7A-05

順序木言語のプリミティブ形式体系に対する１つの正例と所属性質問による学習アルゴリズム

松本哲志† 内田智之‡ 正代隆義* 鈴木祐介‡ 宮原哲浩‡

†東海大学理学部情報数理学科 ‡広島市立大学大学院情報科学研究科

*九州国際大学現代ビジネス学部

1. はじめに

順序木言語[2,3]のプリミティブ形式体系(primitive Formal Ordered Tree System, pFOTS)とは,Uchida[4]らによって提案された形式グラフ体系の一種である．pFOTSはラベル付き順序木言語を定義する形式体系であり，１階述語論理の項の代わりに項木パタンを用いた論理プログラミングシステムである．pFOTSプログラムは，確定節であるグラフ書き換えルールの有限集合として定義される．そのため，pFOTSプログラムは順序木言語の背景知識の表現に適している．質問学習モデルは，計算論的学習理論における質問を用いた学習の数学的モデルである[1]． 本稿では, pFOTSプログラムによる背景知識 Γ と１つのグラフ書き換えルールで定義される順序木言語のクラスが，１つの正例と多項式回の所属性質問を用いた質問学習モデルにより学習可能となる背景知識 Γ の条件を示す.

2. プリミティブ形式体系

Σ と Λ を有限アルファベット， Xを無限アルファベットとする．Xの要素を変数ラベルとよぶ．$(\Sigma\cup\Lambda)\cap X=\phi$ が成り立つとする．$\langle\Sigma,\Lambda\cup X\rangle$上の頂点ラベルと辺ラベル付き順序木 $t=(V_t,E_t)$とは頂点ラベル付け関数 $\Phi_t:V_t\rightarrow\Sigma$ と辺ラベル付け関数 $\Psi_t:E_t\rightarrow\Lambda\cup X$をもつ順序木である．変数ラベルをもつ辺を変数とよぶ. $x\in X$に対して, $o(t,x)$は変数ラベルとしてxをもつtにおける変数の個数を表す. $\langle\Sigma,\Lambda\cup X\rangle$上の線形項木パタン $t=(V_t,E_t)$とは，任意の$x\in X$に対して，$o(t,x)\leqq 1$となる頂点ラベルと辺ラベル付き順序木である．変数をもたない線形項木パタンは頂点ラベルと辺ラベル付き順序木である． 以後， 線形項木パタンを**項木パタン**，頂点ラベルと辺ラベル付き順序木を木とよぶ．木全体集合を***OT***で表す．2つの頂点と１つの変数から成る項木パタンはプリミティブであるという．

Exact Learning Algorithms for Primitive Formal Ordered Tree Systems by One Positive Example and Membership Queries

† Satoshi Matsumoto, Faculty of Science, Tokai University

‡ Tomoyuki Uchida, Yusuke Suzuki, Tetsuhiro Miyahara Graduate School of Information Sciences, Hiroshima City University

*Takayoshi Shoudai, Faculty of Contemporary Business, Kyushu International University

fと g を少なくとも 2 つの頂点をもつ項木パタンとする．$\sigma=[u,v]$を g の異なる頂点の対とし， xを X の変数ラベルとする．形式 $x:=[g,\sigma]$を$\langle\Sigma,\Lambda\cup X\rangle$上の**束縛**とよぶ．束縛 $x:=[g,\sigma]$を f に適用して得られる項木パタン $f\{x:=[g,\sigma]\}$を次のような操作により得られる項木パタンとして定める．$e=\langle s,t\rangle$を変数ラベル x でラベル付けされた fの変数とする．g'を gと同型な項木パタンとし，gのuと vに対応する頂点をそれぞれ u'と v'とする．fから eを取り除き， sを u'と, tを v'と同一視して g'を fに貼り付ける． **代入**とは束縛の有限集合である．項木パタン f に対して代入を適用したとき,代入に含まれる束縛を全て同時に f に適用した項木パタンが得られる．

