Measuring Interestingness of Theorems in Automated Theorem Finding by Forward Reasoning: A Case Study in Tarski's Geometry

1st Hongbiao Gao[1,2]
[1]Department of Information and Computer Sciences, Saitama University, Saitama, Japan
[2]School of Control and Computer Engineering, North China Electric Power University, Beijing, China
gaohongbiao@aise.ics.saitama-u.ac.jp

2nd Jianbin Li
School of Control and Computer Engineering
North China Electric Power University, Beijing, China
lijb87@ncepu.edu.cn

3rd Jingde Cheng
Department of Information and Computer Sciences, Saitama University, Saitama, Japan
cheng@aise.ics.saitama-u.ac.jp

***Abstract*—The problem of automated theorem finding is one of 33 basic research problems in automated reasoning which was originally proposed by Wos in 1988, and it is still an open problem. The problem implicitly requires some metrics to be used for measuring interestingness of found theorems. A set of metrics for measuring interestingness of theorems in automated theorem finding by forward reasoning have been proposed, and case studies to measure interestingness of the theorems of NBG set theory and Peano's arithmetic were performed. This paper presents a case study in Tarski's geometry to show the generality of proposed metrics. In the case study, we use the metrics to measure interestingness of the theorems of Tarski's geometry obtained by using forward reasoning approach, and confirm the effectiveness of the metrics.**

***Index Terms*—Metric; automated theorem finding; forward reasoning; strong relevant logic; Tarski's geometry**

I. Introduction

The problem of automated theorem finding (ATF for short) is one of the 33 basic research problems in automated reasoning which was originally proposed by Wos in 1988 [18], [19], and it is still an open problem until now [11]. The problem of ATF is "What properties can be identified to permit an automated reasoning program to find new and interesting theorems, as opposed to proving conjectured theorems?" [18], [19]. The most important and difficult requirement of the problem is that, in contrast to proving conjectured theorems supplied by the user, it asks for the criteria that an automated reasoning program can use to find some theorems in a field that must be evaluated by theorists of the field as new and interesting theorems. The significance of solving the problem is obvious because an automated reasoning program satisfying the requirement can provide great assistance for scientists in various fields [1]–[3].

A few works aimed to automated theorem discovery (ATD) and automated theorem generation (ATG) have been done [5]–[7], [11], [13]–[15]. However, the problem of ATF is different from the ATD and ATG such that their works are not suitable to be used in ATF. In fact, Wos's problem can be regarded as an attempt to find a systematic methodology in automated reasoning area, but the works on ATD and ATG almost aim to one certain mathematical field. Besides, the works of ATD and ATG rely on the approach of automated theorem proving, however, if we want to find new and interesting theorems, the only way is to use forward reasoning approach [2], [3].

To solve the ATF problem, a systematic methodology for ATF by using forward reasoning approach based on strong relevant logics was proposed [1]–[3], [10]. Following the proposed methodology, we have proposed a set of metrics to help us to find new and interesting theorems, and case studies to measure interestingness of the theorems of NBG set theory and Peano's arithmetic were presented [9], [12].

This paper presents a case study in Tarski's geometry to show the generality of proposed metrics. In the case study, we use the metrics to measure interestingness of the theorems of Tarski's geometry obtained by using forward reasoning approach, and confirm the effectiveness of the metrics.

II. Basic Notions and Notations

A formal logic system L is an ordered pair $(F(L), \vdash_L)$ where $F(L)$ is the set of well formed formulas of L, and $\vdash_L$ is the consequence relation of L such that for a set P of formulas and a formula C, $P \vdash_L C$ means that within the framework of L taking P as premises we can obtain C as a valid conclusion. $Th(L)$ is the set of logical theorems of L such that $\phi \vdash_L T$ holds for any $T \in Th(L)$. According to the representation of the consequence relation of a logic, the logic can be represented as a Hilbert style system, a Gentzen sequent calculus system, a Gentzen natural deduction system, and so on [3].

Let $(F(L), \vdash_L)$ be a formal logic system and $P \subseteq F(L)$ be a non-empty set of sentences. A formal theory with premises P based on L, called a L-theory with premises P and denoted by $T_L(P)$, is defined as $T_L(P) =_{df} Th(L) \cup Th^e_L(P)$ where $Th^e_L(P) =_{df} \{A | P \vdash_L A \ and \ A \notin Th(L)\}$, $Th(L)$ and $Th^e_L(P)$ are called the logical part and the empirical part of the formal theory, respectively, and any element of $Th^e_L(P)$ is called an empirical theorem of the formal theory [3].

