . 115

References . . . . . 115

**9 Conclusions and Outlooks . . . . . 117**

9.1 Structural Imposition onto Empty Vectors . . . . . 118

9.2 Informed Machine Learning . . . . . 118

9.3 A New Building Block for Connectionist Networks . . . . . 120

9.4 Language Acquisition . . . . . 123

References . . . . . 126

**Appendix A: Code List . . . . . 129**

**Appendix B: Sample Task for Membership Validation . . . . . 131**

**Appendix C: Sample Results of Membership Validation . . . . . 135**

**Appendix D: The Nine Laws of Cognition . . . . . 139**

**Bibliography . . . . . 141**

**Index . . . . . 143**

# List of Figures

Fig. 1.1	<b>a</b> Venn diagram for the deduction that (1) $P(u)$ is true; (2) for every $x$ , if $P(x)$ is true, then $Q(x)$ is true; therefore, $Q(u)$ is true; <b>b</b> Venn diagram for the assertion that if all men are mortal and Socrates is a man, then Socrates is mortal. . . . .	5
Fig. 1.2	Four taxonomy structures of <b>apple</b> are merged into one tree. The path to <b>apple</b> , as a kind of <b>tree</b> , from <b>root</b> is uniquely identified as the path-vector (Minsky and Papert 1988; Blank et al.1992; Hilbert and Ackermann 1938) . . . . .	7
Fig. 1.3	Construct $\mathcal{N}$ -Ball embeddings of a simple tree structure of <b>apple</b> , <b>google</b> , <b>company</b> in the blue circle in Fig. 1.2 . . . . .	8
Fig. 1.4	An architecture of integrating connectionism with symbolism through symbol spatialization . . . . .	11
Fig. 2.1	<b>a</b> Square $S$ and point $P$ ; <b>b</b> the distance between $P$ and $AD$ is $PQ$ , the distance between $P$ and $AB$ is $PA$ , so $P$ is nearer to $AD$ than to $AB$ ; <b>c</b> the double cross orientation framework; <b>d</b> a familiar qualitative orientation framework, <b>N</b> , <b>W</b> , <b>S</b> , <b>E</b> represent North, West, South, East, respectively . . . . .	19
Fig. 2.2	By introducing role-specific units (agent, relation, patient) as the hidden layer between the concept layer and the proposition layer, the network can learn to assign concepts to role units. . . . .	23
Fig. 2.3	The letter <b>p</b> is part of the word <b>apple</b> . After training process, <b>p</b> will have different vector embeddings, when located at different positions . . . . .	24
Fig. 2.4	Mapping different part-whole relations into the same network structure, <b>S2</b> has different embedding vectors in (b) and (c). . . . .	25
Fig. 2.5	<b>a</b> A tree structure in which <b>A</b> , <b>B</b> , <b>C</b> have vector embeddings, <b>D</b> does not; <b>b</b> we can compute the vector embedding of <b>D</b> by using back-propagation algorithm . . . . .	26
Fig. 2.6	The architecture of B-RAAM . . . . .	26

Fig. 3.1 If you look at the two words **Bananas** and **vomit**, you may not feel very comfortable (Kahneman 2011, p. 50) . . . . . 32

Fig. 3.2 Filling a letter between **O** and **P** to form a word (Kahneman 2011, p. 52) . . . . . 32

Fig. 3.3 System 2 follows rules to perform calculation. . . . . 33

Fig. 3.4 Quick glancing at these images may have different recognition results . . . . . 34

Fig. 3.5 San Diego (CA) is further east than Reno (NV). Picture copied from Google Map . . . . . 35

Fig. 3.6 A simple semantic network of spatial knowledge among California, Nevada, San Diego, and Reno . . . . . 36

Fig. 3.7 A partially region-based hierarchical structure among California, Nevada, San Diego, and Reno . . . . . 36

Fig. 3.8 How shall we perfectly represent the *is-a* relation? . . . . . 37

Fig. 3.9 Spatializing symbolic structures onto vector space. Connectionist network represents a word as a one-element vector. Biased by training sentence, **Hamburg** vector [348] is closer to **city** vector [327] and **capital** vector [319] than to **harbor-city** vector [300]. To precisely encode symbolic tree structures, we promote them into circles, for example, promoting **harbor-city** vector to a circle with central point [572, 300] with radius = 53, with the aim that inclusion relations among circles encode child-parent relations in the tree structure. . . . . 39

Fig. 4.1 **Dessert** ball partially overlaps with **plant** ball, although they should be disconnected from each other . . . . . 45

