ANZIAM J. 58(2017), 211–219
doi:10.1017/S1446181116000316

AN ALTERNATIVE METHOD OF CONCEPT LEARNING

SEN WANG[1], QINGXIANG FANG[2] and JUN-E FENG[✉1]

(Received 12 June, 2016; accepted 15 October, 2016; first published online 6 March 2017)

Abstract

We solve the problem of concept learning using a semi-tensor product method. All possible hypotheses are expressed under the framework of a semi-tensor product. An algorithm is raised to derive the version space. In some cases, the new approach improves the efficiency compared to the previous approach.

2010 *Mathematics subject classification*: 68T01.

Keywords and phrases: concept learning, version space, semi-tensor product, target concept, all possible hypotheses.

1. Introduction

Concept learning has become a significant problem in artificial learning. It involves deriving general concepts from positive and negative training examples. For instance, there is a task of learning to predict the value of an attribute *EnjoySport* as in Table 1, based on the value of six other attributes, such as *Sky, AirTemp, Humidity, Wind, Water* and *Forecast*. Some of the training examples are provided in Table 1 [6].

Suppose that the concept is a rule by which we determine the value of an attribute based on n other attributes. Since each attribute is two-valued, it can be represented by a Boolean variable. Thus, the concept to be learned can be expressed as a Boolean function $f : \mathcal{D}^n \to \mathcal{D}$, where $\mathcal{D}$ denotes the set $\{0, 1\}$. In this paper, we consider the common case that the concept is composed of a conjunction of constraints on instance attributes. In other words, the concept $y = f(x)$, where $y \in \mathcal{D}$, $x = (x_1, x_2, \ldots, x_n)$, $x_i \in \mathcal{D}$, $i = 1, 2, \ldots, n$, can be equivalently described as

$$y = f_1(x_1) \wedge f_2(x_2) \wedge \cdots \wedge f_n(x_n),$$

[1]School of Mathematics, Shandong University, Jinan 250100, China; e-mail: 17862988175@163.com, fengjune@sdu.edu.cn.
[2]School of Science, China Jiliang University, Hangzhou 310018, China; e-mail: fangqx@cjlu.edu.cn.

TABLE 1. Positive and negative training examples for the target concept *EnjoySport*.

Example	Sky	AirTemp	Humidity	Wind	Water	Forecast	EnjoySport
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	Cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

where $f_i : \mathcal{D} \to \mathcal{D}$, $i = 1, 2, \ldots, n$, are a series of Boolean functions and $\wedge$ denotes the corresponding conjunction operation. Let the scalar form of a Boolean value, 0 and 1, be equivalently expressed in the vector form, $(0, 1)^{\mathrm{T}}$ and $(1, 0)^{\mathrm{T}}$, respectively. The conjunction operation "$\wedge$" of two Boolean values in vector form is defined as

$$\begin{bmatrix}1\\0\end{bmatrix} \wedge \begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}, \quad \begin{bmatrix}1\\0\end{bmatrix} \wedge \begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix},$$
$$\begin{bmatrix}0\\1\end{bmatrix} \wedge \begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix}, \quad \begin{bmatrix}0\\1\end{bmatrix} \wedge \begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0\\1\end{bmatrix}.$$

Define the target concept as the concept to be learned and denote it by c, that is, $c : \mathcal{D}^n \to \mathcal{D}$. The task is to hypothesize or estimate c. We use the symbol H to denote the set of all possible hypotheses. In this paper,

$$H = \{h : \mathcal{D}^n \to \mathcal{D} \mid \exists h_i : \mathcal{D} \to \mathcal{D}, i = 1, 2, \ldots, n, \quad \text{such that}$$
$$h(x_1, x_2, \ldots, x_n) \equiv h_1(x_1) \wedge h_2(x_2) \wedge \cdots \wedge h_n(x_n), x_j \in \mathcal{D},\ j = 1, 2, \ldots, n\}.$$

Note that $c \in H$. Now we can write the ordered pair $\langle x, c(x)\rangle$, with $x \in \mathcal{D}^n$, to describe a training example. Let D be the set of training examples. We define that a hypothesis h is consistent with a set of training examples D if $h(x) = c(x)$ for any $\langle x, c(x)\rangle$ in D. We denote it as $consistent(h, D)$.

The version space, $VS_{H,D}$, is defined as

$$VS_{H,D} \equiv \{h \in H \mid consistent(h, D)\}.$$

A candidate-elimination approach is presented here to derive the version space [7].

