

On the Convergence of Tsetlin Machines for the IDENTITY- and NOT Operators

Xuan Zhang, Lei Jiao, *Senior member, IEEE*, Ole-Christoffer Granmo, and Morten Goodwin

Abstract—The Tsetlin Machine (TM) is a recent machine learning algorithm with several distinct properties, such as interpretability, simplicity, and hardware-friendliness. Although numerous empirical evaluations report on its performance, the mathematical analysis of its convergence is still open. In this article, we analyze the convergence of the TM with only one clause involved for classification. More specifically, we examine two basic logical operators, namely, the “IDENTITY”- and “NOT” operators. Our analysis reveals that the TM, with just one clause, can converge correctly to the intended logical operator, learning from training data over an infinite time horizon. Besides, it can capture arbitrarily rare patterns and select the most accurate one when two candidate patterns are incompatible, by configuring a granularity parameter. The analysis of the convergence of the two basic operators lays the foundation for analyzing other logical operators. These analyses altogether, from a mathematical perspective, provide new insights on why TMs have obtained state-of-the-art performance on several pattern recognition problems.

Index Terms—Tsetlin Automata, Propositional Logic, Tsetlin Machine, Convergence Analysis.

1 INTRODUCTION

The Tsetlin Machine (TM) [1] is a novel machine learning mechanism in which groups of Tsetlin Automata (TAs) [2] operate on binary data using propositional logic, to learn patterns from training samples. A game between the TAs organizes the learning in order to optimize pattern classification accuracy. The TM is an offspring of learning automata (LAs) that have been studied for decades and employed in many applications [3], [4], [5]. The state-of-the-art of LA research is surveyed in [6], [7].

The main advantages of TMs are twofold. The first advantage is their transparency, i.e., the ability to lay open the reasoning behind the decision making process. This advantage addresses a critical challenge in Artificial Intelligence (AI) research – lack of interpretability [8]. Deep neural networks, in particular, mainly support *approximate* local interpretability through post-processing. The approximations do not guarantee model fidelity, however, and global interpretability is currently out of reach [9]. TMs, on the other hand, is inherently interpretable, being founded on propositional logic. That is, one can formulate and manipulate patterns using logical operators. As a result, there exist closed formula expressions for both local and global TM interpretation, akin to SHAP [10].

The second advantage of TMs is simplicity and hardware-friendliness [11]. Because each learning component of a TM is a finite-state automaton (i.e., the TA), the TM only needs to maintain a set of integers to track its state. In contrast to the more extensive arithmetic operations required by most AI techniques, a TA learns using increment and decrement operations only. TMs are thus naturally suited for hardware implementation. Indeed, due to the robustness of TA learning and TM pattern representation, TMs have recently been shown to be inherently fault-tolerant, completely masking stuck-at faults [12].

The TM itself can further act as a building block in more advanced architectures. It has for instance been extended to support convolution [13], regression [14], [15], and relational reasoning [16]. Computation-wise, the native parallelism of TMs supports efficient deployment on Compute Unified Device Architecture (CUDA) [17]. The various TM architectures have been employed in several application domains with promising results. One example is natural language processing where the propositional TM formulas capture textual patterns to solve tasks such as text classification [18], [19], [20], word sense disambiguation [21], novelty detection [22], semantic relation analysis [23], question-answering [16], keyword spotting [24], and aspect-based sentiment analysis [25]. These studies, overall, report that TMs, with smaller memory footprint and higher computational efficiency, obtain better or competitive classification accuracy compared to state-of-the-art AI techniques. Further, the applications demonstrate how the TM facilitates interpretation, contrasted against other machine learning tech-

- X. Zhang is with the Centre for Artificial Intelligence Research, University of Agder, Grimstad, Norway, and with the Norwegian Research Centre (NORCE), Grimstad, Norway, and also with the Confront AS, Norway.
- L. Jiao, O.-C. Granmo, and M. Goodwin are with the Centre for Artificial Intelligence Research, University of Agder, and the Department of ICT, University of Agder, Grimstad, Norway.
- Corresponding author: Lei Jiao. E-mail: lei.jiao@uia.no.

niques.

Although results from computational simulations and empirical evaluations have demonstrated the effectiveness of TMs, analyzing the convergence of the TM learning process is yet to be done. We establish the first step in this undertaking in the present article by analyzing the convergence of TMs via rigorous mathematical tools. The proofs employ a quasi-stationary analysis of the system states together with Markov chain analysis. Overall, we aim to analyze four fundamental logical operators: IDENTITY/NOT, AND, OR, and XOR. The analysis of these operators forms the foundation for analyzing more complex propositional expressions. In this article, we study the convergence of the IDENTITY- and NOT operators, while we will address the AND-, OR-, and XOR-operators in separate articles.

The IDENTITY- and NOT operators are unary, operating on one bit of data. We refer to the analysis of these operators as the “one-bit” case. The “two-bit” case deals with the binary operators AND, OR, and XOR. For the one-bit case, we first prove that the TM can converge surely to the correct pattern when the training data is noise-free. Then we generalize the one-bit case by analyzing the effect of noise, establishing how the noise probability of the data and the granularity parameter of the TM govern convergence.

The remainder of the paper is organized as follows. Section 2 briefly reviews the TM and its training process for the one-bit case. In Section 3, we present our analytical procedure for the one-bit case and the main analytical results. We conclude the paper in Section 4.

2 REVIEW OF THE TSETLIN MACHINE

The fundamentals of TMs have been detailed in [1]. To make this article self-contained, in this section, we do a brief review on TMs, which includes the concept of TA, the TM architecture, and the training process of the TM for the one-bit case.

2.1 Tsetlin Automata (TAs)

TAs are the basic components of the TM. A TA is a fixed structure deterministic automaton that performs actions in a stochastic environment. By interacting with the environment, it aims to learn the optimal action, i.e., the one that has the highest probability of eliciting a reward. A complete description of TA can be found in [2], [26]. Figure 1 shows a simple example of a two-action TA with $2N$ memory states, where $N \in [1, +\infty)$. When the system is in state 1 to N , i.e., on the left-hand side of the state-space, the TA chooses Action 1. If the system is in state $N + 1$ to $2N$, i.e., on the right-hand side of the state-space, the action of the TA is Action 2.

In each interaction with the environment, the TA performs one of the available actions. The environment, in turn, responds with either a reward or a

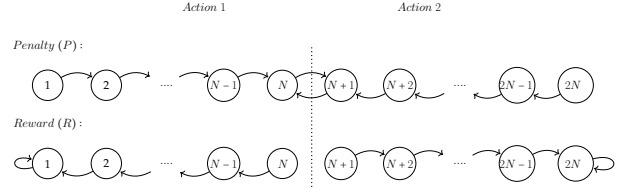


Fig. 1: A two-action Tsetlin Automaton with $2N$ states.

penalty. If the TA receives a penalty, it will move towards the opposite side of the current action. The upper state chain shows this in Figure 1. Conversely, if the TA receives a reward as a response, it will switch to a “deeper” state by moving towards the left or right end of the state chain, depending on whether the current action is Action 1 or Action 2. The lower state chain shows this in Figure 1.

Note that the number of states in a TA can be adjusted. The larger the number, the slower the TA converges, however, the more accurately the TA performs in the stochastic environment. The goal of the TA is to converge to the optimal action by learning. For example, if Action 1 provides a reward with probability 90% and Action 2 has a reward probability of 80%, we expect the TA only to perform Action 1 after the training process.

2.2 Tsetlin Machines (TMs)

Tsetlin Automata Team. A TM consists of m teams of TAs that participate in a decentralized game, designed to solve pattern recognition tasks. The TM trains the TA teams to produce m conjunctive clauses C_j^i , $j = 1, 2, \dots, m$, defined below. These capture the sub-patterns that characterize a particular class i .

The TM takes binary inputs and utilizes propositional logic to represent patterns. In the general case, suppose $X = [x_1, x_2, \dots, x_o]$, with $x_k \in \{0, 1\}$, $k = 1, 2, \dots, o$, is the input of the TM. Then each TA team is formed by o pairs of TAs, each of which takes care of one input value x_k . More specifically, a TA team $\mathcal{G}_j^i = \{\text{TA}_{k'}^{i,j} | 1 \leq k' \leq 2o\}$ consists of $2o$ TAs, where i refers to a specific pattern class, while j refers to a specific clause. The automaton $\text{TA}_{2k-1}^{i,j}$ addresses the input x_k as is, whereas $\text{TA}_{2k}^{i,j}$ addresses the negation of x_k , i.e., $\neg x_k$. Note that the inputs and their negations are jointly referred to as literals. Each TA, in turn, chooses one of two actions. It either “includes” or “excludes” its literal. The actions taken by the TA team members then decide whether the associated literals take part in a conjunction, expressed by the conjunctive clause [1]:

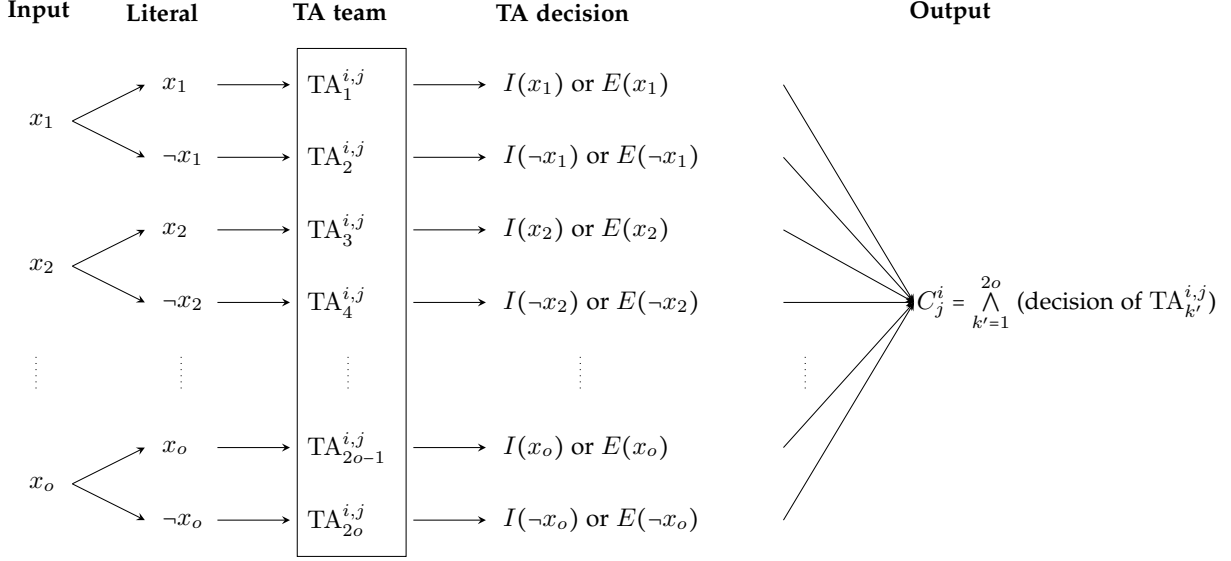


Fig. 2: A TA team G_j^i consisting of $2o$ TAs.

