

進化論的手法によるブール関数の学習について

7N-5

丹羽竜哉, 伊庭斉志, 佐藤泰介, 樋口哲也

電子技術総合研究所

1. はじめに

ブール関数の学習は機械学習の基本的研究分野の一つである。この学習についてはアルゴリズム的アプローチによる健全かつ完全な学習法 (Decision trees, Enumeration 等) が知られており, また計算論的学習理論に基づいた学習能力の評価方法も提案されている。しかし, アルゴリズム的アプローチには計算量の問題があるため, ある程度の完全性を犠牲にして効果的な探索を目指す統計的または進化論的手法が幾つか提案されている。その代表例が, クラシファイヤシステム (以下 CS) や ALN (Adaptive Logic Networks), ニューラルネット (以下 NN) である。しかしながら, 計算論的学習理論の視点による, これらの手法の能力の比較研究はほとんどなされていない。

本稿では, これらの進化論的または統計的な手法による学習方法の能力の, PAC (Probably Approximately Correctly) に基づいた評価方法による評価結果を報告する。さらに, 学習時にノイズを加えた場合についても同様に評価する。

2. 進化論的手法による学習

本稿で評価する学習手法は, CS, ALN, および, それらの比較として NN(back propagation) についても評価を行なった。

CS *if ... then ...* ルールを遺伝的アルゴリズムによって更新し, 適応させて学習していく方法である [1]。本実験では, 集団サイズ:400, 交叉率:12%, 交叉型: 1点交叉, 突然変異率:0.1%, 報酬:1000, 誤り減衰率:80%, ワイルドカードバイアス:4.0 のパラメータを用いた。これは, [1] と同じ値である。

ALN あらかじめ用意した巨大な二分木の各ノードに割り当てられた論理関数 (AND, OR 等) を, 幾つかの法則に従って変化 / 削除して適切な関数を構成する学習方法である [2]。本実験では, 初期ノード数を 29999 個とした。この値は, [2] の一例に従っている。

NN 比較のため, ニューラルネットの誤差逆伝搬法による学習の実験を行なった。パラメータは, 学習率:0.01, モーメント:0.5, 中間層数:1, 中間ノード数:4 (dnf3 では

3) とした。

3. 学習能力の評価方法

学習能力の評価は, PAC に基づいて行なった [3]。N は関数の入力ビット数, K は目的関数を最小 DNF (Disjunctive Normal Form) 表現で記述するために必要なリテラルの数, ϵ は試験フェーズで許される最大誤り率とする。このとき, 我々の使った学習サンプルの数は次の式によって与えられる。

$$\frac{K \times \log_2 N}{\epsilon} \quad (1)$$

実験では $\epsilon = 10\%$ とし, 分類性能を試験するために 2000 個の試験サンプルを用いた。また, 学習の対象として, [3] の中の mx6 (6 マルチプレクサ, 16 ビット入力 (うち無効ビット 10)), par4 (4 ビットパリティ, 16 ビット入力 (うち無効ビット 12)), dnf3 (ランダム DNF, 32 ビット入力 (うち無効ビット 12), 6 項) の 3 つのブール関数について実験を行なった。

学習は収束に達してから停止させた。従って, 学習フェーズにおける繰り返しの回数は 3 つの手法によって異なり, CS では $O(1000)$, ALN では $O(100)$, NN では $O(10000)$ 程度である。しかし, 各々の繰り返しに対する計算は量的に異なっているので, この値は必ずしも計算論的な複雑さを反映するものではない。

ノイズを加えた場合の実験については [3] では考慮されていない。我々は, 学習サンプルの入力ビットを各々一定の確率で反転させたものを学習データとして与えた。試験データについてはノイズを加えていない。学習サンプルの数は, ノイズなしの場合と同じである。

	CS	ALN	NN
mx6	10 100.0 0.0	10 98.8 0.50	9 100.0 0.0
par4	10 100.0 0.0	10 98.6 1.63	3 100.0 0.0
dnf3	0 --- ---	10 87.6 0.94	0 --- ---

表1 ノイズのない場合の学習率

4. 結果および考察

我々は各々の関数に対して, 10 回ずつの独立した試行を行なった。表1はノイズなしの場合, 各々 10 回の試行のうち学習データの学習に成功 (正解率 100%) した回数, 学習データの学習に成功した時の試験データに