Based on the definition above, the problem of ATF can be said as "for any given premises P, how to construct a meaningful formal theory $T_L(P)$ and then find new and interesting theorems in $Th^e_L(P)$ automatically?" [3].

The notion of degree of a logical connective [3] is defined as follows: Let θ be an arbitrary n-ary $(1 \leq n)$ connective of logic L and A be a formula of L, the degree of θ in A, denoted by $D_\theta(A)$, is defined as follows: (1) $D_\theta(A) = 0$ if and only if there is no occurrence of θ in A, (2) if A is in the form $\theta(a_1, a_2, ..., a_n)$ where $a_1, a_2, ..., a_n$ are formulas, then $D_\theta(A) = max\{D_\theta(a_1), D_\theta(a_2), ..., D_\theta(a_n)\} + 1$, (3) if A is in the form $\sigma(a_1, a_2, ..., a_n)$ where σ is a connective different from θ and $a_1, a_2, ..., a_n$ are formulas, then $D_\theta(A) = max\{D_\theta(a_1), D_\theta(a_2), ..., D_\theta(a_n)\}$, and (4) if A is in the form QB where B is a formula and Q is the quantifier prefix of B, then $D_\theta(A) = D_\theta(B)$.

The notion of predicate abstract level [10] is defined as follows: (1) Let $pal(X) = k$ denote that an abstract level of a predicate X is k where k is a natural number, (2) $pal(X) = 1$ if X is the most primitive predicate in the target field, (3) $pal(X) = max(pal(Y_1), pal(Y_2), ..., pal(Y_n)) + 1$ if a predicate X is defined by other predicates $Y_1, Y_2, ..., Y_n$ in the target field where n is a natural number. A predicate X is called k-level predicate, if $pal(X) = k$. If $pal(X) < pal(Y)$, we call the abstract level of predicate X is lower than Y, and Y is higher than X.

The notion of function abstract level [10] is defined as follows: (1) Let $fal(f) = k$ denote that an abstract level of a function f is k where k is a natural number, (2) $fal(f) = 1$ if f is the most primitive function in the target field, (3) $fal(f) = max(fal(g_1), fal(g_2), ..., fal(g_n)) + 1$ if a function f is defined by other functions $g_1, g_2, ..., g_n$ in the target field where n is a natural number. A function f is k-level function, if $fal(f) = k$. If $fal(f) < fal(g)$, we call the abstract level of function f is lower than g, and g is higher than f.

The notion of abstract level [10] of a formula is defined as follows: (1) Let $lfal(A) = (k, m)$ denote that an abstract level of a formula A where $k = pal(A)$ and $m = fal(A)$, (2) $pal(A) = max(pal(Q_1), pal(Q_2), ..., pal(Q_n))$ where Q_i is a predicate and occurs in A $(1 \leq i \leq n)$, or $pal(A) = 0$, if there is not any predicate in A, (3) $fal(A) = max(fal(g_1), fal(g_2), ..., fal(g_n))$ where g_i is a function and occurs in A $(1 \leq i \leq n)$, or $fal(A) = 0$, if there is not any function in A. A formula A is (k, m)-level formula, if $lfal(A) = (k, m)$.

TABLE I
DEGREE OF $\Rightarrow$ OF COLLECTED KNOWN THEOREMS IN NBG SET THEORY

Degree	Appeared time	Appeared rate
$\Rightarrow$,0	242	56%
$\Rightarrow$,1	187	44%
$\Rightarrow$,2	0	0%
$\Rightarrow$,3	0	0%
$\Rightarrow$,4	0	0%

The deduction distance by using Modus Ponens is defined as below: (1) $Dist(A) = 0$, if A an axiom; (2) $Dist(A) = max(Dist(\alpha), Dist(\beta)) + 1$, if A is deduced from two empirical theorems α and β by using Modus Ponens; (3) $Dist(A) = Dist(\alpha) + 1$, if A is deduced from an empirical theorem α and a logical theorem β by using Modus Ponens; (4) $Dist(A) = Dist(\alpha)$, if A is abstracted from α.

The propositional schema of a first-order logical formula can be obtained by removing all of quantifiers and replacing all of atomic formulas of a first-order logical formula with propositional atomic formulas. For example, $\forall x \forall y (((x = y) \Rightarrow (y = x)) \Rightarrow (x \subseteq y))$ is translated into $(A \Rightarrow B) \Rightarrow C$.

III. Factors Related to Interestingness of Theorems

We consider that plural factors relate to the interestingness of found theorems by ATF. The plural factors are the degree of logical connectives in empirical theorems, propositional schema of empirical theorems, abstract level of empirical theorems, and deduction distance of empirical theorems.