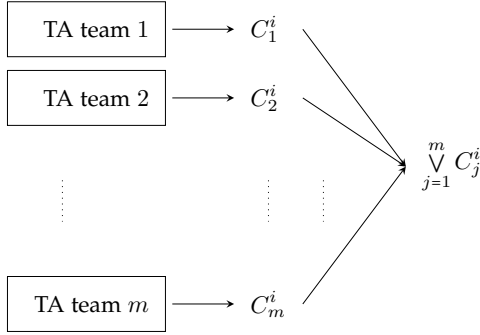


Fig. 3: TM disjunctive normal form architecture.

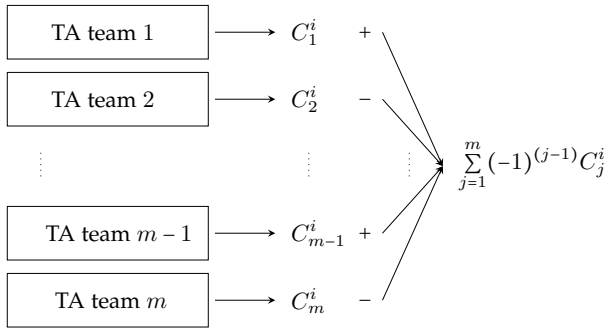


Fig. 4: TM voting architecture.

$$C_j^i(X) = \begin{cases} \left(\bigwedge_{k \in I_j^i} x_k \right) \wedge \left(\bigwedge_{k \in \bar{I}_j^i} \neg x_k \right) \wedge 1 & \text{During training,} \\ \left(\left(\bigwedge_{k \in I_j^i} x_k \right) \wedge \left(\bigwedge_{k \in \bar{I}_j^i} \neg x_k \right) \right) \vee 0 & \text{During testing.} \end{cases} \quad (1)$$

Above, I_j^i and \bar{I}_j^i are the subsets of indexes corresponding to the literals that have been included in the clause. In the subset I_j^i , the included literals are the original input x_k , whereas in \bar{I}_j^i , the included literals are the negated input $\neg x_k$. A clause does not need to include literals, i.e., $I_j^i = \emptyset$ and $\bar{I}_j^i = \emptyset$ are allowed. This happens when all the TAs of the clause choose to exclude their literals. We then define the value of the clause to be $C_j^i(X) = 1$ during the training process and $C_j^i(X) = 0$ during the testing process. Note that an empty clause is usually defined to be 1-valued in propositional logic. However, empirically, it turns out that TMs in general obtain higher test accuracy (on new data) when the empty clauses are 0-valued. In this paper, our theoretical analysis addresses the training process, which means that we operate with the standard definition of empty clauses.

To illustrate the structure of a clause and its relationship to its literals, we show a general clause with $2o$ TAs in Figure 2. For ease of notation, we define the include action as $I(x) = x$ and $I(\neg x) = \neg x$, while we define the exclude action as $E(\cdot) = 1$, the latter meaning that an excluded literal does not contribute to the conjunctive output.

Disjunctive Normal Form (DNF) Architecture. Multiple TA teams are finally assembled into a complete TM. Figure 3 shows a TM composed of m TA teams. The goal of the TM is to learn a pattern representation of a given class i . To this end, the TM decomposes the class into multiple sub-patterns. The outputs of the TA teams, i.e., the conjunctive clauses C_j^i , $1 \leq j \leq m$, are used to represent the sub-patterns of the class. That is, the disjunction of the clause sub-

patterns represents the class:

$$\hat{y}^i = \bigvee_{j=1}^m C_j^i. \quad (2)$$

This is equivalent to saying that if an input “matches” any one of the clauses, i.e., $C_j^i(X) = 1$, for any $j = 1, 2, \dots, m$, then it belongs to class i . Accordingly, the output of the TM is $\hat{y}^i = 1$. Otherwise, the output is $\hat{y}^i = 0$.

Voting Architecture. Besides the DNF architecture, one can organize the TA teams in a “voting” architecture. Figure 4 shows such a TM. In this architecture, the TM assigns a polarity to each TA team. TA teams with odd indexes get positive polarity, and they vote for class i . The remaining TA teams get negative polarity and vote against class i . The voting consists of summing the output of the clauses, according to polarity:

$$f(X) = \left(\sum_{j \in \{1,3,\dots,m-1\}} C_j^i(X) \right) - \left(\sum_{j \in \{2,4,\dots,m\}} C_j^i(X) \right). \quad (3)$$

The output of the TM, in turn, is decided by the unit step function:

$$\hat{y}^i = \begin{cases} 0 & \text{for } f(X) < 0 \\ 1 & \text{for } f(X) \geq 0 \end{cases}. \quad (4)$$

Compared with the DNF architecture, the voting architecture turns out to be more robust to noise and thus more suitable for solving real-world problems [1].

Remark. One drawback of employing the decentralized TA teams for learning is that sub-patterns, in effect, must compete for clauses. It thus can happen that all the available clauses converge to one or some of the sub-patterns. The TM would thus fail to capture the other sub-patterns of the class. To address this drawback, the author of [1] introduced voting summation target T to bound the maximum number of clauses assigned to each sub-pattern. By doing so, the TM distributes the clauses (i.e., its pattern representation capacity) among as many sub-patterns as possible. The one-bit case does not require such a summation target. We, therefore, do not go into the details of this mechanism in the present paper.

2.3 The TM Game for Learning Patterns

2.3.1 The TM Game

The TM training process is built on interconnecting the TAs as players in a decentralized game, with the game being driven by the training data. The training data $(X = [x_1, x_2, \dots, x_o], y^i)$ is obtained from a dataset \mathcal{S} , distributed according to the probability distribution $P(X, y^i)$. In the game, each TA is guided by so-called Type I Feedback and Type II Feedback, as shown in Table 1 and Table 2, respectively. Type I Feedback is activated when the training sample

fed into the system has a positive label, i.e., $y^i = 1$, meaning that the sample belongs to class i . When the training sample is labeled as not belonging to class i , i.e., $y^i = 0$, Type II Feedback is triggered for generating responses. These two types of feedback are designed to reinforce true positive output, i.e., $(\hat{y}^i = 1, y^i = 1)$. Simultaneously, they suppress false positive output, i.e., $(\hat{y}^i = 1, y^i = 0)$, and false negative output, i.e., $(\hat{y}^i = 0, y^i = 1)$.

As shown in Table 1, Type I Feedback encourages a TA to include its literal into its clause whenever the literal has value 1 and the clause evaluates to 1. This is achieved by assigning a relatively large probability $\frac{s-1}{s}$ for rewarding the inclusion of 1-valued literals and for penalizing their exclusion, where $s \geq 1$ is the so called granularity parameter. Conversely, if the clause falsely evaluates to 0, Type I Feedback leads the TA to exclude its literal by rewarding exclude and penalizing include, now with probability $\frac{1}{s}$. Accordingly, excluding a sufficient number of literals makes the clause eventually evaluate to 1. When the forces of including literals and excluding literals are in balance, the TM reaches a Nash equilibrium. In this manner, Type I Feedback reinforces the true positive output, and suppresses the false negative output.

Type II Feedback, shown in Table 2, penalizes excluding literals of value 0 with probability 1.0, however, only when the clause falsely evaluates to 1. This is to encourage including literals of value 0 into the clause, which enhances the ability of the clause to discriminate between different patterns. For the other configurations, the feedback is always inaction (with probabilities 1.0). In particular, if the clause evaluates to 0, Type II Feedback provides no further stimuli for including literals of value 0. Instead, reinforcing include further is left to the Type I Feedback. The purpose of this is to avoid local optima. In this way, the Type II Feedback suppresses false positive output.

Note that the formation of patterns is founded on frequent pattern mining. That is, the parameter s above controls the granularity of the clauses. A larger s allows more literals to be included in each clause, making the corresponding sub-patterns more fine-grained.

A detailed analysis of the two feedback tables can be found in [1]. We now introduce the training process in detail, focusing specifically on the one-bit case.

2.3.2 The training process in the one-bit case

In the one-bit case, the input is reduced to a single variable $X = x$, which takes one of two values, $x \in \{0, 1\}$. Each example x in the training set \mathcal{S} , in turn, has two types of labels, i.e., $y = 0$ or $y = 1$. This results in four input-label configurations: $(x, y) = (0, 0)$, $(x, y) = (0, 1)$, $(x, y) = (1, 0)$, and $(x, y) = (1, 1)$. Based on the probabilities that these input-label pairs appear in the training data set, we define the following two cases:

Value of the clause $C_j^i(X)$		1		0	
Value of the Literal $x_k/\neg x_k$		1	0	1	0
Include Literal	$P(\text{Reward})$	$\frac{s-1}{s}$	NA	0	0
	$P(\text{Inaction})$	$\frac{1}{s}$	NA	$\frac{s-1}{s}$	$\frac{s-1}{s}$
	$P(\text{Penalty})$	0	NA	$\frac{1}{s}$	$\frac{1}{s}$
Exclude Literal	$P(\text{Reward})$	0	$\frac{1}{s}$	$\frac{1}{s}$	$\frac{1}{s}$
	$P(\text{Inaction})$	$\frac{1}{s}$	$\frac{s-1}{s}$	$\frac{s-1}{s}$	$\frac{s-1}{s}$
	$P(\text{Penalty})$	$\frac{s-1}{s}$	0	0	0

TABLE 1: Type I Feedback — Feedback upon receiving a sample with label $y = 1$, for a single TA to decide whether to Include or Exclude a given literal $x_k/\neg x_k$ into C_j^i . NA means not applicable.

Value of the clause $C_j^i(X)$		1		0	
Value of the Literal $x_k/\neg x_k$		1	0	1	0
Include Literal	$P(\text{Reward})$	0	NA	0	0
	$P(\text{Inaction})$	1.0	NA	1.0	1.0
	$P(\text{Penalty})$	0	NA	0	0
Exclude Literal	$P(\text{Reward})$	0	0	0	0
	$P(\text{Inaction})$	1.0	0	1.0	1.0
	$P(\text{Penalty})$	0	1.0	0	0

TABLE 2: Type II Feedback — Feedback upon receiving a sample with label $y = 0$, for a single TA to decide whether to Include or Exclude a given literal $x_k/\neg x_k$ into C_j^i . NA means not applicable.

