

# Contents

Introduction .....	i
List of Abbreviations .....	v
1. Mathematical models of natural language .....	1
1.1. Modelling in linguistics.....	1
1.1.0. Introductory notes.....	1
1.1.1. The basic principles of modelling .....	1
1.1.2. Linguistic models.....	2
1.1.3. Concluding remarks .....	4
1.2. The natural language and the language of mathematics.....	4
1.2.0. Introductory notes.....	4
1.2.1. Two levels of reasoning and generation of speech utterances.....	5
1.2.2. The fuzziness of language and its elements .....	7
1.2.3. Elementary concepts of the Fuzzy Set Theory and their application to linguistic objects .....	11
1.2.3.1. Concept and definition of the Fuzzy Set .....	11
1.2.3.2. A linguistic illustration.....	13
1.2.3.3. Operation on fuzzy sets .....	14
1.2.4. Sign in the natural language.....	17
1.2.4.1. An outline model of the sign in natural language .....	17
1.2.4.2. The language sign.....	20
1.2.4.2.1. The language sign as a bilateral psychic entity ....	21
1.2.4.2.2. Vehicle of meaning (name), conceptum, designatum, and denotatum .....	22
1.2.4.2.3. Relation of “designatum – denotatum” and value (valuer) .....	25
1.2.4.2.4. Linguistic illustration .....	26
1.2.4.2.5. Connotatum.....	28
1.2.4.2.6. Onomasiological and classificationally-semasiologi- cal functions of the language sign .....	29

1.2.4.3. Speech aspect of the sign.....	30
1.2.4.4. Transition from the language sign to the speech sign..	34
1.2.4.4.1. Speech sign. Actualization and pragmatics .....	34
1.2.4.4.2. An outline model of the sign in the natural language. Linguistic illustration .....	34
1.2.4.5. Sign in the language of mathematics.....	37
1.3. Concluding remarks.....	41
2. Application of the sets theory and relational algebra to modelling lexical systems .....	43
2.1. Basic concepts of the sets theory in linguistics .....	43
2.1.0. Introductory notes.....	43
2.1.1. Classical set and its elements .....	44
2.1.2. Subsets, sets, and supersets.....	47
2.1.3. Concluding remarks .....	48
2.2. Operations on sets and their properties .....	48
2.2.0. Introductory notes.....	48
2.2.1. Union of sets .....	49
2.2.1.1. Properties of the union operation.....	50
2.2.2. Intersection of sets .....	51
2.2.2.1. Properties of the intersection operation.....	52
2.2.3. Subtraction of sets .....	54
2.2.3.1. Properties of the subtraction operation.....	55
2.2.4. Concluding remarks .....	57
2.3. Cartesian product of sets.....	57
2.4. Binary relations.....	58
2.4.0. Introductory notes.....	58
2.4.1. A binary relation on a set .....	59
2.4.2. Properties of binary relations.....	62
2.4.2.1. Reflexivity .....	63
2.4.2.2. Mesoreflexivity.....	63

2.4.2.3. Antireflexivity .....	64
2.4.2.4. Symmetry .....	64
2.4.2.5. Mesosymmetry .....	64
2.4.2.6. Asymmetry .....	65
2.4.2.7. Antisymmetry .....	65
2.4.2.8. Transitivity .....	65
2.4.2.9. Mesotransitivity .....	66
2.4.2.10. Non-transitivity .....	66
2.4.3. Concluding remarks .....	66
2.5. Classes of binary relations .....	67
2.5.0. Introductory notes .....	67
2.5.1. Equivalence and tolerance .....	67
2.5.2. Synonymy .....	68
2.5.2.1. Formal definition of synonymy .....	69
2.5.2.2. Absolute synonyms of doublets (a linguistic illustration) .....	70
2.5.2.3. Synonymity and non-synonymity (a linguistic illustration) .....	71
2.5.2.4. Synonymy and paradigmatic meaning relations between words and word combinations .....	71
2.5.2.5. Concluding remarks .....	72
2.5.3. Order Relation .....	73
2.5.3.1. Non-strict order .....	73
2.5.3.2. Strict order .....	74
2.5.4. Concluding remarks .....	77
2.6. Thesaurus modelling of language systems .....	77
2.6.0. Introductory notes .....	77
2.6.1. Linguistic nets .....	77
2.6.2. Thesaurus .....	78
2.6.3. Union of tree thesauri .....	80

2.6.3.1. Union of "Grammatical Categories of Noun" and "Local cases" thesauri.....	81
2.6.3.2. A formal model of the union of thesauri.....	85
2.6.3.2.1. Mathematical setting of the union of thesauri.....	85
2.6.4. Transformation of thesauri.....	88
2.6.4.1. Branching in a tree thesaurus.....	88
2.6.5. Concluding remarks.....	90
3. Modelling semantic translation by means of mapping.....	91
3.1. Semantic pattern recognition of a linguistic unit.....	91
3.1.0. Introductory notes.....	91
3.1.1. Basic concepts of the semantic pattern recognition theory.....	91
3.1.2. Semantic pattern recognition and machine translation....	94
3.1.3. Concluding remarks.....	95
3.2. A mathematical model of machine translation.....	96
3.2.0. Introductory notes.....	96
3.2.1. An outline model of machine translation.....	97
3.2.2. Semantic filter.....	99
3.2.2.1. A binary tree thesaurus of the object region.....	102
3.2.2.2. Resolution of the lexical ambiguities via selecting combinations of thesaurus code symbols.....	103
3.2.3. Concluding remarks.....	106
3.3. Selection of semantic distinctive features in the arrangement of the thesaurus.....	107
3.3.0. Introductory notes.....	107
3.3.1. Application of logico-semantic categories.....	107
3.3.2. An analysis of dictionary verbal definitions of a word....	108
3.3.3. Context analysis.....	112
3.3.4. Concluding remarks.....	113
3.4. The arrangement of the thesaurus semantic space.....	113