Degree of logical connectives

The first factor related to interestingness of theorems is the degree of logical connectives in empirical theorems. We have analyzed more than 400 known theorems of NBG set theory, near 1,000 known theorems of Peano's arithmetic and 87 known theorems of Tarski's geometry in Quaife's book [16] about the degree of logical connectives, and our analysis results are shown in Table I-XII. We found the degrees of the logical connectives of those known theorems are almost lower than 2. Therefore, the degree of logical connectives is related to the interestingness of empirical theorems, and interesting theorems always hold lower degree of logical connectives. The reason is that those theorems holding high degree of logical connectives are hard to be understood and mathematicians always introduce new predicates to abstract the formula holding higher degree of logical connectives. Cheng conjectured that almost all new theorems and questions of a formal theory can be deduced from the premises of that theory by finite inference steps concerned with finite number of low degree entailments [3].

Propositional schema of formula

The second factor is propositional schema of formula that we have defined in Section 2. We consider that the interesting theorems hold some frequent propositional schemata, after

TABLE II
DEGREE OF $\wedge$ OF COLLECTED KNOWN THEOREMS IN NBG SET THEORY

Degree	Appeared time	Appeared rate
$\wedge$,0	356	83%
$\wedge$,1	64	15%
$\wedge$,2	7	<2%
$\wedge$,3	2	<1%
$\wedge$,4	0	0%

TABLE III
DEGREE OF $\vee$ OF COLLECTED KNOWN THEOREMS IN NBG SET THEORY

Degree	Appeared time	Appeared rate
$\vee$,0	381	89%
$\vee$,1	43	10%
$\vee$,2	5	1%
$\vee$,3	0	0%
$\vee$,4	0	0%

TABLE IV
DEGREE OF $\neg$ OF COLLECTED KNOWN THEOREMS IN NBG SET THEORY

Degree	Appeared time	Appeared rate
$\neg$,0	404	94%
$\neg$,1	25	6%
$\neg$,2	0	0%
$\neg$,3	0	0%
$\neg$,4	0	0%

TABLE V
DEGREE OF $\Rightarrow$ OF COLLECTED KNOWN THEOREMS IN PEANO'S ARITHMETIC

Degree	Appeared time	Appeared rate
$\Rightarrow$,0	486	51%
$\Rightarrow$,1	473	49%
$\Rightarrow$,2	0	0%
$\Rightarrow$,3	0	0%
$\Rightarrow$,4	0	0%

TABLE VI
DEGREE OF $\wedge$ OF COLLECTED KNOWN THEOREMS IN PEANO'S ARITHMETIC

Degree	Appeared time	Appeared rate
$\wedge$,0	782	82%
$\wedge$,1	144	15%
$\wedge$,2	28	<3%
$\wedge$,3	4	<1%
$\wedge$,4	1	<1%

TABLE VII
DEGREE OF $\vee$ OF COLLECTED KNOWN THEOREMS IN PEANO'S ARITHMETIC

Degree	Appeared time	Appeared rate
$\vee$,0	795	83%
$\vee$,1	146	15%
$\vee$,2	18	2%
$\vee$,3	0	0%
$\vee$,4	0	0%

TABLE VIII
DEGREE OF $\neg$ OF COLLECTED KNOWN THEOREMS IN PEANO'S ARITHMETIC

Degree	Appeared time	Appeared rate
$\neg$,0	881	92%
$\neg$,1	78	8%
$\neg$,2	0	0%
$\neg$,3	0	0%
$\neg$,4	0	0%

TABLE IX
DEGREE OF $\Rightarrow$ OF COLLECTED KNOWN THEOREMS IN TARSKI'S GEOMETRY

Degree	Appeared time	Appeared rate
$\Rightarrow$,0	22	25%
$\Rightarrow$,1	65	75%
$\Rightarrow$,2	0	0%
$\Rightarrow$,3	0	0%
$\Rightarrow$,4	0	0%

we investigated the propositional schemata of more than 400 known theorems of NBG set theory, near 1,000 known theorems of Peano's arithmetic and 87 known theorems of Tarski's geometry. The most frequent propositional schemata of known theorems is A type. A theorem is always interesting if the theorem does not contain any logical connective, because it holds clear and concise semantics. The second frequent propositional schema is $A \Rightarrow B$. We think the reason is that "if A then B " is a very frequent conditional propositional schema in any fields. Other frequent propositional schemata have been also shown in Table XIII-XV. The analysis results show that known theorems always hold some frequent propositional schemata. We can see known theorems as found interesting theorems, so we consider that the new and interesting theorems may also holds those frequent propositional schemata.

