

Contents

1	Introduction	1
1.1	Coarse-to-Fine Models.....	2
1.2	Coarse-to-Fine Inference	4
2	Latent Variable Grammars for Natural Language Parsing	7
2.1	Introduction	7
2.1.1	Experimental Setup	9
2.2	Manual Grammar Refinement	10
2.2.1	Vertical and Horizontal Markovization	11
2.2.2	Additional Linguistic Refinements	12
2.3	Generative Latent Variable Grammars	13
2.3.1	Hierarchical Estimation	14
2.3.2	Adaptive Refinement	16
2.3.3	Smoothing.....	18
2.3.4	An Infinite Alternative	19
2.4	Inference.....	20
2.4.1	Hierarchical Coarse-to-Fine Pruning	21
2.4.2	Objective Functions for Parsing.....	26
2.5	Additional Experiments.....	29
2.5.1	Experimental Setup	29
2.5.2	Baseline Grammar Variation	30
2.5.3	Final Results WSJ	31
2.5.4	Multilingual Parsing	31
2.5.5	Corpus Variation	33
2.5.6	Training Size Variation	34
2.6	Analysis	35
2.6.1	Lexical Subcategories	35
2.6.2	Phrasal Subcategories	38
2.6.3	Multilingual Analysis.....	40
2.7	Summary and Future Work	42

3	Discriminative Latent Variable Grammars	47
3.1	Introduction	47
3.2	Log-Linear Latent Variable Grammars	47
3.3	Single-Scale Discriminative Grammars	50
3.3.1	Efficient Discriminative Estimation	50
3.3.2	Experiments	52
3.4	Multi-scale Discriminative Grammars	55
3.4.1	Hierarchical Refinement	56
3.4.2	Learning Sparse Multi-scale Grammars	58
3.4.3	Additional Features	61
3.4.4	Experiments	62
3.4.5	Analysis	65
3.5	Summary and Future Work	66
4	Structured Acoustic Models for Speech Recognition	69
4.1	Introduction	69
4.2	Learning	71
4.2.1	The Hand-Aligned Case	72
4.2.2	Splitting	73
4.2.3	Merging	73
4.2.4	Smoothing	74
4.2.5	The Automatically-Aligned Case	75
4.3	Inference	75
4.4	Experiments	76
4.4.1	Phone Recognition	77
4.4.2	Phone Classification	79
4.5	Analysis	79
4.6	Summary and Future Work	82
5	Coarse-to-Fine Machine Translation Decoding	83
5.1	Introduction	83
5.2	Coarse-to-Fine Decoding	85
5.2.1	Related Work	85
5.2.2	Language Model Projections	86
5.2.3	Multipass Decoding	87
5.3	Inversion Transduction Grammars	88
5.4	Learning Coarse Languages	90
5.4.1	Random Projections	90
5.4.2	Frequency Clustering	90
5.4.3	HMM Clustering	91
5.4.4	JCluster	91
5.4.5	Clustering Results	91
5.5	Experiments	92
5.5.1	Clustering	93
5.5.2	Spacing	94
5.5.3	Encoding Versus Order	94

5.5.4 Final Results 95

5.5.5 Search Error Analysis 95

5.6 Summary and Future Work 97

6 Conclusions and Future Work..... 99

References..... 101