

CONTENTS

PREFACE	ix
1 INTRODUCTION	1
2 SEMANTIC EXTRACTION	7
2.1 Historical Overview	7
2.2 Cognitive Science Approaches	8
2.3 Recycling Approaches	17
2.4 Knowledge-Poor Approaches	23
3 SEXTANT	33
3.1 Philosophy	33
3.2 Methodology	34
3.3 Other examples	54
3.4 Discussion	57
4 EVALUATION	69
4.1 Deese Antonyms Discovery	70
4.2 Artificial Synonyms	75
4.3 Gold Standards Evaluations	81
4.4 Webster's 7th	89
4.5 Syntactic vs. Document Co-occurrence	91
4.6 Summary	100
5 APPLICATIONS	101
5.1 Query Expansion	101
5.2 Thesaurus enrichment	114

5.3	Word Meaning Clustering	126
5.4	Automatic Thesaurus Construction	131
5.5	Discussion and Summary	133

6	CONCLUSION	137
6.1	Summary	137
6.2	Criticisms	139
6.3	Future Directions	141
6.4	Vision	147

1	PREPROCESSORS	149
---	---------------	-----

2	WEBSTER STOPWORD LIST	151
---	-----------------------	-----

3	SIMILARITY LIST	153
---	-----------------	-----

4	SEMANTIC CLUSTERING	163
---	---------------------	-----

5	AUTOMATIC THESAURUS GENERATION	171
---	--------------------------------	-----

6	CORPORA TREATED	181
---	-----------------	-----

6.1	ADI	181
6.2	AI	187
6.3	AIDS	192
6.4	ANIMALS	197
6.5	BASEBALL	202
6.6	BROWN	207
6.7	CACM	212
6.8	CISI	219
6.9	CRAN	228
6.10	HARVARD	235
6.11	JFK	240
6.12	MED	247
6.13	MERGERS	252
6.14	MOBYDICK	257
6.15	NEJM	261
6.16	NPL	266

6.17	SPORTS	273
6.18	TIME	278
6.19	XRAY	285
6.20	THESIS	290

INDEX	303
-------	-----