Abstract level

The third factor is the abstract level of predicates and

TABLE X
DEGREE OF $\wedge$ OF COLLECTED KNOWN THEOREMS IN TARSKI'S GEOMETRY

Degree	Appeared time	Appeared rate
$\wedge$,0	46	53%
$\wedge$,1	24	28%
$\wedge$,2	7	8%
$\wedge$,3	3	3%
$\wedge$,4	2	2%
$\wedge$,5	5	6%

TABLE XI
DEGREE OF $\vee$ OF COLLECTED KNOWN THEOREMS IN TARSKI'S GEOMETRY

Degree	Appeared time	Appeared rate
$\vee$,0	62	71%
$\vee$,1	20	23%
$\vee$,2	5	6%
$\vee$,3	0	0%
$\vee$,4	0	0%

TABLE XII
DEGREE OF $\neg$ OF COLLECTED KNOWN THEOREMS IN TARSKI'S GEOMETRY

Degree	Appeared time	Appeared rate
$\neg$,0	82	94%
$\neg$,1	5	6%
$\neg$,2	0	0%
$\neg$,3	0	0%
$\neg$,4	0	0%

TABLE XIII
FREQUENT PROPOSITIONAL SCHEMATA OF COLLECTED KNOWN THEOREMS IN NBG SET THEORY

Propositional schema	Appeared time	Appeared rate
A	186	43%
$A \Rightarrow B$	108	25%
$(A \wedge B) \Rightarrow C$	54	13%
$A \vee B$	26	6%
$A \Rightarrow (B \vee C)$	17	4%
$\neg A$	14	3%
$\neg(A \wedge B)$	10	2%
$(A \wedge B \wedge C) \Rightarrow D$	6	1%
$A \vee B \vee C$	5	1%
$(A \wedge B \wedge C \wedge D) \Rightarrow E$	2	<1%
$\neg(A \wedge B \wedge C)$	1	<1%

functions in one theorem. In the mathematical fields, mathematicians always make definition from simple to complex. For example, the predicate "$\in$" is the most basic predicate in the set theory. Then the mathematicians define the predicate "$\subseteq$" which is a higher level predicate than "$\in$", and abstracts from "$\in$" by the definition of "$\subseteq$": $\forall x \forall y (\forall u ((u \in x) \Rightarrow (u \in y)) \Leftrightarrow (x \subseteq y))$. Then the mathematicians define the predicate "=" which is a higher level predicate than "$\subseteq$", and abstracts from "$\subseteq$" by the axiom: $\forall x \forall y (((x \subseteq y) \wedge (y \subseteq x)) \Leftrightarrow (x = y))$. Based on the fact, we can consider that a theorem holds higher abstract level predicates and functions, the theorem is more interesting from the viewpoint of the meaning of the theorem.

Deduction distance

The fourth factor is deduction distance. If a theorem can be reasoned out by several steps, the theorem is easy to be found and is obvious to be understood by observing used premises. The interesting theorems are those theorems which are difficult to be reasoned out from premises. Therefore, if the deduction distance of an obtained theorem is long, the theorem may be interesting.

IV. A SET OF METRICS FOR MEASURING INTERESTINGNESS OF THEOREMS

Our metrics to measure the interestingness of obtained empirical theorems consists of parameters about the degree of logical connectives, propositional schema of formula, abstract level and deduction distance, and we use four variables Vd, Vp, Va, Ve to represent four parameters respectively. In detail, the parameter about the degree of logical connective is defined as Vd = $Value_{\Rightarrow} * Value_{\wedge} * Value_{\vee} * Value_{\neg}$. We showed the value of parameter about the degree of logical connective in Table XVI-XIX. Second, we presented the value of parameter about the propositional schemata of formula in Table XX. We assign the value 0 for empirical theorems containing a tautology part, because if one theorem contains a tautology part, this empirical theorem must not be an interesting empirical theorem. Third, if the abstract level of one empirical theorem is (k, m), then the value of parameter about abstract level of one theorem is defined as $Va = k{+}m$. Fourth, if the deduction distance of one empirical theorem is $Dist(A)$, then the value about the parameter about deduction distance is defined as $Ve = Dist(A)$. By using four parameters Vd, Vp, Va and Ve, we can use several metrics to measure the interestingness of an obtained empirical theorems, such as: Vd, Vp, Va, Ve, $Vd * Vp$, $Vd * Va$, $Vd * Ve$, $Vp * Va$, $Vp{*}Ve$, $Va{*}Ve$, $Vd{*}Vp{*}Va$, $Vd{*}Vp{*}Ve$, $Vp{*}Va{*}Ve$, $Vd{*}Va{*}Ve$, $Vd{*}Vp{*}Va{*}Ve$. The value is bigger, theorem is more interesting.

