

Contents

Part I A Comprehensive Mathematical Framework for the Development of Semantic Technologies

- 1 Mathematical Models for Designing Natural Language Processing Systems as a New Field of Studies for Systems Science 3**
 - 1.1 An Idea of a Bridge Between Systems Science and Engineering of Semantics-Oriented Natural Language Processing Systems 3
 - 1.2 The Models of Types 1–4 5
 - 1.2.1 The Models of Type 1 5
 - 1.2.2 The Models of Type 2 7
 - 1.2.3 The Models of Type 3 7
 - 1.2.4 The Models of Type 4 10
 - 1.3 The Models of Type 5 11
 - 1.4 The Significance of the Models for the Design of Linguistic Processors 16
 - 1.5 The Context of Cognitive Linguistics for the Formal Study of Natural Language 17
 - 1.6 Early Stage of Natural Language Formal Semantics 18
 - 1.7 The Significance of Highly Expressive Formal Systems of Semantic Representations 22
- 2 Introduction to Integral Formal Semantics of Natural Language 27**
 - 2.1 The Basic Principles of Integral Formal Semantics of Natural Language 27
 - 2.1.1 Basic Principles 27
 - 2.1.2 The Notion of a Broadly Applicable Conceptual Metagrammar 29
 - 2.2 The Components of Integral Formal Semantics of Natural Language 30
 - 2.3 The Theory of S-Calculuses and S-Languages 32
 - 2.4 A Model of a Correspondence Between NL-Texts and Their Semantic Representations 37

2.5	The Theory of T-Calculuses and T-Languages	38
2.6	The Initial Version of the Theory of K-Calculuses and K-Languages	39
2.7	The Theory of K-Representations as the Kernel of the Current Version of Integral Formal Semantics	41
3	A Mathematical Model for Describing a System of Primary Units of Conceptual Level Used by Applied Intelligent Systems	43
3.1	Global Task Statement	43
3.2	Local Task Statement	47
3.3	Basic Denotations and Auxiliary Definitions	48
3.3.1	General Mathematical Denotations	48
3.3.2	The Preliminary Definitions from the Theory of Formal Grammars and Languages	48
3.3.3	The Used Definitions from the Theory of Algebraic Systems	49
3.4	The Basic Ideas of the Definition of a Sort System	50
3.5	The Formal Definition of a Sort System	52
3.6	Types Generated by a Sort System	55
3.7	The Concretization Relation on the Set of Types	58
3.7.1	Basic Ideas	58
3.7.2	Formal Definitions	60
3.8	Concept-Object Systems	64
3.9	Systems of Quantifiers and Logical Connectives: Conceptual Bases	67
3.10	A Discussion of the Constructed Mathematical Model	71
3.10.1	Mathematical Peculiarities of the Model	71
3.10.2	The Comparison of the Model with Related Approaches ...	72
4	A Mathematical Model for Describing Structured Meanings of Natural Language Sentences and Discourses	77
4.1	The Essence of a New Approach to Formalizing Semantic Structure of Natural Language Texts	77
4.1.1	Toward Expanding the Universe of Formal Study	77
4.1.2	The Algebraic Essence of the Model Describing Conceptual Operations	80
4.1.3	Shortly About the Rules for Building Semantic Representations of Natural Language Texts	81
4.1.4	The Scheme of Determining Three Classes of Formulas Generated by a Conceptual Basis	84
4.2	The Use of Intensional Quantifiers in Formulas	86
4.3	The Use of Relational Symbols and the Marking-Up of Formulas	92
4.3.1	The Rules for Employing Relational Symbols	92
4.3.2	The Rule for Marking Up the Formulas	96