Fig. 4.2 You are in this maze, equipped with a computer, and given a partial route instruction. Now, the instruction reaches the end, you are still in the maze, which direction shall you take in the next crossing point?. . . . . 45

Fig. 4.3 **a** A connectionist network with one hidden layer, two input nodes, and two output nodes; **b** matrix computation of the second layer; **c** matrix computation of the first layer. . . . . 46

Fig. 4.4 A small knowledge-base with two tree structures . . . . . 52

Fig. 4.5 Diagrammatic representation of the initial entity balls. . . . . 53

Fig. 4.6 The structure of a ball in 2-dimensional space is defined as an open region . . . . . 54

Fig. 4.7 Ball  $\mathbb{B}_1$  is inside Ball  $\mathbb{B}_2$ . . . . . 54

Fig. 4.8 Ball  $\mathbb{B}_1$  disconnects from  $\mathbb{B}_2$ . . . . . 55

Fig. 4.9 The back-propagation approach sometimes achieves zero-energy cost . . . . . 58

Fig. 4.10 The back-propagation approach fails to achieve zero-energy cost . . . . . 59

Fig. 5.1 Each color marks one possible sequence to traverse the tree. . . . 62

Fig. 5.2 **a** Google ball connects with **apple** ball; **b** A homothetic transformation is applied for **google** ball . . . . . 63

Fig. 5.3 Ball  $O_1$  disconnects from ball  $O_2$ , and contains ball  $O_3$ . After homothetic transformation, both relations are kept: ball  $O'_1$  disconnects from ball  $O'_2$ , and contains ball  $O'_3$  . . . . . 63

Fig. 5.4 Ball  $O_1$  contains the origin point  $O$ , and overlaps with Ball  $O_2$ . After we apply the homothetic transformation for Ball  $O_1$  with  $k = 3$ , it still overlaps with Ball  $O_2$  . . . . . 64

Fig. 5.5 A shift transformation on Ball  $O_1$  will change the central point vectors of child balls, e.g. Ball  $O_2$  . . . . . 65

Fig. 5.6 **a** Without adding a new dimension, hyponyms of **harbor-city** include **Hamburg**, as well as **Berlin**, **capital**, and **city**; **b** by adding a new dimension, the hyponyms of **harbor-city** only contain **Hamburg** . . . . . 67

Fig. 5.7 Construct **company** ball that contains **apple** ball and **google** ball . . . . . 70

Fig. 6.1 A tree structure in WordNet 3.0 for three word-senses of **flower** . . . . . 74

Fig. 6.2 The diagram of  $\mathcal{N}$ -Ball embeddings of the tree structure in Fig. 6.1 . . . . . 74

Fig. 6.3 Three components of the center point of an  $\mathcal{N}$ -Ball embedding . . . . . 75

Fig. 6.4 Recall increases rapidly with the increasing of the radius . . . . . 86

Fig. 6.5 Precision and recall without margin extension . . . . . 87

Fig. 7.1 Subspaces inside **city** ball . . . . . 91

Fig. 7.2 Transforming a labeled tree into an unlabeled tree . . . . . 92

Fig. 7.3 The structure of central point vector and a labeled tree . . . . . 93

Fig. 7.4 Precision of Triple classification using WN11  $\mathcal{N}$ -Ball dataset. When  $\gamma$  increases, the precision will drop . . . . . 99

Fig. 7.5 Recall of Triple classification using WN11  $\mathcal{N}$ -Ball dataset. When  $\gamma$  increases, the recall will increase . . . . . 99

Fig. 7.6 Accuracy of Triple Classification using WN11  $\mathcal{N}$ -Ball with different  $\gamma$  values . . . . . 99

Fig. 7.7 Precision versus length of type chains in WN11  $\mathcal{N}$ -Ball Dataset, when  $\gamma = 1$  . . . . . 100

Fig. 7.8 Accuracy versus different lengths of type chains in WN11  $\mathcal{N}$ -Ball Dataset; when lengths increases, the accuracy will have the strong tendency to increase; numbers in the plots represent the  $\gamma$  value with which the accuracy reaches maximum . . . . . 100

Fig. 7.9 With the unbalanced WN18  $\mathcal{N}$ -Ball testing dataset, the accuracy reaches 98% using TEKE\_E pre-trained entity-embeddings, when  $\gamma = 1$  . . . . . 101