- **Noise-free case:**

- $P(y = 1|x = 1) = 1$,
- $P(y = 0|x = 1) = 0$,
- $P(y = 1|x = 0) = 0$,
- $P(y = 0|x = 0) = 1$.

- **Noisy case, where $a, b \in (0, 1)$:**

- $P(y = 1|x = 1) = a$,
- $P(y = 0|x = 1) = 1 - a$,
- $P(y = 1|x = 0) = b$,
- $P(y = 0|x = 0) = 1 - b$.

For the sake of simplicity, and without loss of generality, we present the training process based on the noise-free case. In the noise-free case, the output $y = 1$ is always associated with the input value $x = 1$. That is, only one sub-pattern is associated with output $y = 1$. We thus set the number of clauses $m = 1$, meaning that the resulting TM only consists of one TA team. Accordingly, since we only have one input variable, x , the TA team contains two TAs. We thus have simplified the entire system into a game between two TAs.

We further define the two operators we are targeting in this paper, IDENTITY and NOT, as:

$$\begin{aligned} \text{IDENTITY}(x) &= x, \\ \text{NOT}(x) &= 1 - x. \end{aligned}$$

Accordingly, the above noise-free case covers the IDENTITY-operator, which can be converted into the NOT-operator by inverting the output y . We now describe the training process for the noise-free case step-by-step:

- 1) We initialize the TAs by giving each of them a random state among the states associated with the exclude action.
- 2) We obtain a new training sample (x, y) and evaluate our single clause C according to Eq. (1). Depending on the actions of the two TAs, we get the following four possible clause configurations:

$$C(x) = \begin{cases} I(x) \wedge I(\neg x) = 0 & (5a) \\ I(x) \wedge E(\neg x) = I(x) = x & (5b) \\ E(x) \wedge I(\neg x) = I(\neg x) = \neg x & (5c) \\ E(x) \wedge E(\neg x) = 1. & (5d) \end{cases}$$

- 3) Based on the label y , the clause value $C(x)$, and the value of each individual literal x or $\neg x$, we update the states of the TAs according to Table 1 and Table 2. The details on how the states of the TAs are updated will be elaborated in Section 3.1.
- 4) Repeat from Step 2 until a given stopping criteria is met.

Note that the training process defined here does not involve any “voting”, this is because the targeted TM only involves a single clause. This simplification will make it easier for us to capture the core of the learning mechanism underlying a TM, i.e., how the feedback tables are able to guide the TAs to converge to the expected pattern.

3 ANALYTICAL PROCEDURE AND MAIN RESULTS FOR ONE-BIT CASE

In this section, we analyze the convergence of the TM in the one-bit case. The target of the analysis is to decide whether the clause C converges to the correct relation, reflected by the training samples. This is equivalent to saying that we are to find out whether the clause converges to $C = x$ if and only if $P(y = 1|x = 1) > P(y = 0|x = 1)$, or the clause converges to $C = \neg x$ if and only if $P(y = 1|x = 0) > P(y = 0|x = 0)$.

3.1 The noise-free case

Recall that the samples are distributed according to the following probabilities for the noise-free case: $P(y = 1|x = 1) = 1$ which implies $P(y = 0|x = 1) = 0$, and $P(y = 0|x = 0) = 1$ which implies $P(y = 1|x = 0) = 0$. There are hence two types of samples, i.e., $(x, y) = (1, 1)$ and $(x, y) = (0, 0)$, and it is the IDENTITY relation that captures this pattern. We further assume that the input $x = 1$ appears in the training data set with probability $P(x = 1) = c$, which means that $x = 0$ appears with probability $P(x = 0) = 1 - c$. We are now ready to prove that the TM can converge to the IDENTITY relation after training.

Theorem 1. *In the noise-free case where $P(y = 1|x = 1) = 1$ and $P(y = 0|x = 0) = 1$, the TM converges surely to the IDENTITY relation in an infinite time horizon, i.e.,*

$$C(x) \rightarrow I(x) \wedge E(\neg x) = x, \text{ when } c \in (0, 1]. \quad (6)$$

Proof: The TM studied in this article consists of two TAs: TA_1 for literal x and TA_2 for literal $\neg x$. There are thus four possible action configurations: $((E(x), E(\neg x)), (E(x), I(\neg x)), (I(x), E(\neg x)),$ or $(I(x), I(\neg x))$. By combining these four configurations with the two types of training samples in the noise-free case, we get altogether eight scenarios. Table 5 displays the state transitions of TA_1 and TA_2 for each of these scenarios¹. The state transitions are generated from the Type I and Type II Feedback tables, showing how TA_1 and TA_2 move from one state to another. We use Scenario 2 and Scenario 5 to exemplify how the state transitions are generated.

In Scenario 2, we assume that the current states of the TAs are $S_1, S_2 \in (N + 1, 2N)$. This means that both TAs select the exclude action, i.e., TA_1 selects action $E(x)$ and TA_2 selects action $E(\neg x)$. According to Eq. (5d), the resulting clause becomes $C(x) = E(x) \wedge E(\neg x) = 1$. When the sample $(0, 0)$ is fed to the system, we have $C(x) = 1$. Label $y = 0$ triggers Type II Feedback. We thus refer to Table 2 for the reward/inaction/penalty probabilities:

- We first consider TA_1 . As seen in the column where $C = 1$ and $x = 0$, TA_1 receives a penalty with probability 1 for choosing the action “Exclude Literal”. This means that TA_1 changes state to $S_1 - 1$. Accordingly, the state change weakens action “Exclude Literal” by moving towards the “Include Literal”-side of the state space. As the input $x = 0$ appears with probability $1 - c$, the effective penalty probability is $1 - c$ in the state transition diagram.
- We now examine feedback for TA_2 . The column where $C = 1$ and $\neg x = 1$ shows that TA_2 receives inaction feedback with probability 1 for choosing the action “Exclude Literal”. The effective inaction probability is thus $1 - c$. For the sake of simplicity, all the self loops for inaction probabilities are omitted in all the state transition diagrams, which makes plotting the state transition for TA_2 in this scenario trivial, as the state of TA_2 simply remains unchanged.

In Scenario 5, we assume the current state of TA_1 is $S_1 \in (1, N)$, which makes the TA select action $I(x)$. We further assume that TA_2 is in State $S_2 \in (N + 1, 2N)$, producing action $E(\neg x)$. Then, when the sample $(1, 1)$ is fed to the system, we evaluate the clause according to Eq. (5b). We thus get $C(x) = I(x) \wedge E(\neg x) = I(x) \wedge 1 = x = 1$. Since label $y = 1$ triggers Type I Feedback, we refer to Table 1 for the reward/inaction/penalty probabilities:

- We first consider TA_1 . According to the column where $C = 1$ and $x = 1$, TA_1 will receive a reward with probability $\frac{s-1}{s}$ for choosing the

action “Include Literal”. This means that TA_1 will change state to $S_1 + 1$ with probability $\frac{s-1}{s}$, reinforcing the “Include Literal” action. Because $P(x = 1) = c$, the effective reward probability is $\frac{(s-1)c}{s}$, as shown by the state transition diagram. Conversely, TA_1 does not change state with probability $\frac{1}{s} \times c = \frac{c}{s}$, according to the inaction probability for “Include Literal”.

- We now examine feedback for TA_2 . The column where $C = 1$ and $\neg x = 0$ shows that TA_2 will receive a reward with probability $\frac{1}{s}$ for choosing the action “Exclude Literal”. This means that TA_2 changes state to $S_2 + 1$ with probability $\frac{1}{s}$, reinforcing the “Exclude Literal” action. Again, by multiplying the probability c of observing $x = 1$, the effective reward probability becomes $\frac{c}{s}$. The state of TA_2 will thus remain unchanged with probability $\frac{(s-1)c}{s}$ according to the inaction probability for “Exclude Literal”.

For the other scenarios, the state transitions for TA_1 and TA_2 can be generated in the same manner. Furthermore, based on Table 5, the state transitions of TA_1 and TA_2 can be combined to Markov chains, as shown in Figure 5. We analyze the convergence of the TM by investigating the Markov chains.

- We first consider $0 < c \leq 1$, i.e., the case with two types of samples $(1, 1)$ and $(0, 0)$, or only one type, $(1, 1)$. The analysis can then be conducted in two steps:
 - 1) As $s > 1$ by definition, we have $\frac{c}{s} > 0$. The Markov chains for TA_2 shows that no matter which action TA_1 takes, TA_2 moves and converges to $E(\neg x)$.
 - 2) When TA_2 converges to $E(\neg x)$, Chain 1 shows that TA_1 converges to $I(x)$.

We thus can confirm that when $0 < c \leq 1$, $C(x) \rightarrow I(x) \wedge E(\neg x) = x$, reflecting the IDENTITY relation.

- The remaining case is $c = 0$, i.e., when there is only one type of samples $(0, 0)$. Again, the analysis consists of two steps:
 - 1) The Markov chain for TA_2 shows that its state does not change, meaning the final state of TA_2 depends only on its initial state. TA_2 can thus be choosing either $I(\neg x)$ or $E(\neg x)$.
 - 2) If $TA_2 = E(\neg x)$, Chain 1 shows that TA_1 converges to $I(x)$ with probability 1, and we get $C \rightarrow I(x) \wedge E(\neg x) = x$. However, if $TA_2 = I(\neg x)$, then Chain 2 shows that the decision of TA_1 depends fully on its initial state, which can be either $I(x)$ or $E(x)$.

1. We have simplified all the state transition chains in the paper to a chain depth of 2. However, our analysis addresses the general depth N .