The semi-tensor product (STP), presented by Cheng, becomes a powerful tool to study Boolean networks [1]. In recent years, many fruitful results have been obtained via STP [2–4, 8]. This paper provides an effective approach to derive the version space using STP.

The rest of the paper is organized as follows. Section 2 introduces some definitions and notations of STP. In Section 3, main results are derived. Based on them, an algorithm is presented to obtain the version space. Finally, a comparison between the new approach and the candidate-elimination algorithm is given in the concluding remarks in Section 4.

https://doi.org/10.1017/S1446181116000316 Published online by Cambridge University Press

2. Preliminaries

In this section, we introduce some symbols and definitions used in this paper.

Let δ_n^i be the ith column of the identity matrix I_n and

$$\Delta_n = \{\delta_n^1, \delta_n^2, \ldots, \delta_n^n\}.$$

We simply use $\Delta = \Delta_2$ when $n = 2$.

For a matrix A, let $\mathrm{Col}(A)$ and $\mathrm{Row}(A)$ be the sets of columns and rows of A, respectively. A matrix $L \in M_{n\times s}$ is called a logical matrix if $\mathrm{Col}(L) \subset \Delta_n$. Denote the set of $n \times s$ logical matrices by $\mathcal{L}_{n\times s}$. Write the ith column of matrix A as $\mathrm{col}_i(A)$ and the ith row of matrix A as $\mathrm{row}_i(A)$.

Let $A = (a_{ij}) \in R_{m\times n}$ and $B = (b_{ij}) \in R_{p\times q}$, and denote the least common multiple of n and p by t. Then the STP of A and B is defined as

$$A \ltimes B = (A \otimes I_{t/n})(B \otimes I_{t/p}) \in R_{mt/n\times qt/p},$$

where $\otimes$ is the Kronecker product [1].

Since STP is a generalization of the conventional matrix product, n is omitted from the symbol $\ltimes_{i=1}^n$ when no ambiguity occurs.

Lemma 2.1. *If $x \in \Delta_{2^n}$ is given, there exist $x_1, x_2, \ldots, x_n \in \Delta$ such that $x = \ltimes_{i=1}^n x_i$, and each x_i is uniquely determined.*

For any set C, let $|C|$ denote the cardinality (the number of elements) of C.

3. Main results

In this section, we adopt the vector form of Boolean values. Consider the necessary and sufficient condition where a function

$$y = Mx, \tag{3.1}$$

$M \in \mathcal{L}_{2\times 2^n}$, $x = \ltimes_{i=1}^n x_i$, $x_i \in \Delta$, can also be equivalently expressed as

$$y = (M_1x_1) \wedge (M_2x_2) \wedge \cdots \wedge (M_nx_n), \tag{3.2}$$

$M_k \in \mathcal{L}_{2\times 2}$, $k = 1, 2, \ldots, n$. Here x is the argument. Each M_k consists of two columns. If there exists a column $(1, 0)^{\mathrm{T}}$ in M_k, we write the corresponding column number as i_j^k, $1 \leqslant j \leqslant 2$. Set $C_k = \{i_1^k\}$ when M_k has only one $(1, 0)^{\mathrm{T}}$ column, or $C_k = \{i_1^k, i_2^k\}$ when M_k has two $(1, 0)^{\mathrm{T}}$ columns, or $C_k = \emptyset$ when M_k has no $(1, 0)^{\mathrm{T}}$ column. Note that given equation (3.2), there always exists $M \in \mathcal{L}_{2\times 2^n}$ such that $Mx = (M_1x_1) \wedge (M_2x_2) \wedge \cdots \wedge (M_nx_n)$ for all $x = \ltimes_{i=1}^n x_i, x_i \in \Delta, i = 1, 2, \ldots, n$.

Proposition 3.1. *Given $M_k \in \mathcal{L}_{2\times 2}$, $k = 1, 2, \ldots, n$, there exists $x = \ltimes_{i=1}^n x_i, x_i \in \Delta, i = 1, 2, \ldots, n$, such that $(M_1x_1) \wedge (M_2x_2) \wedge \cdots \wedge (M_nx_n) = (1, 0)^{\mathrm{T}}$ if and only if for all $k \in \{1, 2, \ldots, n\}$, $C_k \neq \emptyset$.*