Fig. 7.10	With the unbalanced WN18 $\mathcal{N}$ -Ball testing dataset, the accuracy increases along with the length of type-chains, when $\gamma = 1$ . . . . .	102
Fig. 8.1	Symbol spatialization reaches zero energy cost through geometric transformations . . . . .	106
Fig. 8.2	Blue points represent British colonial islands with Left Hand Traffic System (LHTS); Red points represent French colonial islands with Right Hand Traffic System (RHTS). Without imposing any external knowledge, autonomous driving cars will not have concepts of different traffic systems. Picture is copied from Wikipedia . . . . .	107
Fig. 8.3	An instruction that the letter Q shall be pronounced $ ch $ , if it appears in Mandarin . . . . .	110
Fig. 8.4	An implementation using GCMs for an additional instruction . . . . .	111
Fig. 9.1	Synergistic integration of two-system model of the mind through Geometric Connectionist Machine (GCM) . . . . .	118
Fig. 9.2	“White as snow” is translated into “white as pelican” in the Natemba language . . . . .	119
Fig. 9.3	Semantic structures of <i>yesterday John bought apples; last week John bought pears</i> , respectively . . . . .	121
Fig. 9.4	The integrated semantic network . . . . .	121
Fig. 9.5	An $\mathcal{N}$ -Ball solution for the crosstalk problem . . . . .	121
Fig. 9.6	A connectionist network to learn the identity relation. This diagram is copied from Marcus (2003, p. 46) . . . . .	122
Fig. 9.7	An $\mathcal{N}$ -Ball solution, in the slow thinking model, to predict the identity relation . . . . .	123
Fig. 9.8	The $\mathcal{N}$ -Ball semantic representation of the sentence “Mary believes Jane to like Bill”. . . . .	124
Fig. 9.9	The $\mathcal{N}$ -Ball of <i>believes-actor</i> is enlarged to partially overlap with $\mathcal{N}$ -Ball of <i>like-actor</i> , so that $\mathcal{N}$ -Ball of <i>Mary</i> could coincide with $\mathcal{N}$ -Ball of <i>Jane</i> . This ends with the denotation of <i>herself</i> . . . . .	125
Fig. 9.10	If the $\mathcal{N}$ -Ball of <i>believes-actor</i> is enlarged to partially overlap with the $\mathcal{N}$ -Ball of <i>like-object</i> , it must overlap with the $\mathcal{N}$ -Ball of <i>like-actor</i> . . . . .	126

# List of Tables

Table 1.1	Top-6 $\mathcal{N}$ -Ball nearest neighbors compared with top-N GloVe nearest neighbors. $\mathcal{N}$ -Ball neighbors possess clean semantics . . . . .	9
Table 4.1	Two tree structures in the small knowledge-base . . . . .	52
Table 4.2	Each entity is represented by a 2-dimensional ball . . . . .	53
Table 6.1	Word embedding part $WE_{\mathcal{N}\text{-Ball}}$ in $\mathcal{N}$ -Balls produce the same Spearman's correlation as the pertained word-embedding $WE_{GloVe}$ . . . . .	79
Table 6.2	Top-5 nearest neighbors measured by $Sim_{\tau}$ and $\cos$ , respectively . . . . .	81
Table 6.3	Nearest neighbors using GloVe embeddings, and ConceptNet embeddings . . . . .	82
Table 6.4	Upper-level categories based on $f_{\mathbf{p}}$ . $-_k$ represents the $k$ th largest negative value of $f_{\mathbf{p}}$ . . . . .	82
Table 6.5	The table of recall with different parameters. “% in Training” represents how much percentage is used as training data; “Ratio $\times$ Radius” represents the radius $r$ being enlarged by multiplying a ratio . . . . .	85
Table 7.1	Mapping WN18/WN11/FB13 relations to $\mathcal{N}$ -Ball relations . . . . .	96
Table 7.2	Datasets extracted from WN11, FB13, WN18 . . . . .	97
Table 8.1	Connectionist architectures are inadequate as being lack of a number of important features (Prince and Pinker 1988) . . . . .	109
Table 8.2	Prince and Pinker (1988) claimed that a number of features that Smolensky (1988b) used to distinguish connectionism from symbolism are not correct. . . . .	109

Table 9.1 Sample dataset for the identity mapping from **X** to **Y**. This example is copied from Marcus (2003, p. 37) . . . . . 122

Table C.1 List of membership validation. Column A is the total number of children; Column B is the number of training set; TP represents the number of true-positive predictions; FP represents the number of false-positive predictions; FN represents the number of false-negative predictions . . . . . 135