

Contents

1	Introduction	1
1.1	Motivation	2
1.2	Approach	2
1.3	Overview	4
2	Reinforcement Learning	7
2.1	Reinforcement Learning Framework	7
2.2	Temporal Difference Methods	9
2.3	Policy Search Methods	11
3	On-Line Evolutionary Computation	17
3.1	ϵ -Greedy Evolution	20
3.2	Softmax Evolution	21
3.3	Interval Estimation Evolution	22
3.4	Testbed Domains	22
3.4.1	Mountain Car	23
3.4.2	Server Job Scheduling	24
3.5	Results	26
3.6	Discussion	28
4	Evolutionary Function Approximation	31
4.1	NEAT+Q	33
4.2	Results	34
4.2.1	Comparing Manual and Evolutionary Function Approximation	34
4.2.2	Combining On-Line Evolution with Evolutionary Function Approximation	37
4.2.3	Comparing to Other Approaches	38
4.2.4	Comparing Darwinian and Lamarckian Approaches	42

4.2.5	Continual Learning Tests	43
4.3	Discussion	45
5	Sample-Efficient Evolutionary Function	
	Approximation	47
5.1	Sample-Efficient NEAT+Q	47
5.2	Results	49
5.3	Discussion	52
6	Automatic Feature Selection for Reinforcement	
	Learning	53
6.1	FS-NEAT	54
6.2	Testbed Domain	56
6.3	Results	58
6.4	Discussion	61
7	Adaptive Tile Coding	65
7.1	Background	66
7.1.1	Tile Coding	66
7.2	Method	68
7.2.1	When to Split	69
7.2.2	Where to Split	70
	subsubsection.7.2.2.1 subsubsection.7.2.2.2	
7.3	Testbed Domains	72
7.4	Results	73
7.5	Discussion	74
8	Related Work	77
8.1	Optimizing Representations	77
8.1.1	Supervised Learning	77
8.1.2	Reinforcement Learning	78
8.1.3	Evolutionary Computation	82
8.2	Combining Evolution and Learning	84
8.2.1	Applications to Supervised Learning	84
8.2.2	Applications to Reinforcement Learning	86
8.3	Balancing Exploration and Exploitation	89
8.3.1	k -Armed Bandit Problem	89
8.3.2	Associative Search	90
8.3.3	Reinforcement Learning	90
8.4	Feature Selection	92
8.4.1	Filters	92
8.4.2	Wrappers	93
8.4.3	FS-NEAT	93

9	Conclusion	95
9.1	Primary Conclusions	95
9.2	Negative Results	96
9.2.1	Combining FS-NEAT with NEAT+Q	96
9.2.2	Feature Selection in Adaptive Tile Coding	97
9.2.3	Fitness Functions Based on Bellman Error	97
9.3	Broader Implications	98
9.3.1	Stochastic vs. Deterministic Domains	98
9.3.2	The Value Function Gamble	99
9.3.3	The Role of Search in Adaptive Representations	100
9.4	Future Work	102
9.4.1	Non-stationarity	102
9.4.2	Steady-State Evolutionary Computation	102
9.4.3	Model-Based Reinforcement Learning	103
9.5	Final Remarks	104
A	Statistical Significance	105
	References	107