PROOF. Consider $M_ix_i = \mathrm{col}_1(M_i)$ if $x_i = (1,\ 0)^{\mathrm{T}}$, and $M_ix_i = \mathrm{col}_1(M_i)$ if $x_i = (0,\ 1)^{\mathrm{T}}$. There exists $x = \ltimes_{i=1}^{n} x_i$ such that $(M_1x_1) \wedge (M_2x_2) \wedge \cdots \wedge (M_nx_n) = (1,\ 0)^{\mathrm{T}}$ if and only if each M_i has a column $(1,\ 0)^{\mathrm{T}}$. □

Suppose that equation (3.1) is equivalently written as equation (3.2). In fact, if there exists a $(1,\ 0)^{\mathrm{T}}$ column in M, we can obtain that each M_k has at least one $(1,\ 0)^{\mathrm{T}}$ column and each corresponding $(1,\ 0)^{\mathrm{T}}$ column is determined. It is demonstrated in the following examples.

EXAMPLE 3.2. Assume that $M = \delta_2[1,\ 2,\ 2,\ 2]$. Since $\mathrm{col}_1(M) = (1,\ 0)^{\mathrm{T}}$, when $x = \delta_4^1$, $Mx = 1$. Consider that $x = x_1 \ltimes x_2$, so $x_1 = \delta_2^1$ and $x_2 = \delta_2^1$. Then we can conclude that $\mathrm{col}_1(M_1) = (1,\ 0)^{\mathrm{T}}$ and $\mathrm{col}_1(M_2) = (1,\ 0)^{\mathrm{T}}$. Thus, $C_1 = \{1\}$, and $C_2 = \{1\}$.

EXAMPLE 3.3. Suppose that $M = \delta_2[1, 1, 2, 2]$. Similarly, we have $C_1 = \{1\}$, $C_2 = \{1,\ 2\}$.

EXAMPLE 3.4. If $M = \delta_2[1, 2, 1, 1]$, we will show that equation (3.1) cannot be written as equation (3.2). From $\mathrm{col}_1(M) = (1, 0)^{\mathrm{T}}$, we can derive that $\mathrm{col}_1(M_1) = (1,\ 0)^{\mathrm{T}}$, $\mathrm{col}_1(M_2) = (1, 0)^{\mathrm{T}}$. Similarly, $\mathrm{col}_2(M_1) = (1, 0)^{\mathrm{T}}$ and $\mathrm{col}_2(M_2) = (1, 0)^{\mathrm{T}}$. If $Mx = (M_1x_1) \wedge (M_2x_2) \wedge \cdots \wedge (M_nx_n)$, $x = \ltimes_{i=1}^{n} x_i$, $x_i \in \Delta$, $i = 1,\ 2, \ldots,\ n$, then $\mathrm{col}_2(M) = (1, 0)^{\mathrm{T}}$, which is a contradiction.

PROPOSITION 3.5. *If equation (3.1) is equivalently written as equation (3.2) and* $C_k \neq \emptyset, k = 1,\ 2, \ldots,\ n$, *then* $|\{i \mid \mathrm{col}_i(M) = (1,\ 0)^{\mathrm{T}}\}| = 2^{|\{k \mid |C_k|=2\}|}$.

PROOF. From equation (3.2), in order that $Mx = (1,\ 0)^{\mathrm{T}}$, let $x = \ltimes_{i=1}^{n} x_i$ take values as follows. If $|C_k| = 1$, we take $x_k = i_1^k$. If $|C_k| = 2$, we take $x_k = i_1^k$ or i_2^k. Therefore, there are two values to be taken for x_k when $|C_k| = 2$. Thus, x can take $2^{|\{k \mid |C_k|=2\}|}$ values to make $Mx = (1,\ 0)^{\mathrm{T}}$. Now, the result follows immediately. □

Note that even if $|\{i \mid \mathrm{col}_i(M) = (1,\ 0)^{\mathrm{T}}\}| = 2^{|\{k \mid |C_k|=2\}|}$, we cannot conclude that equation (3.1) can be equivalently expressed as equation (3.2).