4.4	The Use of Logical Connectives NOT, AND, OR	97
4.5	Building Compound Designations of Notions and Objects	99
4.5.1	Compound Designations of Notions	99
4.5.2	Compound Designations of Objects	101
4.6	Final Rules	102
4.6.1	The Use of Existential and Universal Quantifiers	102
4.6.2	The Representations of Finite Sequences	104
4.6.3	A Summing-Up Information about the Rules P[0]–P[10] ...	105
4.7	SK-Languages: Mathematical Investigation of Their Properties	107
5	A Study of the Expressive Possibilities of SK-Languages	113
5.1	A Convenient Method of Describing Events	113
5.2	Formalization of Assumptions About the Structure of Semantic Representations of Sets	115
5.3	Semantic Representations of Questions with the Role Interrogative Words	118
5.4	Semantic Representations of Questions About the Quantity of Objects and Events	120
5.5	Semantic Representations of Questions with an Interrogative Pronoun Attached to a Noun	121
5.6	Semantic Representations of General Questions	122
5.7	Describing Semantic Structure of Commands	123
5.8	Representation of Set-Theoretical Relationships and Operations on Sets	123
5.9	Semantic Representations of Phrases with Subordinate Clauses of Purpose and Indirect Speech	124
5.10	Explicit Representation of Causal and Time Relations in Discourses	125
5.11	Semantic Representations of Discourses with the References to the Meanings of Phrases and Larger Parts of the Text	127
5.12	Representing the Pieces of Knowledge About the World	127
5.13	Object-Oriented Representations of Knowledge Pieces	128
5.14	The Marked-Up Conceptual Bases	129
5.15	Related Approaches to Representing Semantic Structure of NL-Texts	131
6	The Significance of a New Mathematical Model for Web Science, E-science, and E-commerce	137
6.1	The Problem of Semantic Data Integration	137
6.1.1	The Purpose of Semantic Data Integration in E-Science and Other E-Fields	137
6.1.2	Ontologies in Modern Information Society	138
6.1.3	The Language UNL and the Problem of Sharing Knowledge	139

6.2	Building Semantic Annotations of Web Data	140
6.3	Conceptual Descriptions of Visual Images	143
6.4	Representation of Knowledge in Biology and Ecology	145
6.5	Representation of Knowledge in Medicine	147
6.6	Representation of Semantic Content in Business	148
6.7	SK-Languages as a Tool for Building E-Contracts	149
6.7.1	Formal Languages for E-Contracting	149
6.7.2	The Possibilities of Forming Contracts and Records of E-Negotiations by Means of SK-Languages	150
6.8	Simulation of the Expressive Mechanisms of RDF, RDFS, and OWL	152
6.8.1	Simulation of the Expressive Mechanisms of RDF and RDFS	152
6.8.2	Simulation of OWL Expressive Mechanisms	155
6.9	A Metaphor of a Kitchen Combine for the Designers of Semantic Technologies	156
6.10	The Significance of the Theory of K-Representations for Semantic Web and Web Science	158
6.10.1	Theory of K-Representations as a Universal Resources and Agents Framework	158
6.10.2	The Need for the Incentives for Semantic Web	159
6.10.3	Toward a New Language Platform for Semantic Web	160
6.10.4	A Possible Strategy of Developing Semantic Web of a New Generation	160

Part II Formal Methods and Algorithms for the Design of Semantics-Oriented Linguistic Processors

7	A Mathematical Model of a Linguistic Database	165
7.1	The Principles of Designing Semantics-Oriented Linguistic Processors	165
7.2	Morphological Bases	167
7.3	Text-Forming Systems	171
7.4	Lexico-semantic Dictionaries	173
7.5	Dictionaries of Verbal – Prepositional Semantic-Syntactic Frames ..	175
7.6	The Dictionaries of Prepositional Frames	179
7.7	Linguistic Bases	181
7.7.1	Semantic Information Associated with the Role Interrogative Words	181
7.7.2	The Notion of a Linguistic Basis	182
8	A New Method of Transforming Texts into Semantic Representations	185
8.1	A Component-Morphological Representation of an NL-text	185
8.1.1	Morphological Representation	185
8.1.2	Classifying Representation	187