Πを１引数述語記号の集合とする． １引数述語記号 $p\in\Pi$と項木パタン tからなる表現，$p(t)$をアトムとよぶ． $A, B_1, \cdots, B_m$をアトムとする($m\geqq 0$)．$\langle\Pi,\Sigma,\Lambda\cup X\rangle$上の**グラフ書き換えルール**(graph rewriting rule)とは， $A\leftarrow B_1,\cdots,B_m$という形式の確定節である． 以後, **ルール**とよぶ． A を**頭部**, $B_1,\cdots,B_m$を**本体**とよぶ．$m=0$のときのルールを**ファクト**とよぶ．ルール$p(t)\leftarrow q_1(f_1),\cdots,q_m(f_m)$は次の条件(1)-(3)を満たすとき, **プリミティブ**であるという： (1) $m\geqq 1$のとき， $f_i(1\leqq i\leqq m)$はプリミティブな項木パタンである， (2) $m=0$のとき，tは頂点数２の木である， (3)任意の変数ラベル $x\in X$に対して，$o(f_1,x)+\cdots+o(f_m,x)=1$となる必要十分条件は$o(t,x)=1$である．プリミティブなルールの有限集合を**プリミティブ形式体系(pFOTS)プロ**

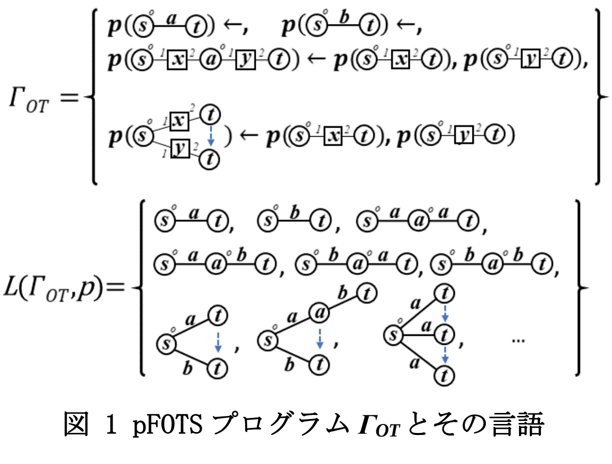

図 1 pFOTS プログラム Γ_{OT} とその言語

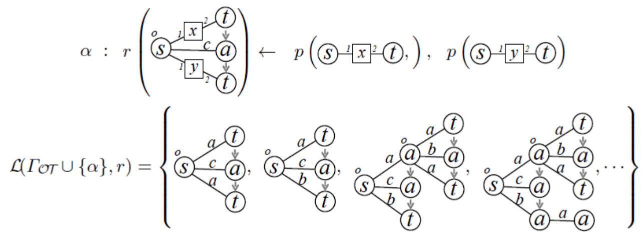

図 2 pFOTS プログラム Γ_{OT} とルール α によって生成される言語 $L(\Gamma_{OT} \cup \{\alpha\}, r)$

グラムとよぶ. 図 1 に pFOTS プログラム Γ_{OT} とその言語の例を与える. $\langle \Pi, \Sigma, \Lambda \cup X \rangle$ 上の pFOTS プログラム Γ と Π の述語記号 p に対して, $L(\Gamma, p)$={$g \in$ ***OT*** | $p(g)$は Γ のルールと代入, モーダスポネンスを使って証明可能なアトム}と定める.

3. 質問する学習モデル

$\langle \Pi, \Sigma, \Lambda \cup X \rangle$ 上の Γ を pFOTS プログラムとし, 背景知識とする. 学習アルゴリズムは事前に Γ に関する情報を得ているものとする. $\Pi^h(\Gamma)$ を Γ に含まれるルールの頭部に現れる述語記号全体とし, $\Pi^b(\Gamma)$ を Γ に含まれるルールの本体に現れる述語記号全体とする. さらに, $\Pi(\Gamma) = \Pi^h(\Gamma) \cup \Pi^b(\Gamma)$ と定める. $\Pi \setminus \Pi(\Gamma)$ の述語記号 r に対して, ***pGRR***(Γ, r) は $\Pi^h(\{\alpha\})=\{r\}$ かつ $\Pi^b(\{\alpha\}) \subseteq \Pi^h(\Gamma)$ となるプリミティブなルール α の全体集合を表す. 本稿では, 言語クラス $\{L(\Gamma \cup \{\alpha\}, r) \mid \alpha \in$ ***pGRR***$(\Gamma, r)\}$ の学習可能性について考える. 学習モデルとして Angluin[1]によって提唱された質問学習モデルを用いる. ルール α_* を ***pGRR***(Γ, r) の目標ルールとする. $L(\Gamma \cup \{\alpha_*\}, r)$ に含まれる木を正例とよぶ. 図 2 に例を与える. 学習アルゴリズムはオラクルを用いて目標ルール α_* に関する情報を獲得することが可能である. 本稿で用いる所属性質問では, オラクルに木 t を入力し, t が正例ならばオラクルは yes を返し, それ以外ならば no を返す.