PROPOSITION 3.6. *Suppose that* $|\{i \mid \mathrm{col}_i(M) = (1,\ 0)^{\mathrm{T}}\}| = 2^m$, $m \in \mathbb{Z}^+$ *and* $\mathrm{col}_{i_j}(M) = (1,\ 0)^{\mathrm{T}}$, $j = 1,\ 2, \ldots,\ 2^m$. *Let* $\ltimes_{k=1}^{n} z_k^j = \delta_{2^n}^{i_j}$, $j = 1,\ 2, \ldots,\ 2^m$, $z_k^j \in \Delta$. *Write* $z_k^j = \delta_2^{p_{k,j}}$. *If* $Mx = (M_1x_1) \wedge (M_2x_2) \wedge \cdots \wedge (M_nx_n)$ *for all* $x = \ltimes_{i=1}^{n} x_i$, $x_i \in \Delta, i = 1,\ 2, \ldots,\ n$, *then* $\mathrm{col}_{p_{k,j}}(M_k) = (1,\ 0)^{\mathrm{T}}$, $j = 1,\ 2, \ldots,\ 2^m, k = 1,\ 2, \ldots,\ n$.

PROOF. Consider that $\mathrm{col}_{i_j}(M) = (1, 0)^{\mathrm{T}}$. Thus, $M\delta_{2^n}^{i_j} = (1,\ 0)^{\mathrm{T}}$. That is, when $x = \ltimes_{k=1}^{n} x_k = \delta_{2^n}^{i_j}$, $Mx = (1, 0)^{\mathrm{T}}$. Therefore, when $x_k = \delta_{2^n}^{p_{k,j}}$, $k = 1,\ 2, \ldots,\ n$, $Mx = (M_1x_1) \wedge (M_2x_2) \wedge \cdots \wedge (M_nx_n) = (1, 0)^{\mathrm{T}}$. It is clear that $\mathrm{col}_{p_{k,j}}(M_k) = (1,\ 0)^{\mathrm{T}}$. □

EXAMPLE 3.7. Consider the matrix $M = \delta_2[1, 2, 1, 1]$ with $\mathrm{col}_3(M) = (1, 0)^{\mathrm{T}}$, and $\delta_4^3 = \delta_2^2 \ltimes \delta_2^1$. Then $\mathrm{col}_2(M_1) = (1, 0)^{\mathrm{T}}$ and $\mathrm{col}_1(M_2) = (1, 0)^{\mathrm{T}}$. Parallel results about other columns of M can be similarly derived.

https://doi.org/10.1017/S1446181116000316 Published online by Cambridge University Press

Given $\delta_{2^n}^{i_j}$ and $k \in \{1, 2, \ldots, n\}$, $z_k^j = S_k^n \delta_{2^n}^{i_j}$, where S_k^n is defined as follows: [2]

$$S_1^n = \delta_2[\underbrace{1, \ldots, 1}_{2^{n-1}}, \underbrace{2, \ldots, 2}_{2^{n-1}}],$$

$$S_2^n = \delta_2[\underbrace{1, \ldots, 1}_{2^{n-2}}, \underbrace{2, \ldots, 2}_{2^{n-2}}, \underbrace{1, \ldots, 1}_{2^{n-2}}, \underbrace{2, \ldots, 2}_{2^{n-2}}],$$

$$\vdots$$

$$S_n^n = \delta_2[1, 2, 1, 2, \ldots, 1, 2].$$

Combining this with Proposition 3.6, we obtain the following result.

PROPOSITION 3.8. *Suppose that*

$$\{i \mid \mathrm{col}_i(M) = (1, 0)^{\mathrm{T}}\} = \{i_1, i_2, \ldots, i_{2^m}\}$$

and

$$Mx = (M_1x_1) \wedge (M_2x_2) \wedge \cdots \wedge (M_nx_n) \textit{ for all } x = \ltimes_{i=1}^n x_i, x_i \in \Delta, i = 1, 2, \ldots, n.$$

For any $k \in \{1, 2, \ldots, n\}$, *if* $\mathrm{col}(S_k^n[\delta_{2^n}^{i_1}, \delta_{2^n}^{i_2}, \ldots, \delta_{2^n}^{i_{2^m}}])$ *contains* δ_2^1, *then* $\mathrm{col}_1(M_k) = (1, 0)^{\mathrm{T}}$ *and, if* $\mathrm{col}(S_k^n[\delta_{2^n}^{i_1}, \delta_{2^n}^{i_2}, \ldots, \delta_{2^n}^{i_{2^m}}])$ *contains* δ_2^2, *then* $\mathrm{col}_2(M_k) = (1, 0)^{\mathrm{T}}$.