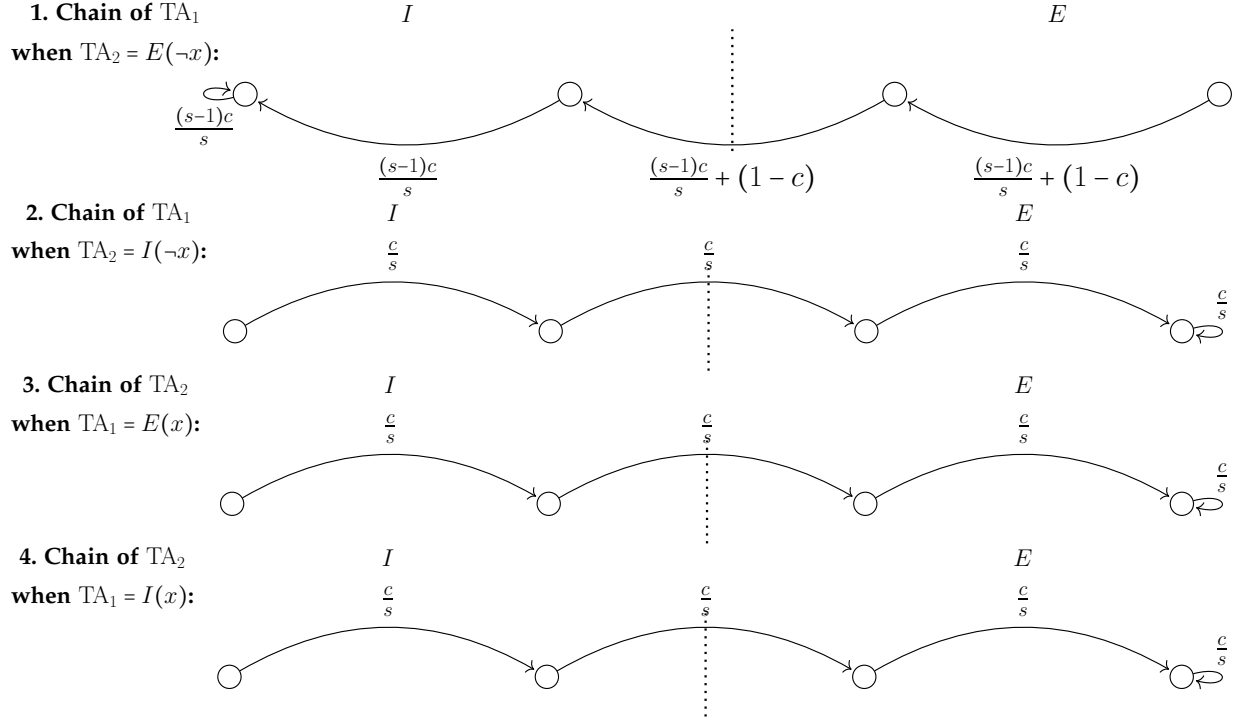


Fig. 5: Markov chains in the noise-free case, when $P(y = 1|X = 1) = 1$, $P(y = 0|X = 0) = 1$, and $P(X = 1) = c$, $c \in (0, 1)$.

In other words, when $c = 0$, $C(x)$ does not converge consistently to $I(x) \wedge E(-x) = x$. This is simply because there are no samples where $y = 1$ to learn from.

In summary, as long as there are samples with label $y = 1$, $C(x)$ will converge to $I(x) \wedge E(-x) = x$, reflecting the IDENTITY relation in pattern $y = 1$. Theorem 1 is thus proven. ■

Remark 1. Theorem 1 does not have any requirements on the depths of the TAs. The TAs can learn the correct pattern even if $N = 1$.

3.2 The noisy case

In the case with noise, we assume the samples are distributed according to the probabilities: $P(y = 1|x = 1) = a$, which implies $P(y = 0|x = 1) = 1 - a$, and $P(y = 1|x = 0) = b$, which implies $P(y = 0|x = 0) = 1 - b$. There are hence four types of samples, i.e., $(1, 1)$, $(1, 0)$, and $(0, 0)$, $(0, 1)$. These provide two types of candidate relations for predicting y , either IDENTITY or NOT. We are to find out whether the TM can converge to the correct relation over an infinite training horizon.

We start with the static state analysis of the Markov chain shown in Figure 6. The Markov chain is indeed a two-action TA that has $2N$ -state, with α , β , and γ being the transition probabilities between states. Note that we have left out the self-loops of the Markov chain in the diagram. We are to derive the

conditions for the TA to converge to one of the actions with probability 1.

Lemma 1. When $N \rightarrow \infty$, action "Include" will be selected with probability 1 if $\alpha < \min(\beta, \gamma)$. Similarly, action "Exclude" will be selected with probability 1 if $\alpha > \max(\beta, \gamma)$.

Proof: Denoting the static state probability of State i as π_i , $i = 0, 1, \dots, 2N$, the balance equations of the Markov chain are:

$$\begin{aligned}
 \alpha\pi_0 &= \beta\pi_1 \implies \pi_1 = \frac{\alpha}{\beta}\pi_0, \\
 \alpha\pi_1 &= \beta\pi_2 \implies \pi_2 = \frac{\alpha}{\beta}\pi_1 = \left(\frac{\alpha}{\beta}\right)^2 \pi_0, \\
 &\dots, \\
 \alpha\pi_{N-1} &= \beta\pi_N \implies \pi_N = \frac{\alpha}{\beta}\pi_{N-1} = \left(\frac{\alpha}{\beta}\right)^N \pi_0, \\
 \alpha\pi_N &= \gamma\pi_{N+1} \implies \pi_{N+1} = \frac{\alpha}{\gamma}\pi_N = \frac{\alpha}{\gamma}\left(\frac{\alpha}{\beta}\right)^N \pi_0, \\
 \alpha\pi_{N+1} &= \gamma\pi_{N+2} \implies \pi_{N+2} = \frac{\alpha}{\gamma}\pi_{N+1} = \left(\frac{\alpha}{\gamma}\right)^2 \left(\frac{\alpha}{\beta}\right)^N \pi_0, \\
 &\dots, \\
 \alpha\pi_{2N-1} &= \gamma\pi_{2N} \implies \pi_{2N} = \frac{\alpha}{\gamma}\pi_{2N-1} = \left(\frac{\alpha}{\gamma}\right)^N \left(\frac{\alpha}{\beta}\right)^N \pi_0.
 \end{aligned} \tag{7}$$

As $(\pi_0 + \pi_1 + \pi_2 + \dots + \pi_N + \pi_{N+1} + \dots + \pi_{2N}) = 1$, we

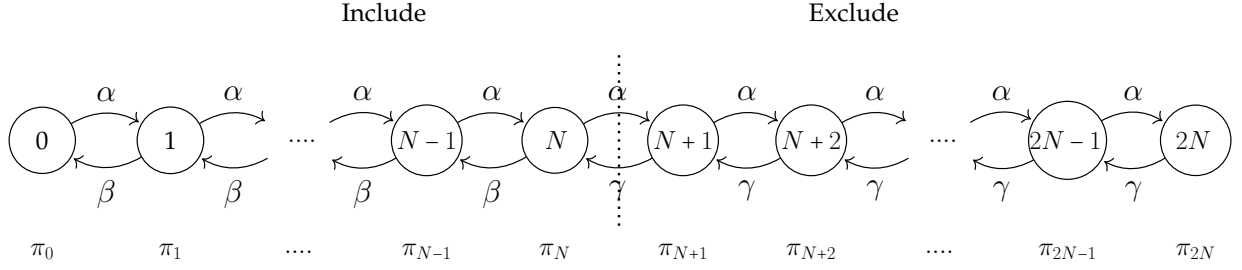


Fig. 6: A Markov chain for Lemma 1.

have:

$$\begin{aligned} & \pi_0 \left(1 + \frac{\alpha}{\beta} + \left(\frac{\alpha}{\beta}\right)^2 + \dots + \left(\frac{\alpha}{\beta}\right)^N \right) + \\ & \pi_0 \left(\frac{\alpha}{\beta}\right)^N \left(\left(\frac{\alpha}{\gamma}\right) + \left(\frac{\alpha}{\gamma}\right)^2 + \dots + \left(\frac{\alpha}{\gamma}\right)^N \right) = 1 \\ \implies & \pi_0 \left(\frac{1 - \left(\frac{\alpha}{\beta}\right)^{N+1}}{1 - \frac{\alpha}{\beta}} + \left(\frac{\alpha}{\beta}\right)^N \left(\frac{1 - \left(\frac{\alpha}{\gamma}\right)^{N+1}}{1 - \frac{\alpha}{\gamma}} \right) \right) = 1. \quad (8) \end{aligned}$$

If $\alpha < \min(\beta, \gamma)$, then when $N \rightarrow \infty$, Eq. (8) is $\pi_0 \left(\frac{1}{1 - \frac{\alpha}{\beta}} \right) = 1$, which implies $\pi_0 = 1 - \frac{\alpha}{\beta}$. The probability that the state of the Markov chain stays on the left-hand side, can thus be calculated as $\lim_{N \rightarrow \infty} (\pi_0 + \pi_1 + \dots + \pi_N) = \lim_{N \rightarrow \infty} \pi_0 \left(1 + \frac{\alpha}{\beta} + \left(\frac{\alpha}{\beta}\right)^2 + \dots + \left(\frac{\alpha}{\beta}\right)^N \right) = \lim_{N \rightarrow \infty} \pi_0 \frac{1 - \left(\frac{\alpha}{\beta}\right)^{N+1}}{1 - \frac{\alpha}{\beta}} = \pi_0 \left(\frac{1}{1 - \frac{\alpha}{\beta}} \right) = 1$. Accordingly, we have proved that action ‘‘Include’’ will be selected with probability 1 if $\alpha < \min\{\beta, \gamma\}$ when $N \rightarrow \infty$.

The above procedure can be modified to prove that action ‘‘Exclude’’ will be selected with probability 1 if $\alpha > \max\{\beta, \gamma\}$ when $N \rightarrow \infty$. ■

We again use the Markov chain to analyze the convergence of the TM. Following the same process as in the noise-free case, we plot the state transitions for both TAs, per unique scenario. In the noisy case, there are in total 16 scenarios. Table 6 and Table 7 list the state transitions for TA_1 and TA_2 in each of these scenarios. By combining the listed transitions, we get the Markov chains for TA_1 and TA_2 , as shown in Figure 7.

We examine TA_1 first:

- 1) When $TA_2 = E(-x)$, the state transition of TA_1 is shown by the first Markov chain in Figure 7.

- a) In order for $TA_1 \rightarrow E(x)$, by Lemma 1, we must have:

$$\begin{aligned} & \frac{b(1-c)}{s} > \\ & \max \left(\frac{(s-1)ac}{s}, \frac{(s-1)ac}{s} + (1-b)(1-c) \right) \\ \implies & \frac{b(1-c)}{s} > \frac{(s-1)ac}{s} + (1-b)(1-c) \\ \implies & s < \frac{ac + b(1-c)}{ac + (1-b)(1-c)} \xrightarrow{\text{define to be}} s_1 \quad (9) \end{aligned}$$

- b) In order for $TA_1 \rightarrow I(x)$, by Lemma 1, we need to have:

$$\begin{aligned} & \frac{b(1-c)}{s} < \\ & \min \left(\frac{(s-1)ac}{s}, \frac{(s-1)ac}{s} + (1-b)(1-c) \right) \\ \implies & \frac{b(1-c)}{s} < \frac{(s-1)ac}{s} \\ \implies & s > \frac{ac + b(1-c)}{ac} \xrightarrow{\text{define to be}} s_2 \quad (10) \end{aligned}$$

- 2) When $TA_2 = I(-x)$, the state transition of TA_1 is shown by the second Markov chain in Figure 7.