3.5. An algorithm of semantic pattern recognition of a linguistic unit via the thesaurus semantic space .....	132
3.5.0. Introductory notes .....	132
3.5.1. Recognition filters .....	133
3.5.2. A flow diagram of machine translation including the semantic pattern recognition of the input of current words and word forms .....	133
3.5.2.1. The semantic pattern recognition of current words in the noun phrase $N_1 + N_2(A + N)$ .....	140
3.5.2.2. Semantic pattern recognition of current words in the verb-noun phrase $(V + N)$ .....	141
3.5.2.3. The pattern recognition of semantico-syntactic bonds between verb and noun segments .....	142
3.5.3. Concluding remarks .....	144
4. A formal mathematical theory of artificial language .....	146
4.0. Introductory notes .....	146
4.1. The basic concepts of the ETALON language .....	147
4.1.0. Concluding remarks .....	149
4.2. The formal theory and model of the ETALON language ..	150
4.2.0. Introductory notes .....	150
4.2.1. The syntax and the semantics .....	151
4.2.2. The sets and the relations of the information representation model in the ETALON language .....	151
4.2.3. The model and theory of semantic information representation .....	153
4.2.4. The model and the theory of utterances: the model of the ETALON syntagmatics .....	155
4.2.5. The frame model of the ETALON paradigmatics .....	158
4.3. Concluding remarks .....	161
5. Probability .....	162
5.1. Linguistics and chance .....	162

5.1.0. Introductory notes .....	162
5.1.1. Linguistic trials and events .....	162
5.1.2. Ontology and epistemology of chance in linguistics .....	165
5.1.3. Concluding remarks .....	167
5.2. Algebra of events .....	168
5.2.0. Introductory notes .....	168
5.2.1. Elementary and complex events .....	168
5.2.2. Operations on linguistic events and on relations of events .....	168
5.2.3. Concluding remarks .....	171
5.3. Probability of a linguistic event .....	171
5.3.0. Introductory notes .....	171
5.3.1. The classical definition of probability (the chance scheme) .....	172
5.3.1.1. Calculation of probabilities of lexical units in a separate text and the complete works of an author....	174
5.3.1.2. The use of the theory of combinations (combina- torics) in the classical definition of probability .....	175
5.3.1.2.1. Permutations .....	176
5.3.1.2.2. Repetitive permutation .....	177
5.3.1.2.3. Rearrangements .....	177
5.3.1.2.4. Repetitive rearrangements .....	178
5.3.1.2.5. Combinations .....	179
5.3.1.2.6. Repetitive combinations .....	180
5.3.2. Statistical definition of probability and sampling frequency description of a text .....	181
5.3.3. Subjective probability .....	186
5.3.3.1. Experiments on collective guessing of the letters of an unknown word or a text .....	188
5.3.4. Axiomatic construction of the probability theory .....	191
5.3.5. Probabilities of complex linguistic events .....	201

5.3.5.1.	Addition of probabilities .....	201
5.3.5.2.	Prediction of probabilities of linguistic events of repetitious trials .....	202
5.3.5.3.	Dependent linguistic events and conditional probabilities .....	203
5.3.5.4.	Probability multiplication rule and calculation of probabilities of linguistic chains .....	204
5.3.5.5.	Calculation of the total probability of a linguistic event by the complete probability theorem .....	206
5.3.5.6.	Prior and posterior probabilities. Calculation of the probabilities of a linguistic hypothesis .....	208
5.3.6.	Concluding remarks .....	210
5.4.	Probability and fuzziness .....	211
5.4.0.	Introductory notes .....	211
5.4.1.	Fuzzy measure and probability. Fuzzy set .....	211
5.4.1.1.	Linguistic example .....	212
5.4.2.	Fuzzy relations .....	213
5.4.2.1.	Linguistic example .....	214
5.4.3.	Concluding remarks .....	216
6.	Natural language message and measurement of its information .....	218
6.1.	Information process and information .....	218
6.1.0.	Introductory notes .....	218
6.1.1.	Information process .....	218
6.1.1.1.	Message generation .....	220
6.1.2.	Types of information .....	222
6.1.3.	Concluding remarks .....	224
6.2.	Uncertainty and potential information .....	224
6.2.0.	Introductory notes .....	224
6.2.1.	Uncertainty and entropy .....	224

6.2.2. Combinatorial approach to the evaluation of the quantity of information .....	227
6.2.2.1. Measurement of constraints imposed by the language system and norm on the occurrence of a linguistic unit .....	228
6.2.3. Probabilistic approach to the measurement of information quantity .....	230
6.2.4. Potential information .....	233
6.2.5. Concluding remarks .....	235
6.3. Syntactic and meaning information .....	235
6.3.0. Introductory notes .....	235
6.3.1. Syntactic information .....	235
6.3.1.1. Peculiarities of the syntactic information distribution in texts and vocabularies .....	238
6.3.1.2. Contextual constraints and text redundancy .....	239
6.3.2. Measuring of meaning information in a text .....	240
6.3.3. Concluding remarks .....	246
Conclusion .....	247
Bibliography .....	253