EXAMPLE 3.9. For the matrix $M = \delta_2[1, 2, 1, 1]$, observe that $\mathrm{col}_1(M) = \mathrm{col}_3(M) = \mathrm{col}_4(M) = (1, 0)^{\mathrm{T}}$. We calculate that

$$S_1^2[\delta_4^1, \delta_4^3, \delta_4^4] = [\delta_2^1, \delta_2^2, \delta_2^2].$$

Then $\mathrm{col}_1(M_1) = \mathrm{col}_2(M_1) = (1, 0)^{\mathrm{T}}$. Similarly, from $S_2^2[\delta_4^1, \delta_4^3, \delta_4^4] = [\delta_2^1, \delta_2^1, \delta_2^2]$, we can derive that $\mathrm{col}_1(M_2) = \mathrm{col}_2(M_2) = (1, 0)^{\mathrm{T}}$.

Conversely, suppose that there is a set $\{i_1, i_2, \ldots, i_s\} \subset \{1, 2, \ldots, 2^n\}$ such that $\mathrm{col}_{i_j}(M) = (1, 0)^{\mathrm{T}}$, $j = 1, 2, \ldots, s$, and $M_k \in \mathcal{L}_{2\times 2}$ satisfying the following conditions: for any $k \in \{1, 2, \ldots, n\}$, if $\mathrm{col}(S_k^n[\delta_{2^n}^{i_1}, \delta_{2^n}^{i_2}, \ldots, \delta_{2^n}^{i_s}])$ contains δ_2^1, then $\mathrm{col}_1(M_k) = (1, 0)^{\mathrm{T}}$ and, if $\mathrm{col}(S_k^n[\delta_{2^n}^{i_1}, \delta_{2^n}^{i_2}, \ldots, \delta_{2^n}^{i_s}])$ contains δ_2^2, then $\mathrm{col}_2(M_k) = (1, 0)^{\mathrm{T}}$, and other columns of M_k are $(0, 1)^{\mathrm{T}}$. Let $X_1 = \{x \mid x = \ltimes_{i=1}^n x_i, x_i \in \Delta, i = 1, 2, \ldots, n, (M_1x_1) \wedge (M_2x_2) \wedge \cdots \wedge (M_nx_n) = (1, 0)^{\mathrm{T}}\}$; then $Mx = (M_1x_1) \wedge (M_2x_2) \wedge \cdots \wedge (M_nx_n)$ for all $x = \ltimes_{i=1}^n x_i \in X_1$, $x_i \in \Delta, i = 1, 2, \ldots, n$.

Suppose that $\{i \mid \mathrm{col}_i(M) = (1, 0)^{\mathrm{T}}\} = \{i_1, i_2, \ldots, i_{2^m}\}$ and $\{i \mid \mathrm{col}_i(M) = (0, 1)^{\mathrm{T}}\} = \{q_1, q_2, \ldots, q_t\}$. Now set $\overline{C}_k$ as follows.

$$\overline{C}_k = \{i \mid i \in \{1, 2\}, \delta_2^i \in \mathrm{col}(S_k^n[\delta_{2^n}^{i_1}, \delta_{2^n}^{i_2}, \ldots, \delta_{2^n}^{i_{2^m}}])\}.$$

Given $j \in \{1, 2, \ldots, t\}$, let $S_k^n \delta_{2^n}^{q_j} = \delta_2^{\gamma_j^k}$, $k = 1, 2, \ldots, n$. Then we have the following result.

PROPOSITION 3.10. *If* $Mx = (M_1x_1) \wedge (M_2x_2) \wedge \cdots \wedge (M_nx_n)$ *for all* $x = \ltimes_{i=1}^n x_i, x_i \in \Delta, i = 1, 2, \ldots, n$, *then there exists* $k \in \{1, 2, \ldots, n\}$ *such that* $\mathrm{col}_{\gamma_j^k}(M_k) = (0, 1)^{\mathrm{T}}$.

https://doi.org/10.1017/S1446181116000316 Published online by Cambridge University Press

PROOF. For all $k \in \{1, 2, \dots, n\}$, assume that $\mathrm{col}_{\gamma_j^k}(M_k) = (1, 0)^{\mathrm{T}}$. It is natural that $\mathrm{col}_{q_j}(M) = (1, 0)^{\mathrm{T}}$, which contradicts the construction of $\{q_1, q_2, \dots, q_t\}$. □

Conversely, suppose that there are some matrices $M_k \in \mathcal{L}_{2\times 2}$, $k = 1, 2, \dots, n$. Consider $(1, 2)\delta_2^{\gamma} = \gamma$, $\gamma \in \{1, 2\}$. Define two sets

$$K = \{i \mid i \in \{1, 2, \dots, 2^n\}, \exists k \in \{1, 2, \dots, n\} \text{ such that } \mathrm{col}_{(1,2)S_k^n \delta_{2^n}^i}(M_k) = (0, 1)^{\mathrm{T}}\}$$

and

$$X_2 = \{x \mid x = \ltimes_{i=1}^n x_i,\ x_i \in \Delta, i = 1, 2, \dots, n, \\ (M_1 x_1) \wedge (M_2 x_2) \wedge \cdots \wedge (M_n x_n) = (0, 1)^{\mathrm{T}}\}.$$