- a) In order for $TA_1 \rightarrow E(x)$, by Lemma 1, we have:

$$\begin{aligned} & \frac{ac}{s} + \frac{b(1-c)}{s} > (1-b)(1-c) \\ \implies & s < \frac{ac + b(1-c)}{(1-b)(1-c)} \xrightarrow{\text{define to be}} s_3 \quad (11) \end{aligned}$$

- b) TA_1 will never converge to $I(x)$ when $TA_2 = I(-x)$.

We now apply the same analysis on TA_2 :

- 1) When $TA_1 = E(x)$, the state transition of TA_2 is shown by the third Markov chain in Figure 7.

- a) According to Lemma 1, in order for $TA_2 \rightarrow E(-x)$, we must have:

$$\begin{aligned} & \frac{ac}{s} > \\ & \max \left(\frac{(s-1)b(1-c)}{s}, \frac{(s-1)b(1-c)}{s} + (1-a)c \right) \\ \implies & \frac{ac}{s} > \frac{(s-1)b(1-c)}{s} + (1-a)c \\ \implies & s < \frac{ac + b(1-c)}{b(1-c) + (1-a)c} \xrightarrow{\text{define to be}} s_4 \quad (12) \end{aligned}$$

- b) Similarly, for $TA_2 \rightarrow I(-x)$, we have:

$$\begin{aligned} & \frac{ac}{s} < \\ & \min \left(\frac{(s-1)b(1-c)}{s}, \frac{(s-1)b(1-c)}{s} + (1-a)c \right) \\ \implies & \frac{ac}{s} < \frac{(s-1)b(1-c)}{s} \\ \implies & s > \frac{ac + b(1-c)}{b(1-c)} \xrightarrow{\text{define to be}} s_5 \quad (13) \end{aligned}$$

TABLE 3: Conditions of s for TAs to converge.

Actions of TA ₂	For TA ₁ → $E(x)$:	For TA ₁ → $I(x)$:
TA ₂ = $E(\neg x)$	$s < s_1$	$s > s_2$
TA ₂ = $I(\neg x)$	$s < s_3$	$s > s_3$ NA
Actions of TA ₁	For TA ₂ → $E(\neg x)$:	For TA ₂ → $I(\neg x)$:
TA ₁ = $E(x)$	$s < s_4$	$s > s_5$
TA ₁ = $I(x)$	$s < s_6$	$s > s_6$ NA

2) When TA₁ = $I(x)$, the state transition of TA₂ is shown by the last Markov chain in Figure 7.

a) For TA₂ → $E(\neg x)$, by Lemma 1, we have:

$$\frac{ac}{s} + \frac{b(1-c)}{s} > (1-a)c$$

$$\implies s < \frac{ac + b(1-c)}{(1-a)c} \stackrel{\text{define to be}}{\longrightarrow} s_6 \quad (14)$$

b) TA₂ will never converge to $I(\neg x)$ when TA₁ = $I(x)$.

We list here the definition of s_1, s_2, s_3, s_4, s_5 and s_6 for easy reference:

$$s_1 = \frac{ac + b(1-c)}{ac + (1-b)(1-c)} \xrightarrow{\text{when } c=0.5} s_1 = \frac{a+b}{a+1-b},$$

$$s_2 = \frac{ac + b(1-c)}{ac} \xrightarrow{\text{when } c=0.5} s_2 = \frac{a+b}{a},$$

$$s_3 = \frac{ac + b(1-c)}{(1-b)(1-c)} \xrightarrow{\text{when } c=0.5} s_3 = \frac{a+b}{1-b},$$

$$s_4 = \frac{ac + b(1-c)}{b(1-c) + (1-a)c} \xrightarrow{\text{when } c=0.5} s_4 = \frac{a+b}{b+1-a},$$

$$s_5 = \frac{ac + b(1-c)}{b(1-c)} \xrightarrow{\text{when } c=0.5} s_5 = \frac{a+b}{b},$$

$$s_6 = \frac{ac + b(1-c)}{(1-a)c} \xrightarrow{\text{when } c=0.5} s_6 = \frac{a+b}{1-a}.$$

The conditions for TA₁ and TA₂ to converge to different actions are summarized in Table 3, from which, we have:

- 1) For (TA₁, TA₂) → ($E(x), E(\neg x)$), we need $s < \min(s_1, s_4)$. As $s \geq 1$ by definition, we have $\min(s_1, s_4) > 1$ that implies $a > 0.5, b > 0.5$.
- 2) For (TA₁, TA₂) → ($I(x), E(\neg x)$), we need $s_2 < s < s_6$. In this case, s_6 must be greater than s_2 , which implies $a > 0.5$.
- 3) For (TA₁, TA₂) → ($E(x), I(\neg x)$), we need $s_5 < s < s_3$, which implies $s_5 < s_3$ that leads to $b > 0.5$.

Three convergence possibilities of the TM are summarized in Table 4. This table provides important insights, based on which we analyze how a, b, c , and s affect the convergence of the TM. We organize our analysis according to the following two cases, unbiased and biased training data.

Unbiased data. We first simplify the analysis by rendering $c = 0.5$. This means training samples are unbiased. The conditions for s are listed in the third column in Table 4.

1) When $a < 0.5, b > 0.5$, the NOT relation is dominant in the training data.

In this case, $s_4 = \min(s_1, s_4) < 1, s_6 < s_2$, which means the intervals $(1, \min(s_1, s_4))$ and (s_2, s_6) do not exist. Therefore, only when $s \in (s_5, s_3)$, the TM can converge to ($E(x), I(\neg x)$), reflecting the NOT relation.

2) When $a > 0.5, b < 0.5$, IDENTITY is the dominant relation in the training data.

In this case, $s_1 = \min(s_1, s_4) < 1, s_3 < s_5$, the TM converges only when $s \in (s_2, s_6)$, to ($I(x), E(\neg x)$), i.e., the IDENTITY relation.

3) When $a > 0.5, b > 0.5$, both IDENTITY and NOT are dominating in the training data. Figure 8 shows the positions of s_1, \dots, s_6 when $a > b > 0.5$ and $b > a > 0.5$ respectively.

a) When $a > b > 0.5$, we have $s_4 < s_5 < s_3$ and $s_4 < s_6$, implying two possible positions of s_4 that are visualized by item 1 in Fig. 8. Similarly, when $b > a > 0.5$, we have $s_1 < s_2 < s_6$ and $s_1 < s_3$. Item 2 in Fig. 8 depicts the possible positions of s_1 . Indeed, Fig. 8 shows that s_1 and s_4 cannot be greater than $\min(s_2, s_3, s_5, s_6)$ at the same time, i.e., $\min(s_1, s_4) < \min(s_2, s_3, s_5, s_6)$. Therefore, when $s \in (1, \min(s_1, s_4))$, the TM converges only to ($E(x), E(\neg x)$). In other words, the s satisfying the condition for converging to ($E(x), E(\neg x)$) will not result in the convergence to ($I(x), E(\neg x)$) or ($E(x), I(\neg x)$). In the classifying process, $C(x) = E(x) \wedge E(\neg x)$ means all inputs, whether $x = 0$ or $x = 1$, can be classified into pattern $y = 1$, reflecting the fact that both IDENTITY and NOT are the sub-patterns associated with pattern $y = 1$ in this case.

b) When $a > b > 0.5$, Figure 8 shows that the interval of (s_2, s_6) contains the interval of (s_5, s_3) . This means:

- When $s \in (s_2, s_5) \cup (s_3, s_6)$, the TM will converge to ($I(x), E(\neg x)$).
- When $s \in (s_5, s_3)$, the TM can converge to either ($I(x), E(\neg x)$) or ($E(x), I(\neg x)$).

We take the interval $s \in (s_2, s_5)$ as an example to show how the TM converges to ($I(x), E(\neg x)$).

When s is bounded by (s_2, s_5) , the transition probabilities in the Markov chains are bounded accordingly, as shown in Figure 9. The ranges of the transition probabilities imply the trend of movement of the TAs. In Chain 1,