8.2	The Projections of the Components of a Linguistic Basis on the Input Text	189
8.3	Matrix Semantic-Syntactic Representations of NL-Texts	192
8.4	A New Method of Transforming NL-Texts into Semantic Representations	194
8.4.1	Formulation of the Method	194
8.4.2	The Principles of Selecting the Form of a Text's Semantic Representation	196
9	Algorithm of Building a Matrix Semantic-Syntactic Representation of a Natural Language Text	201
9.1	Task Statement	201
9.1.1	Purpose of Development and Non-detailed Structure of the Algorithm	201
9.1.2	Some Peculiarities of the Approach	202
9.2	Initial Stages of Developing the Algorithm BuildMatr1	204
9.2.1	External Specification of the Algorithm BuildMatr1	204
9.2.2	Development of a Plan of the Algorithm BuildMatr1	204
9.3	Description of the Algorithm Classifying the Input Texts.....	205
9.3.1	The Purpose of the Algorithm.....	205
9.3.2	Compound Designations of the Subclasses of Lexical Units	207
9.3.3	External Specification of the Algorithm “Defining-form-of-text”	208
9.3.4	Algorithm “Defining-form-of-text”	209
9.4	Principles of Processing the Role Interrogative Words	212
9.5	An Algorithm of Searching for Semantic Connections of the Verbs	215
9.5.1	Key Ideas of the Algorithm	216
9.5.2	Description of an Algorithm of Searching for Semantic Connections Between a Verb and a Noun Group	217
9.5.3	Description of the Algorithm Processing the Constructs	221
9.5.4	Description of the Algorithm “Find-set-thematic-roles”	223
9.5.5	An Algorithm of Searching for a Semantic Relationship Between a Verb and a Dependent Expression	224
9.5.6	Final Part of a Description of an Algorithm of Processing Verbs	227
9.6	Processing of Adjectives, Prepositions, Cardinal Numerals, Names and Nouns	229
9.6.1	Processing of Adjectives	229
9.6.2	Processing of Prepositions, Cardinal Numbers, and Names.....	230
9.6.3	An Algorithm Searching for Possible Semantic Connections Between Two Nouns with Respect to a Preposition	231

9.7	Finding the Connections of a Noun with Other Text Units	235
9.8	Final Step of Developing an Algorithm Building a Matrix	
	Semantic-Syntactic Representation of the Input Text	244
9.8.1	Description of the Head Module of the Algorithm	244
9.8.2	External Specifications of Auxiliary Algorithms	245
9.8.3	Algorithm of Building an MSSR of the Input Text	246
10	An Algorithm of Semantic Representation Assembly	249
10.1	Initial Step of Building Semantic Representations of Input Texts . .	249
	10.1.1 Description of the Algorithm	
	“Preparation-to-constr-SemRepr”	250
10.2	Semantic Representations of Short Fragments of the Input Texts . .	254
	10.2.1 Description of the Algorithm	
	“Calculation-of-the-kind-of-case”	254
	10.2.2 External Specification of the Algorithms	
	BuildSemdes1 – BuildSemdes7	256
	10.2.3 Description of the Algorithms	
	BuildSemdes1 – BuildSemdes7	257
	10.2.4 Description of the Algorithm ProcessSit	262
	10.2.5 Description of the Algorithm “Begin-of-constr-SemRepr” . .	264
10.3	Development of the Algorithm “Representation-of-situations”	265
	10.3.1 The Key Ideas of the Algorithm	265
	10.3.2 Description of the Algorithm	
	“Representation-of-situations”	268
10.4	Final Stages of Developing an Algorithm of Semantic	
	Representation Assembly	273
	10.4.1 External Specification of the Algorithm	
	“Final-operations”	274
	10.4.2 Algorithm “Final-operations”	275
	10.4.3 An Algorithm of Semantic Representation Assembly	280
10.5	A Multilingual Algorithm of Semantic-Syntactic Analysis	
	of Natural Language Texts	281
	10.5.1 Description of the Algorithm SemSynt1	281
	10.5.2 Some Possible Directions of Using the Algorithm	
	SemSynt1	282
11	Natural Language Processing Applications	285
11.1	The Structure of a Computer Intelligent Agent for Semantic	
	Classification of E-mail Messages	285
	11.1.1 The Problem of Semantic Classification	
	of E-mail Messages	285
	11.1.2 An Outline of the Computer Intelligent Agent Mailagent1 . .	286
	11.1.3 General Structure of Computer Intelligent Agent	
	Mailagent1	288
	11.1.4 The Structure of Semantic-Syntactic Patterns	
	and Lexico-Semantic Dictionary	288
	11.1.5 Implementation Data	293

11.2 A Transformer of Natural Language Knowledge Descriptions into OWL-Expressions	293
A Proofs of Lemmas 1, 2 and Theorem 4.5 from Chapter 4	299
A.1 Proof of Lemma 1	299
A.2 Proof of Lemma 2	302
A.3 Proof of Theorem 4.5	304
Glossary	313
References	315
Index	327

Semantics-Oriented Natural Language Processing
Mathematical Models and Algorithms

Fomichov A., V.

2010, XXIII, 328 p., Hardcover

ISBN: 978-0-387-72924-4