If there exists a matrix $M \in \mathcal{L}_{2\times 2^n}$ such that for any $i \in K$, $\mathrm{col}_i(M) = (0, 1)^{\mathrm{T}}$, then $Mx = (M_1 x_1) \wedge (M_2 x_2) \wedge \cdots \wedge (M_n x_n)$ for all $x = \ltimes_{k=1}^n x_k \in X_2$, $x_k \in \Delta, k = 1, 2, \dots, n$.

EXAMPLE 3.11. Let $M = \delta_2[2, 2, 1, 2]$. Since $\mathrm{col}_1(M) = \delta_2^2$, $S_1^2 \delta_4^1 = \delta_2^1$ and $S_2^2 \delta_4^2 = \delta_2^1$, we can see that $(0, 1)^{\mathrm{T}} \in \{\mathrm{col}_1(M_1), \mathrm{col}_1(M_2)\}$.

Now we calculate

$$\begin{bmatrix} S_1^n \\ S_2^n \\ \vdots \\ S_n^n \end{bmatrix} \begin{bmatrix} \delta_{2^n}^{q_1} \ \delta_{2^n}^{q_2} \cdots \delta_{2^n}^{q_t} \end{bmatrix} = \begin{bmatrix} \delta_2^{\gamma_1^1} \ \delta_2^{\gamma_2^1} \cdots \delta_2^{\gamma_t^1} \\ \delta_2^{\gamma_1^2} \ \delta_2^{\gamma_2^2} \cdots \delta_2^{\gamma_t^2} \\ \vdots \quad \vdots \qquad \vdots \\ \delta_2^{\gamma_1^n} \ \delta_2^{\gamma_2^n} \cdots \delta_2^{\gamma_t^n} \end{bmatrix} \tag{3.3}$$

and denote

$$N = \begin{bmatrix} \gamma_1^1 \ \gamma_2^1 \cdots \gamma_t^1 \\ \gamma_1^2 \ \gamma_2^2 \cdots \gamma_t^2 \\ \vdots \quad \vdots \qquad \vdots \\ \gamma_1^n \ \gamma_2^n \cdots \gamma_t^n \end{bmatrix}.$$

Combining these with Proposition 3.10, we obtain the following result.

PROPOSITION 3.12. *If $Mx = (M_1 x_1) \wedge (M_2 x_2) \wedge \cdots \wedge (M_n x_n)$ for all $x = \ltimes_{i=1}^n x_i$, $x_i \in \Delta, i = 1, 2, \dots, n$, then for any $\mathrm{col}_j(N) = [\gamma_j^1, \gamma_j^2, \dots, \gamma_j^n]^{\mathrm{T}}$, there exists $k \in \{1, 2, \dots, n\}$ such that $\mathrm{col}_{\gamma_j^k}(M_k) = (0, 1)^{\mathrm{T}}$.*

Denote the training example as $T_e \subset \Delta_{2^n}$. For any $x \in T_e$, let y_x be the target function value of x in vector form, where the target function means the Boolean function representing the target concept.

Now we introduce an element $\oslash$, called the *null element*.

DEFINITION 3.13. For a set A, define the nominal set of A as $A^{\oslash} = A \cup \{\oslash\}$, where $\oslash$ is the null element.

https://doi.org/10.1017/S1446181116000316 Published online by Cambridge University Press

Assume that no operation is defined between $\oslash$ and other elements in a nominal set. Let

$$P = \left(\{1,2\} \setminus \overline{C}_1\right)^{\oslash} \times \left(\{1,2\} \setminus \overline{C}_2\right)^{\oslash} \times \cdots \times \left(\{1,2\} \setminus \overline{C}_n\right)^{\oslash}. \tag{3.4}$$

Define the operator "$\leftrightarrow$" as

$$x \leftrightarrow y = \begin{cases} 1, & x = y, \\ 0, & x \neq y, \end{cases}$$

where $x, y \in \mathbb{R}$. For two matrices $A = (a_{ij})$, $B = (b_{ij})$ of the same dimensions, let matrix $A \leftrightarrow B = (a_{ij} \leftrightarrow b_{ij})$. Assume that there is at most one $(0, 1)^{\mathrm{T}}$ column in each M_k.