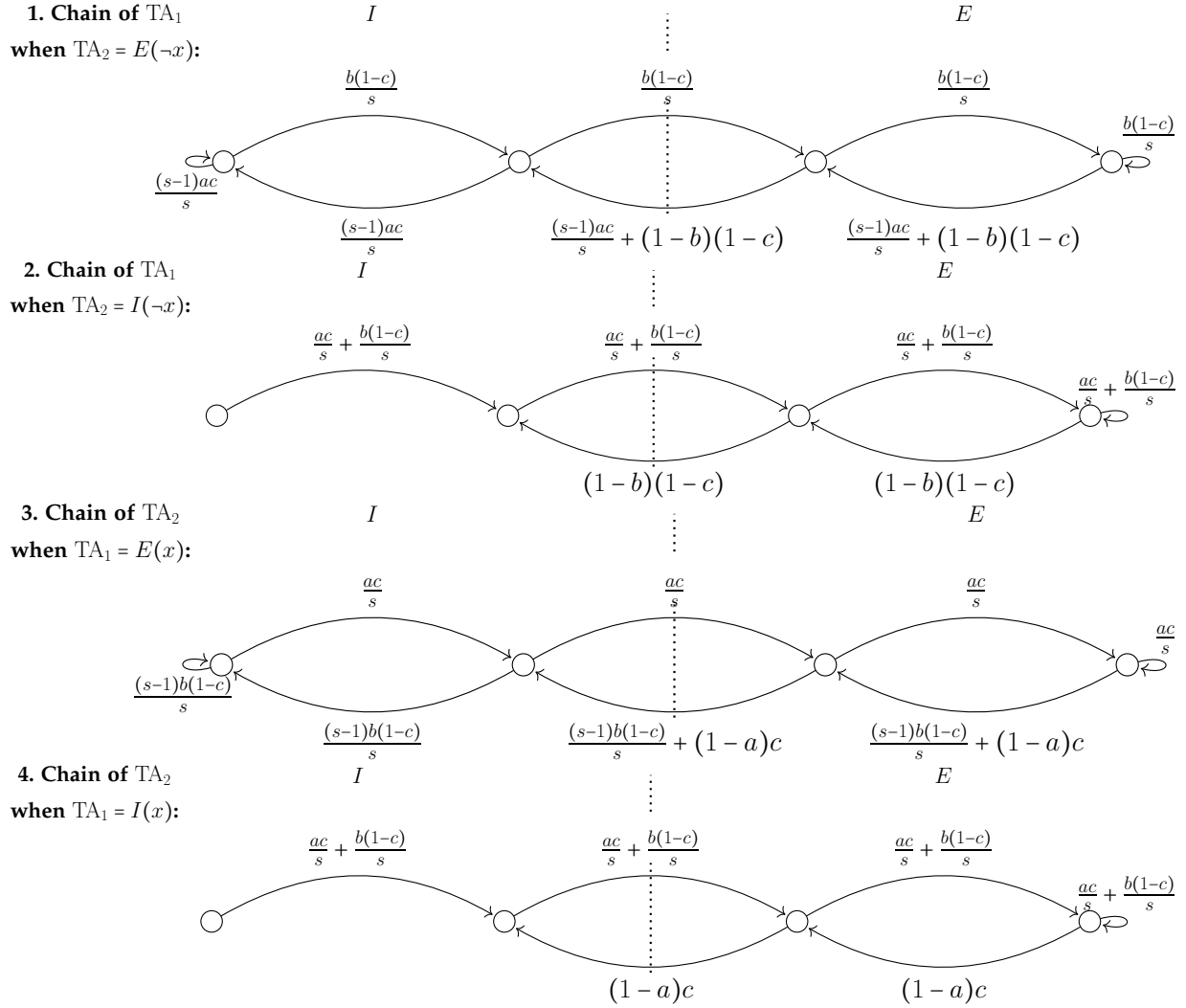


Fig. 7: Markov chains in the noisy case, where $P(y = 1|X = 1) = a$, and $P(y = 1|X = 0) = b$; $P(X = 1) = c$; $a, b, c \in (0, 1)$.

TABLE 4: Three convergence possibilities of the TM.

(TA_1, TA_2)	s in general case	s when $c = 0.5$	a, b
$(TA_1, TA_2) \rightarrow (E(x), E(-x))$	$1 < s < \min(s_1, s_4)$	$1 < s < \min(\frac{a+b}{a+1-b}, \frac{a+b}{b+1-a})$	$a > 0.5, b > 0.5$
$(TA_1, TA_2) \rightarrow (I(x), E(-x))$	$s_2 < s < s_6$	$\frac{a+b}{a} < s < \frac{a+b}{1-a}$	$a > 0.5$
$(TA_1, TA_2) \rightarrow (E(x), I(-x))$	$s_5 < s < s_3$	$\frac{a+b}{b} < s < \frac{a+b}{1-b}$	$b > 0.5$

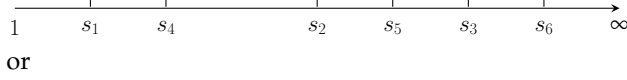
the trend is TA_1 moving towards $I(x)$. In Chain 2 and Chain 4, the trend is that TA_1 and TA_2 both move towards excluding literal. The situation in Chain 3 is different. When TA_2 is on the left-hand side of the chain, the trend is that it moves towards $E(-x)$. However, when TA_2 is on the right-hand side of the chain, the trend depends on the values of a and b . In brief, if $a > \frac{b+1}{2}$, TA_2 tends to move towards $E(-x)$, whereas if $a < \frac{b+1}{2}$, TA_2 tends

to move towards $I(-x)$.

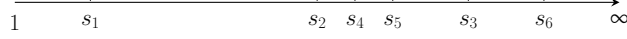
Suppose both the TAs initially are in one of the include literal-states, i.e., $(TA_1, TA_2) = (I(x), I(-x))$. Chain 2 and Chain 4 tell that both TAs tend to move towards $E(\cdot)$.

- If TA_2 reaches $E(-x)$ first, then Chain 1 shows that TA_1 tends to move towards $I(x)$. Once TA_1 reaches $I(x)$, according to Chain 4, TA_2 tends to move towards $E(-x)$. The TM thus con-

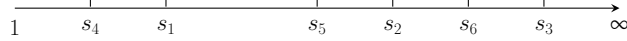
1. When $a > b > 0.5$, the positions can be:



or



2. When $b > a > 0.5$, the positions can be:



or

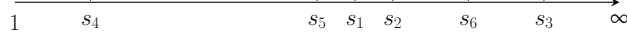


Fig. 8: Positions of $s_1, s_2, s_3, s_4, s_5, s_6$ when $a > b > 0.5$ or $b > a > 0.5$, with $c \in (0, 1)$.

verges to $(I(x), E(\neg x))$.

- If TA_1 reaches $E(x)$ first, then Chain 3 shows that TA_2 tends to first move towards $E(\neg x)$. When TA_2 reaches the node I_N , that is in the middle left of the chain, TA_2 tends to move towards either $I(\neg x)$ or $E(\neg x)$:
 - when $a > \frac{b+1}{2}$, TA_2 continues the trend of moving towards $E(\neg x)$. Once TA_2 reaches the state of $E(\neg x)$, TA_1 tends to move towards $I(x)$ according to Chain 1. This means the TM is moving towards $(I(x), E(\neg x))$, and once $TA_1 = I(x)$ and $TA_2 = E(\neg x)$, the TAs tend to go deeper and deeper towards the end states and the TM converges to $(I(x), E(\neg x))$.
 - when $a < \frac{b+1}{2}$, the probability that TA_2 moves towards $E(\neg x)$ is less than the probability it moves towards $I(\neg x)$. Therefore, TA_2 tends to linger around at the node I_N . When it happens that TA_2 moves towards $E(\neg x)$ and reaches the node that is on the right-hand side of the chain, then according to Chain 1, TA_1 tends to move towards $I(x)$. This means that it is still possible for the TM to move towards $(I(x), E(\neg x))$, but the process that the TM reaches $(I(x), E(\neg x))$ involves more “back and forth” movements, and thus it takes longer time

for the TM to converge. This makes sense, as when $a < \frac{b+1}{2}$, the difference between a and b is smaller than when $a > \frac{b+1}{2}$, it is reasonable that the TM will need more time to converge to the more advantageous relation.

Indeed, depending on the range of s , the TM converges through different paths. One can examine how different choices of s influence the converging process by following the same Markov analysis on the above example.

- c) When $b > a > 0.5$, the situation is completely opposite to the case when $a > b > 0.5$. Figure 8 shows the interval of (s_2, s_6) is contained in (s_5, s_3) .
 - When $s \in (s_5, s_2) \cup (s_6, s_3)$, the TM converges only to $(E(x), I(\neg x))$, i.e., the NOT operator.
 - When $s \in (s_2, s_6)$, the TM converges to either $(I(x), E(\neg x))$ or $(E(x), I(\neg x))$.

When $a > 0.5, b > 0.5$, there are basically two sub-patterns (IDENTITY and NOT) in the pattern $y = 1$. By properly configuring s , a TM with only one clause is able to converge to $(E(x), E(\neg x))$ to classify both sub-patterns into the correct pattern $y = 1$. However, the one-clause TM does not distinguish between sub-patterns. To distinguish between sub-patterns, we need to configure the TM with at least two clauses, each of which being used to represent one sub-pattern. These clauses can be organized in a voting architecture, where the threshold $T = 1$ can be applied to allocate clause resource.

- 4) When $a = b = 0.5$, we have $s_1 = s_4 = 1$, and $s_2 = s_6 = s_5 = s_3 = 2$. No s -value will cause the TM to converge. This makes sense as in this case, the training samples have the greatest uncertainty, there are basically no pattern present for the TM to learn.
- 5) When $a < 0.5, b < 0.5$, we have $s_1 < 1, s_4 < 1, s_6 < s_2$, and $s_3 < s_5$. No s -value meets the convergence conditions listed in Table 4. The TM accordingly does not learn, because it is designed to learn the pattern with label $y = 1$. For this scenario, $y = 1$ is not the dominant pattern in the training set because both a and b are less than 0.5. Indeed, the dominant pattern becomes $y = 0$. So if we swap Type I and Type II Feedback for the TM, it is able to learn the pattern $y = 0$, with the analysis being exactly the same as that on the case when $a > 0.5, b > 0.5$ for pattern $y = 1$.