4. プリミティブなルールの学習可能性

Γ を $\langle \Pi, \Sigma, \Lambda \cup X \rangle$ 上の pFOTS プログラムとする. a を Λ の辺ラベルとする. P(a)は a を辺ラベルとしてもつファクトの述語記号の集合と定める. 任意の述語記号 $p \in \Pi(\Gamma)$ に対して, $P(a_1) \cap P(a_2)=\{p\}$ となる相異なる辺ラベル a_1, a_2 が存在するとき, Γ は条件 1 を満たすという. Γ に含まれるファクト以外のルールの頭部に現れる項木パタンの集合を $H(\Gamma)$ とする.

定理 1: $|\Lambda| \geqq 2$ とする. Γ を $|\Pi(\Gamma)|=1$ となる pFOTS プログラムとし, r を $\Pi \setminus \Pi(\Gamma)$ の述語記号とする. 正例となる木 t が与えられたとき,学習アルゴリズム LEARNING_pFTOS(図 3)は ***pGRR***(Γ, r) の目標ルールを $O(N^2)$回の所属性質問を用いて同定可能である. ただし, N は t の辺の個数を表す.

同様なアルゴリズムを用いて次の系が得られる.

アルゴリズム LEARNING_pFTOS
入力:正例となる木 $t \in$ ***OT***
出力:α_*と同じ言語を生成できるルール α
1:**do**
2: t のある部分木が $H(\Gamma)$ の中のある項木パタンと形が同じであるならば, その部分木を 1 つの辺に置き換えたものを t' とする.
3: t' に対して所属性質問を行い, yes が返ってきた場合 t' を t とする.
4:**while** t が変更される.
5:h を t のコピーとする.
6: t のそれぞれの辺に対して, 別ラベルをもつ辺に置き換えて所属性質問を行い, yes が返ってきた場合は, その辺に対応する h の辺を変数とする.
7:h をルール α に書き換えて出力する.

図 3 学習アルゴリズム LEARNING_pFTOS

系 1: Γ を条件 1 と $|\Pi(\Gamma)| \geqq 2$ を満たす pFOTS プログラムとし, r を $\Pi \setminus \Pi(\Gamma)$ の述語記号とする. 正例となる木 t が与えられたとき, ***pGRR***(Γ, r) の目標ルールは $O(N^2)$回の所属性質問を用いて同定可能である. ただし, N は t の辺の個数を表す.

系 2: k を正整数とする. $\Gamma_1, \ldots, \Gamma_k$ を $\cup_{1 \leqq j \leqq k} \Gamma_j$ が条件 1 を満たし, 任意の i ($1 \leqq i \leqq k$)に対して $|\Pi(\Gamma_i)|=1$,かつ $i \neq j$ となる任意の i,j ($1 \leqq i,j \leqq k$)に対して $\Pi(\Gamma_i) \cap \Pi(\Gamma_j)=\phi$ となる pFOTS プログラムとする. r を $\Pi \setminus \cup_{1 \leqq j \leqq k} \Pi(\Gamma_i)$ の述語記号とする. 正例となる木 t が与えられたとき, ***pGRR***$(\cup_{1 \leqq j \leqq k} \Gamma_j, r)$ の目標ルールは $O(N^2)$回の所属性質問を用いて同定可能である. ただし, N は t の辺の個数を表す.

5. おわりに

本稿では, pFOTS プログラムによる背景知識 Γ と 1 つのルールで定義される順序木言語のクラスが, 1 つの正例と所属性質問を用いた質問学習モデルにより学習可能となる Γ の条件を示した.

謝辞 本研究は JSPS 科研費 JP15K00312, JP15K00313, JP17K00321 の助成を受けたものです.

参考文献

1. D. Angluin: Queries and concept learning. Machine Learning 2(4), pp.319–342 (1988)
2. S. Matsumoto et al. : Learning of finite unions of tree patterns with internal structured variables from queries. IEICE Trans. Inf. & Syst. E91-D(2), pp. 222–230 (2008)
3. Y. Suzuki et al.: Ordered term tree languages which are polynomial time inductively inferable from positive data. TCS 350(1), pp. 63–90 (2006)
4. T. Uchida et al.: Parallel algorithms for refutation tree problem on formal graph systems. IEICE Trans. Inf. & Syst. E78-D(2), pp. 99–112 (1995)