Theorem 3.14. *Equation (3.1) is equivalent to equation (3.2) and there exists $x \in \Delta_{2^n}$ such that $Mx = (1, 0)^{\mathrm{T}}$ if and only if:*

(1) *for all $k \in \{1, 2, \ldots, n\}$, $\overline{C}_k \neq \emptyset$;* (3.5)

(2) *there exists $p \in P$, for all $j \in \{1, 2, \ldots, t\}$, $p \leftrightarrow \mathrm{col}_j(N)$ contains an element 1.* (3.6)

Proof. If (3.6) holds, let the corresponding

$$p = (p_1, p_2, \ldots, p_n)^{\mathrm{T}}.$$

When $p_k \neq \oslash$, set $\mathrm{col}_{p_k}(M_k) = (0, 1)^{\mathrm{T}}$ and $\mathrm{col}_{\{1,2\}\setminus p_k}(M_k) = (1, 0)^{\mathrm{T}}$. And, when $p_k = \oslash$, write $M_k = \delta_2[1, 1]$. Let

$$X_1 = \{x \mid x = \ltimes_{i=1}^n x_i,\ x_i \in \Delta, i = 1, 2, \ldots, n,\ (M_1x_1) \wedge (M_2x_2) \wedge \cdots \wedge (M_nx_n) = (1, 0)^{\mathrm{T}}\}$$

and

$$X_2 = \{x \mid x = \ltimes_{i=1}^n x_i,\ x_i \in \Delta, i = 1, 2, \ldots, n,\ (M_1x_1) \wedge (M_2x_2) \wedge \cdots \wedge (M_nx_n) = (0, 1)^{\mathrm{T}}\}.$$

From the discussion above, we can verify that $Mx = (M_1x_1) \wedge (M_2x_2) \wedge \cdots \wedge (M_nx_n)$ for all $x = \ltimes_{i=1}^n x_i \in X_1$, $x_i \in \Delta, i = 1, 2, \ldots, n$, if $p \in P$. Similarly, $Mx = (M_1x_1) \wedge (M_2x_2) \wedge \cdots \wedge (M_nx_n)$ for all $x = \ltimes_{i=1}^n x_i \in X_2$, $x_i \in \Delta, i = 1, 2, \ldots, n$, if $p \leftrightarrow \mathrm{col}_j(N)$ contains an element 1. Then suppose that (3.5) holds. From Proposition 3.8, note that $C_k = \overline{C}_k$, $k = 1, 2, \ldots, n$. Combining this with Proposition 3.1, we see that there exists $x \in \Delta_{2^n}$ such that $Mx = (1, 0)^{\mathrm{T}}$.

Conversely, suppose that $Mx = (M_1x_1) \wedge (M_2x_2) \wedge \cdots \wedge (M_nx_n)$ for all $x = \ltimes_{i=1}^n x_i$, $x_i \in \Delta, i = 1, 2, \ldots, n$. From Proposition 3.8, it follows that $C_k = \overline{C}_k$, $k = 1, 2, \ldots, n$. Since there exists $x \in \Delta_{2^n}$ such that $Mx = (1, 0)^{\mathrm{T}}$, from Proposition 3.1, we obtain (3.5). Construct $p' = (p'_1, p'_2, \ldots, p'_n)^{\mathrm{T}}$ as follows. If $M_k = \delta_2[1, 1]$, set $p'_k = \oslash$. Otherwise, let p'_k satisfy $\mathrm{col}_{p'_k}(M_k) = (0, 1)^{\mathrm{T}}$. Then we can verify that $p' \in P$. By Proposition 3.12, we also have that $p' \leftrightarrow \mathrm{col}_j(N)$ contains an element 1 for any $j \in \{1, 2, \ldots, t\}$. □

We obtain the version space by using Algorithm 1. The correctness of this algorithm follows from the proof of Theorem 3.14.

https://doi.org/10.1017/S1446181116000316 Published online by Cambridge University Press

Algorithm 1

`Step 1.` Construct M satisfying for all $x \in T_e$, $Mx = y_x$. The columns of M which are not involved are undetermined.