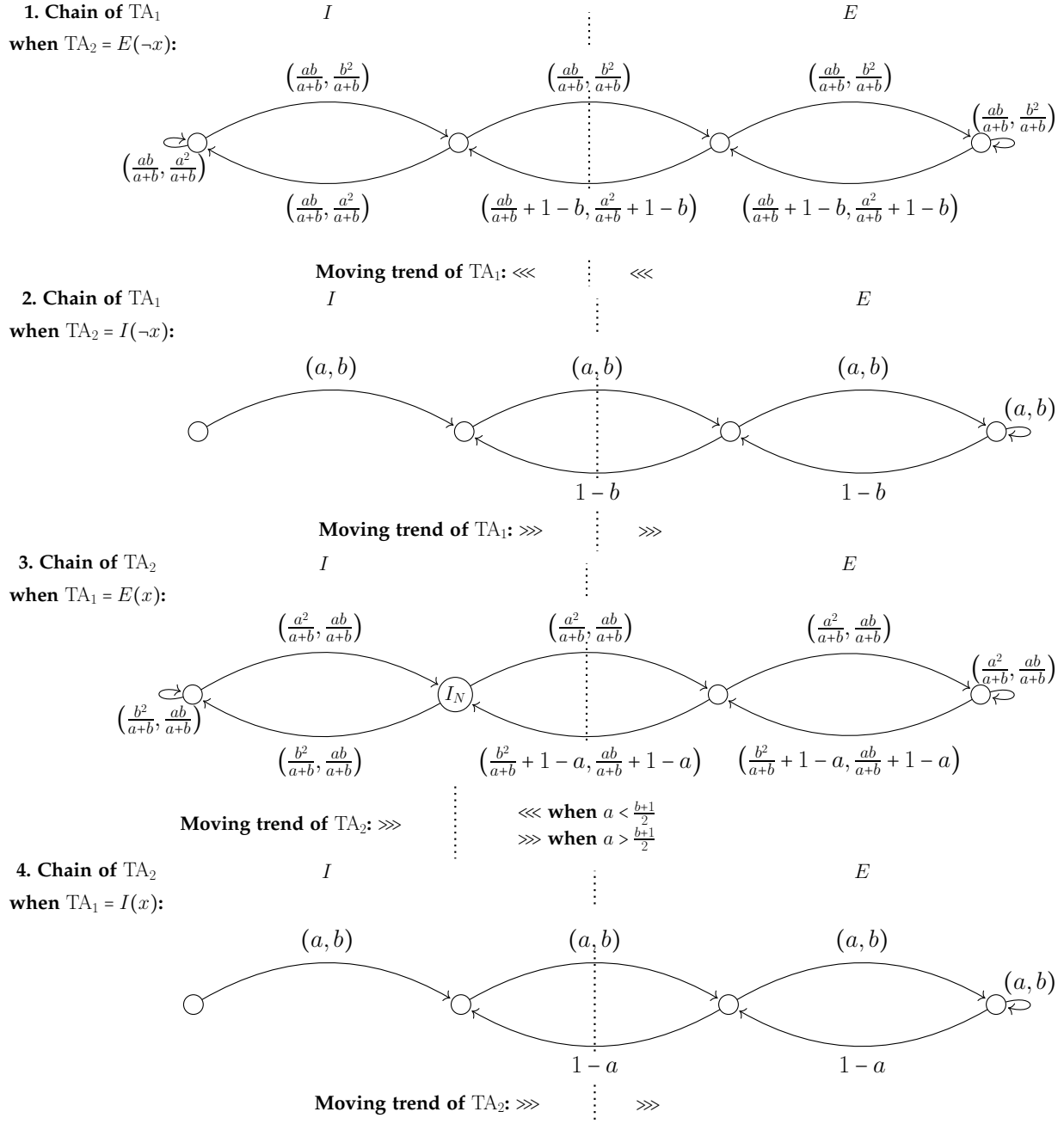


Fig. 9: Moving trend when $s \in (s_2, s_5)$. In the noisy case, where $P(y = 1|X = 1) = a$, and $P(y = 1|X = 0) = b$; $P(X = 1) = c$; $0.5 < b < a < 1$, $c \in (0, 1)$.

Biased data. We now analyze the convergence of the TM in a more generic environment where training samples can be biased, i.e., when $0 < c < 1$ and $c \neq 0.5$. We state that the general case share the same conclusions with the unbiased case for the following environments: (1) $a < 0.5, b > 0.5$, (2) $a > 0.5, b < 0.5$, (4) $a = 0.5, b = 0.5$, and (5) $a < 0.5, b < 0.5$. The difference lies in the environment where $a > 0.5, b > 0.5$. We omit the deductive process and summarize the analytical results for the environment where $a > 0.5$ and $b > 0.5$ below.

- 1) When $s < \min(s_1, s_4)$, the TM converges to $(E(x), E(\neg x))$ in both cases where $a > b > 0.5$ and $b > a > 0.5$. This part is the same as in the unbiased environment.
- 2) We then observe how the positions of s_2, s_3, s_5 , and s_6 change as c changes. We plot the changes in the case where $a > b > 0.5$ in Figure 10. There we have defined

$$c_1 = \frac{1-b}{a+1-b}, c_2 = \frac{b}{a+b}, c_3 = \frac{1-b}{2-a-b}, c_4 = \frac{b}{1-a+b}.$$

The intervals of c :

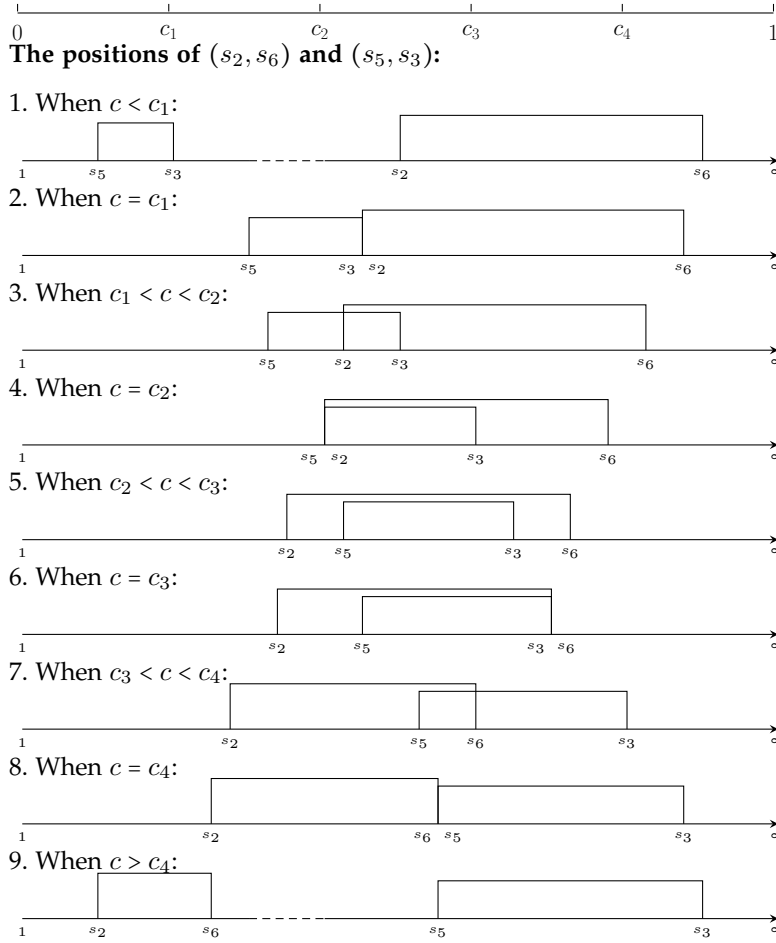


Fig. 10: How the positions of s_2, s_3, s_5, s_6 change with c , in the environment $a > b > 0.5$.

As can be seen from Figure 10, the lengths and the positions of the intervals (s_5, s_3) and (s_2, s_6) depend on where c is. Briefly speaking, the smaller the c , the shorter is the interval (s_5, s_3) , the longer is the interval (s_2, s_6) , and the farther away are these two intervals from each other. When $c \rightarrow 0$, we have $(s_2, s_6) \rightarrow (\infty, \infty)$, $(s_5, s_3) \rightarrow (1, \frac{b}{1-b})$, and the distance between these two intervals becomes the largest. When $c < c_1$, there is no overlap between intervals (s_5, s_3) and (s_2, s_6) . When c increases, the distance between these two intervals decreases, and when $c = c_1$, they become next to each other. If c continues to increase and moves into (c_1, c_2) , the intervals (s_5, s_3) and (s_2, s_6) start overlapping each other. When $c_2 \leq c \leq c_3$, (s_5, s_3) “merges” into (s_2, s_6) .

If c continues on increasing, (s_5, s_3) starts moving out of (s_2, s_6) . When $c > c_4$, (s_5, s_3) and (s_2, s_6) part from each other again and become farther away from each other. When

$c \rightarrow 1$, we have $(s_5, s_3) \rightarrow (\infty, \infty)$, $(s_2, s_6) \rightarrow (1, \frac{a}{1-a})$, and the distance between (s_5, s_3) and (s_2, s_6) reaches the greatest again.

Wherever c is, if $s \in (s_5, s_3)$, the TM converges to the NOT relation, i.e., $(TA_1, TA_2) = (E(x), I(-x))$; if $s \in (s_2, s_6)$, the TM converges to the IDENTITY relation, i.e., $(TA_1, TA_2) = (I(x), E(-x))$; and if s is in the overlapping region, the TM converges to either $(I(x), E(-x))$ or $(E(x), I(-x))$.

Figure 10 indeed reveals a crucial property of TMs: *by properly selecting a value for s , the TM is able to pick up the pattern represented by very rare samples*. For example, in the environment where c is very small, most of the samples have input $x = 0$, whose label is $y = 1$ with probability $b > 0.5$. Yet, there are very rare input $x = 1$ samples, whose label is $y = 1$ with probability $a > b > 0.5$. In this environment, if we set $s \in (s_5, s_3)$, then the TM will converge to $(E(x), I(-x))$, capturing the NOT relation represented by the dominating samples $x = 0$.

However, if we configure s to be sufficiently large, that is, within the interval (s_2, s_6) , then the TM will still converge to the IDENTITY relation $(I(x), E(\neg x))$, despite the low value of c . This ability to directly control the frequency of the patterns captured is a unique property of TMs.

Remark 2. *The convergence of TM in the noisy case do have requirements on the depths of the TAs. The tolerance of noise in TM relies on the depth of states associated with each action. The larger the depth, the more the robustness to noise.*

4 CONCLUSIONS

In this article, we studied the convergence of single-clause Tsetlin Machines (TMs) for learning two of the most fundamental logical operators, i.e., the IDENTITY and the NOT operators. The analytical results reveal that the TM, in its simplest form, can learn the above two propositional logic operators over an infinite training horizon. Besides, by appropriately configuring the granularity parameter s , the TM is able to capture the pattern represented by rare samples and select the most accurate one when two candidate patterns are incompatible. This ability to directly control the frequency and the accuracy of the patterns captured is a unique property of TMs.

The analytical approach proposed in this paper provides a foundation for formal treatment of TM learning in more complex scenarios. To this end, in our further work, we aim to extend our analysis to also cover the binary logical operators: AND, OR, and XOR. Such an analysis will involve the voting mechanism of the TM in full extent.

ACKNOWLEDGMENT

This work is supported by the project Spacetime Vision: Towards Unsupervised Learning in the 4D World financed by the EEA and Norway Grants 2014-2021 under the grant number EEA-RO-NO-2018-04.