ノイズ	関数	CS				ALN				NN			
		学習		試験		学習		試験		学習		試験	
0%	mx6	100.0	0.0	100.0	0.0	100.0	0.0	98.9	0.7	99.0	1.2	98.7	1.4
	par4	100.0	0.0	100.0	0.0	100.0	0.0	98.6	1.3	89.1	12.7	85.9	18.3
	dnf3	90.0	1.8	87.8	3.1	100.0	0.0	87.6	1.6	96.7	0.8	92.7	3.0
2%	mx6	100.0	0.0	100.0	0.0	96.4	0.6	95.5	3.5	95.7	1.0	95.6	1.0
	par4	98.2	2.3	97.1	3.3	92.6	0.8	99.9	0.3	84.8	11.4	81.9	17.2
	dnf3	71.0	27.7	66.2	31.9	96.4	0.6	86.4	1.2	94.5	1.4	92.4	1.7
5%	mx6	98.3	2.5	98.4	2.4	90.9	1.2	99.8	0.4	90.7	1.1	90.1	0.5
	par4	44.8	2.2	36.6	1.8	74.4	6.7	71.5	23.7	76.8	8.8	74.7	13.1
	dnf3	27.0	29.0	20.3	32.2	90.1	0.8	89.9	2.5	92.0	0.8	90.7	0.8

表2 ノイズのある場合の学習率

対する正答率の平均、およびその分散(--- は10回とも失敗したことを示す)を示す。

表2は、3つの手法による学習および試験データの正答率の平均と標準偏差である。ノイズ0%の環境は表1と同じであるが、表2ではノイズの影響を評価するために、10回すべての試行の平均と標準偏差を計算した。

CS mx6とpar4において他の2つの手法に比べて優れ、完全なルールセットを導き出している。このルールセットは、無効なビットは無視し、関係のあるビットだけを抽出したものである。これに対して、dnf3の学習には失敗している。これは、CSにとってdnf3の環境をクラシファイアのルールで表現することが困難なことによる。すなわち、mx6やpar4では#(ワイルドカード)が効果的に作用するのに対し、dnf3ではほとんどのルールが0を出力するものとなっている。ルールの数を多くする(約2000個)と、dnf3でもかなり良く学習できる[4]。

ノイズを加えた場合、ノイズがあるレベルを越えると、突然に性能が落ちる。また、CSの性能は、個々の試行毎に不安定(標準偏差が大)である。

ALN 3つすべての関数に対して比較的良い性能を示している。しかし、表1に見られるように(dnf3に対する性能は90%以下)、学習データの汎化には成功していない。この結果は、ALNが単純に学習サンプルを記憶し、汎化の能力を欠いていることを示す。また、一般にALNは非常に多くの初期ノードを要する(例えば、[2]では $O(60000)$ のものもある)。最終的なノード数は妥当な数($O(100)$)であるが、初期ノードが少ないと学習は成功しない。

ノイズを加えた場合、mx6(2%と5%ノイズ)とpar4(0%と2%ノイズ)に見られるように、過学習の現象が起きている。

NN 表1に見られるように、NNは学習データの学

習に常に成功するとは限らない。NNは、mx6やpar4に関しては貧弱な結果を示している。これは、NNの分散表現がmx6やpar4の無効ビットの区別の妨げになっているからである。また、dnf3はNNに対して難しい問題である。

ノイズに対しては、比較的うまく処理している。つまり、ノイズのレベルの上昇に対して学習性能の降下が穏やかである。

5. おわりに

本稿では、進化論的/統計的な代表的な3種類の学習方法の能力を、PACに基づいた評価方法によって評価した。さらに、ノイズを加えた場合についても評価を行なった。

これら3つの手法は、各々長所と短所を持っている。現在我々は、これらの手法各々の特徴を合わせ持つ、複合的な学習方法を研究中である[4]。

参考文献

- [1] S.W.Wilson, "Classifier Systems and the Animat Problem", *Machine Learning*, vol.2, no.3, 1987.
- [2] W.W.Armstrong, "Learning and Generalization in Adaptive Logic Networks", *Artificial Neural Networks*, Elsevier Science Pub., 1991.
- [3] G.Pagallo, and D.Hausslear, "Boolean Feature Discovery in Empirical Learning", *Machine Learning*, vol.5, 1990.
- [4] 丹羽, 伊庭, 佐藤, 樋口, "ブール関数の進化論的手法による学習について", 電気学会情報処理研究会論文集 IP-93-14, 1993.