`Step 2.` Suppose that in `step 1`, we determine that $\{i \mid \mathrm{col}_i(M) = (1,0)^{\mathrm{T}}\} = \{i_1, i_2, \ldots, i_s\}$ and $\{i \mid \mathrm{col}_i(M) = (0,1)^{\mathrm{T}}\} = \{q_1, q_2, \ldots, q_t\}$. Then let

$$C_k = \overline{C}_k = \{i \mid i \in \{1, 2\}, \delta_2^i \in \mathrm{col}(S_k^n[\delta_{2^n}^{i_1}, \delta_{2^n}^{i_2}, \ldots, \delta_{2^n}^{i_s}])\}.$$

Calculate P and N as those from (3.3) and (3.4).

`Step 3.` Set $P_{\mathrm{ad}} = \emptyset$. For all the elements $p \in P$, verify whether for all $j \in \{1, 2, \ldots, t\}$, $p \leftrightarrow \mathrm{col}_j(N)$ contains the element 1. If so, add p to P_{ad}.

`Step 4.` We can derive the version space $\{y = (M_1x_1) \wedge (M_2x_2) \wedge \cdots \wedge (M_nx_n) \mid p = (p_1, p_2, \ldots, p_n)^{\mathrm{T}} \in P_{\mathrm{ad}}$. If $p_k \neq \varnothing$, then $\mathrm{col}_{p_k}(M_k) = (0, 1)^{\mathrm{T}}$. Otherwise, $M_k = \delta_2[1, 1]$, $k = 1, 2, \ldots, n\}$.

4. Conclusion

In this paper, an alternative method of concept learning has been established within the new framework of STP. To develop this theory, the core problem is to obtain the necessary and sufficient condition where a function in form (3.1) can be equivalently expressed in form (3.2). Since it is solved, the algorithm for finding the version space naturally evolves as a byproduct.

Here, we give a comparison between our algorithm and an existing method. In the process of the candidate-elimination approach, which is the traditional way to derive the version space, two sets called general boundary and specific boundary need to be maintained. For each element in training examples D, the two sets are changed accordingly. Thus, the iteration times are given by $|D|$. Besides, Haussler [5] concluded that the dimension of the general boundary increases exponentially according to the scale of $|D|$ (see [7] for more details).

Algorithm 1 is required to store a 2×2^n matrix M. In step 2, matrix C_k is computed n times, and each time it involves the product of two matrices whose dimensions are 2×2^n and $2^n \times 2^n$. At most 2^n elements are contained in the set P and the dimension of N is at most $n \times 2^n$. The iteration times in step 3 are $|P|$ and, at a time, the running time is $O(n)$. Therefore, the computation complexity increases exponentially according to the number of attributes. So, whether Algorithm 1 is more efficient than the existing one depends on different situations.

Acknowledgement

This work was partially supported by NNSF of China (61374025).

References

[1] D. Cheng, "Semi-tensor product of matrices and its application to Morgen's problem", *Sci. China Inf. Sci.* **44** (2001) 195–212; doi:10.1007/BF02714570.

[2] D. Cheng and H. Qi, "Controllability and observability of Boolean control networks", *Automatica* **45** (2009) 1659–1667; doi:10.1016/j.automatica.2009.03.006.

[3] D. Cheng and H. Qi, "A linear representation of dynamics of Boolean networks", *IEEE Trans. Automat. Control* **55**(10) (2010) 2251–2258; doi:10.1109/TAC.2010.2043294.

[4] D. Cheng, H. Qi, Z. Li and J. Liu, "Stability and stabilization of Boolean networks", *Internat. J. Robust Nonlinear Control* **21** (2011) 134–156; doi:10.1002/rnc.1581.

[5] D. Haussler, "Quantifying inductive bias: AI learning algorithms and Valiant's learning framework", *Artificial Intelligence* **36** (1988) 177–221; doi:10.1016/0004-3702(88)90002-1.

[6] T. M. Mitchell, *Machine learning* (McGraw-Hill, Boston, MA, 1997); https://www.iitgn.ac.in/sites/default/files/library_files/2016/19032016.pdf.

[7] T. M. Mitchell, "Version spaces: a candidate elimination approach to rule learning", in: *IJCAI'77 Proceedings of the 5th international joint conference on Artificial intelligence - Volume 1* (Morgan Kaufmann Publishers Inc., San Francisco, CA, 1977) 305–310; http://dl.acm.org/citation.cfm?id=1624501.

[8] Y. Zhao, H. Qi and D. Cheng, "Input-state incidence matrix of Boolean control networks and its applications", *Systems Control Lett.* **59** (2010) 767–774; doi:10.1016/j.sysconle.2010.09.002.