REFERENCES

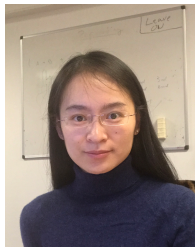
- [1] O.-C. Granmo, "The Tsetlin machine - a game theoretic bandit driven approach to optimal pattern recognition with propositional logic," *arXiv:1804.01508*, Apr 2018. [Online]. Available: <https://arxiv.org/abs/1804.01508>
- [2] M. L. Tsetlin, "On behaviour of finite automata in random medium," *Avtomat. i Telemekh.*, vol. 22, no. 10, pp. 1345–1354, 1961.
- [3] S. Sarkar and P. Soundararajan, "Supervised learning of large perceptual organization: Graph spectral partitioning and learning automata," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 5, pp. 504–525, 2000.
- [4] L. Jiao, X. Zhang, B. J. Oommen, and O.-C. Granmo, "Optimizing channel selection for cognitive radio networks using a distributed Bayesian learning automata-based approach," *Applied Intelligence*, vol. 44, no. 2, pp. 307–321, 2016.
- [5] Z. Yang, Y. Liu, Y. Chen, and L. Jiao, "Learning automata based Q-learning for content placement in cooperative caching," *IEEE Transactions on Communications*, vol. 68, no. 6, pp. 3667–3680, 2020.
- [6] X. Zhang, L. Jiao, B. J. Oommen, and O. Granmo, "A conclusive analysis of the finite-time behavior of the discretized pursuit learning automaton," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 1, pp. 284–294, 2020.
- [7] A. Yazidi, X. Zhang, L. Jiao, and B. J. Oommen, "The hierarchical continuous pursuit learning automation: A novel scheme for environments with large numbers of actions," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 2, pp. 512–526, 2020.
- [8] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you?: Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. ACM, 2016, pp. 1135–1144.
- [9] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nature Machine Intelligence*, vol. 1, no. 5, pp. 206–215, 2019. [Online]. Available: <https://doi.org/10.1038/s42256-019-0048-x>
- [10] C. D. Blakely and O.-C. Granmo, "Closed-Form Expressions for Global and Local Interpretation of Tsetlin Machines," in *34th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems (IEA/AIE 2021)*. Springer, 2021.
- [11] A. Wheeldon, R. Shafik, T. Rahman, J. Lei, A. Yakovlev, and O.-C. Granmo, "Learning automata based energy-efficient ai hardware design for IoT," *Philosophical Transactions of the Royal Society A*, 2020.
- [12] R. Shafik, A. Wheeldon, and A. Yakovlev, "Explainability and dependability analysis of learning automata based ai hardware," in *IEEE 26th International Symposium on On-Line Testing and Robust System Design (IOLTS)*. IEEE, 2020.
- [13] O.-C. Granmo, S. Glimsdal, L. Jiao, M. Goodwin, C. W. Omlin, and G. T. Berge, "The convolutional Tsetlin machine," *arXiv preprint arXiv:1905.09688*, 2019.
- [14] K. D. Abeyrathna, O.-C. Granmo, X. Zhang, L. Jiao, and M. Goodwin, "The regression Tsetlin machine - a novel approach to interpretable non-linear regression," *Philosophical Transactions of the Royal Society A*, vol. 378, 2019.
- [15] K. D. Abeyrathna, O.-C. Granmo, and M. Goodwin, "A regression Tsetlin machine with integer weighted clauses for compact pattern representation," in *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*. Springer, 2020.
- [16] R. Saha, O.-C. Granmo, V. I. Zadorozhny, and M. Goodwin, "A Relational Tsetlin Machine with Applications to Natural Language Understanding," *arXiv preprint arXiv:2102.10952*, 2021. [Online]. Available: <https://arxiv.org/abs/2102.10952>
- [17] K. D. Abeyrathna, B. Bhattarai, M. Goodwin, S. Gorji, O.-C. Granmo, L. Jiao, R. Saha, and R. K. Yadav, "Massively parallel and asynchronous Tsetlin machine architecture supporting almost constant-time scaling," *arXiv preprint arXiv:2009.04861*, 2020.
- [18] G. T. Berge, O.-C. Granmo, T. O. Tveit, M. Goodwin, L. Jiao, and B. V. Matheussen, "Using the Tsetlin machine to learn human-interpretable rules for high-accuracy text categorization with medical applications," *IEEE Access*, vol. 7, pp. 115 134–115 146, 2019.
- [19] R. K. Yadav, L. Jiao, O.-C. Granmo, and M. Goodwin, "Distributed Word Representation in Tsetlin Machine," *arXiv preprint arXiv:2104.06901*, 2021. [Online]. Available: <https://arxiv.org/abs/2104.06901>
- [20] D. C. Nicolae, "Question Classification using Interpretable Tsetlin Machine," in *The 1st International Workshop on Machine Reasoning (MRC 2021)*, 2021.
- [21] R. K. Yadav, L. Jiao, O.-C. Granmo, and M. Goodwin, "Interpretability in word sense disambiguation using Tsetlin machine," in *13th International Conference on Agents and*

Artificial Intelligence (ICAART), Vienna, Austria. INSTICC, 2021.

- [22] B. Bhattarai, O.-C. Granmo, and L. Jiao, "Measuring the novelty of natural language text using the conjunctive clauses of a Tsetlin machine text classifier," in *13th International Conference on Agents and Artificial Intelligence (ICAART), Vienna, Austria.* INSTICC, 2021.
- [23] R. Saha, O.-C. Granmo, and M. Goodwin, "Mining Interpretable Rules for Sentiment and Semantic Relation Analysis using Tsetlin Machines," in *Proceedings of the 40th International Conference on Innovative Techniques and Applications of Artificial Intelligence (SGAI), Cambridge, UK.* Springer International Publishing, 2020.
- [24] J. Lei, T. Rahman, R. Shafik, A. Wheeldon, A. Yakovlev, O.-C. Granmo, F. Kawsar, and A. Mathur, "Low-Power Audio Keyword Spotting Using Tsetlin Machines," *Journal of Low Power Electronics and Applications*, vol. 11, 2021. [Online]. Available: <https://www.mdpi.com/2079-9268/11/2/18>
- [25] R. K. Yadav, L. Jiao, O.-C. Granmo, and M. Goodwin, "Human-level interpretable learning for aspect-based sentiment analysis," in *Proceedings of AAAI, Vancouver, Canada.* AAAI, 2021.
- [26] K. S. Narendra and M. A. L. Thathachar, *Learning Automata: An Introduction.* Prentice-Hall, Inc., 1989.



Ole-Christoffer Granmo is a Professor and Founding Director of Centre for Artificial Intelligence Research (CAIR), University of Agder, Norway. He obtained his master's degree in 1999 and the PhD degree in 2004, both from the University of Oslo, Norway. Dr. Granmo has authored in excess of 140 refereed papers with numerous best paper awards, encompassing learning automata, bandit algorithms, Tsetlin machines, Bayesian reasoning, reinforcement learning, and computational linguistics. He has further coordinated 7+ Norwegian Research Council projects and graduated more than 60 master- and PhD students. Dr. Granmo is also a co-founder of the Norwegian Artificial Intelligence Consortium (NORA). Apart from his academic endeavours, he co-founded Anzyz Technologies AS..



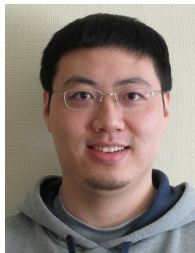
Xuan Zhang received her PhD degree in Information and Communication Technology from the University of Agder (UiA) in 2015. She has a Master's degree in Signal and Information Processing and a Bachelor's degree in Electronics and Information Engineering. She is now a senior researcher in Norwegian Research Center (NORCE). At the same time, she is a scientific researcher at Centre of Artificial Intelligence Research (CAIR) in

UiA. She is currently serving as a Board Member of Norwegian Association for Image Processing and Machine Learning. Her research interests include: Learning Automata, Mathematical Analysis on Learning Algorithms, Stochastic Modeling and Optimization, Deep Learning, Natural Language Processing and Computer Vision.



Morten Goodwin received the B.Sc. and M.Sc. degrees from the University of Agder, Norway, in 2003 and 2005, respectively, and the Ph.D. degree from Aalborg University Department of Computer Science, Denmark, in 2011, on applying machine learning algorithms on eGovernment indicators which are difficult to measure automatically. He is a Professor with the Department of ICT, the University of Agder, deputy director

for Centre for Artificial Intelligence Research, a public speaker, and an active researcher. His main research interests include machine learning, including swarm intelligence, deep learning, and adaptive learning in medicine, games, and chatbots. He has more than 100 peer reviews of scientific publications. He has supervised more than 110 student projects, including Master and Ph. D. theses within these topics, and more than 90 popular science public speaking events, mostly in Artificial Intelligence.



Lei Jiao (M'12-SM'18) received his BE degree from Hunan University, China, in 2005. He received his ME degree from Shandong University, China, in 2008. He obtained his PhD degree in Information and Communications Technology from University of Agder (UiA), Norway, in 2012. He is now an Associated Professor in the Department of Information and Communication Technology, UiA. His research interests include reinforcement

learning, learning automata, natural language processing, resource allocation and performance evaluation for communication and energy systems.

TABLE 5: State transitions per scenarios, in the noise-free case, where $P(y = 1|X = 1) = 1$, and $P(y = 0|X = 0) = 1$; $P(X = 1) = c$; $c \in (0, 1)$.

(TA_1, TA_2)	Scenario index	(X, y)	$C(X)$	State transitions for TA_1 :	State transitions for TA_2 :
$(E(x), E(\neg x))$	Scenario 1	$(1, 1)$	1		
	Scenario 2	$(0, 0)$	1		No change.
$(E(x), I(\neg x))$	Scenario 3	$(1, 1)$	0		
	Scenario 4	$(0, 0)$	1		No change.
$(I(x), E(\neg x))$	Scenario 5	$(1, 1)$	1		
	Scenario 6	$(0, 0)$	0	No change.	No change.
$(I(x), I(\neg x))$	Scenario 7	$(1, 1)$	0		
	Scenario 8	$(0, 0)$	0	No change.	No change.

TABLE 6: State transitions per senarios, part 1, in the case with noise, where $P(y = 1|X = 1) = a$, and $P(y = 1|X = 0) = b$; $P(X = 1) = c$; $a, b, c \in (0, 1)$.

(TA_1, TA_2)	Scenario index	(X, y)	$C(X)$	State transitions for TA_1 :	State transitions for TA_2 :
$(E(x), E(\neg x))$	Scenario 1	(1,1)	1		
	Scenario 2	(1,0)	0	No change.	
	Scenario 3	(0,0)	1		No change.
	Scenario 4	(0,1)	1		
$(E(x), I(\neg x))$	Scenario 5	(1,1)	0		
	Scenario 6	(1,0)	0	No change.	No change.
	Scenario 7	(0,0)	1		No change.
	Scenario 8	(0,1)	1		

TABLE 7: State transitions per scenarios, part 2, in the case with noise, where $P(y = 1|X = 1) = a$, and $P(y = 1|X = 0) = b$; $P(X = 1) = c$; $a, b, c \in (0, 1)$.

(TA_1, TA_2)	Scenario index	(X, y)	$C(X)$	State transitions for TA_1 :	State transitions for TA_2 :
$(I(x), E(\neg x))$	Scenario 9	$(1, 1)$	1		
	Scenario 10	$(1, 0)$	1	No change.	
	Scenario 11	$(0, 0)$	0	No change.	No change.
	Scenario 12	$(0, 1)$	0		
$(I(x), I(\neg x))$	Scenario 13	$(1, 1)$	0		
	Scenario 14	$(1, 0)$	0	No change.	No change.
	Scenario 15	$(0, 0)$	0	No change.	No change.
	Scenario 16	$(0, 1)$	0		