TABLE XIV
FREQUENT PROPOSITIONAL SCHEMATA OF COLLECTED KNOWN THEOREMS IN PEANO'S ARITHMETIC

Propositional schema	Appeared time	Appeared rate
A	321	33%
$A \Rightarrow B$	270	28%
$(A \wedge B) \Rightarrow C$	101	11%
$A \vee B$	72	8%
$A \Rightarrow (B \vee C)$	65	7%
$\neg A$	37	4%
$\neg(A \wedge B)$	35	4%
$(A \wedge B \wedge C) \Rightarrow D$	21	2%
$A \vee B \vee C$	15	2%
$(A \wedge B) \Rightarrow (C \vee D)$	8	<1%
$\neg(A \wedge B \wedge C)$	6	<1%
$(A \wedge B \wedge C \wedge D) \Rightarrow E$	3	<1%
$A \Rightarrow (B \vee C \vee D)$	2	<1%
$(A \wedge B \wedge C \wedge D) \Rightarrow (E \vee F \vee G)$	1	<1%
$(A \wedge B \wedge C \wedge D \wedge E) \Rightarrow F$	1	<1%
$(A \wedge B \wedge C) \Rightarrow (D \vee E)$	1	<1%

TABLE XV
FREQUENT PROPOSITIONAL SCHEMATA OF COLLECTED KNOWN THEOREMS IN TARSKI'S GEOMETRY

Propositional schema	Appeared time	Appeared rate
$A \Rightarrow B$	22	25%
A	16	18%
$(A \wedge B) \Rightarrow C$	11	13%
$(A \wedge B) \Rightarrow (C \vee D)$	10	11%
$\neg A$	5	6%
$(A \wedge B \wedge C) \Rightarrow (D \vee E)$	4	5%
$(A \wedge B \wedge C) \Rightarrow D$	3	3%
$(A \wedge B \wedge C \wedge D) \Rightarrow E$	3	3%
$(A \wedge B) \Rightarrow (C \vee D \vee E)$	3	3%
$(A \wedge B \wedge C \wedge D \wedge E \wedge F) \Rightarrow (G \vee H)$	3	3%
$(A \wedge B \wedge C \wedge D \wedge E) \Rightarrow (F \vee G \vee H)$	2	2%
$A \Rightarrow (B \vee C)$	2	2%
$(A \wedge B \wedge C \wedge D \wedge E \wedge F) \Rightarrow G$	2	2%
$B \vee C$	1	1%

V. CASE STUDY IN TARSKI'S GEOMETRY

"In his 1926 - 1927 lectures at the University of Warsaw, Alfred Tarski gave an axiomatic development of elementary Euclidean geometry, the part of plane Euclidean geometry that is not based upon set-theoretical notions, or, in other words,

TABLE XVI
THE VALUE OF PARAMETER ABOUT DEGREE OF $\Rightarrow$

Logical connective	Degree	Value
$\Rightarrow$	0	1
$\Rightarrow$	1	1
$\Rightarrow$	2	1/2
$\Rightarrow$	3	1/3
$\Rightarrow$	n	1/n

TABLE XVII
THE VALUE OF PARAMETER ABOUT DEGREE OF $\wedge$

Logical connective	Degree	Value
$\wedge$	0	1
$\wedge$	1	1
$\wedge$	2	1/2
$\wedge$	3	1/3
$\wedge$	n	1/n

the part that can be developed within the framework of first-order logic" [17]. We have not used our metric to measure the empirical theorems of Tarski's geometry.

The purpose of the case study was to confirm the generality of the proposed metrics. In the case study, we applied the proposed metrics of interestingness in empirical theorems of Tarski's geometry obtained by forward reasoning approach. We collected the axioms and definitions of Tarski's geometry from Quaife's book [16]. We used all of axioms and definitions of Tarski's geometry in Quaife's book as premises, performed automated forward reasoning by using FreeEnCal [4], and obtained empirical theorems of Tarski's geometry [8]. Then, we applied the proposed metrics to the obtained empirical theorems as same as case studies in NBG set theory and Peano's arithmetic [9], [12]. In detail, we measured Vd and Vp of those empirical theorems. To measure Va, we summarized the abstract levels of the predicates of Tarski's geometry in Quaife's book. Then, we also summarized the abstract levels of the functions of Tarski's geometry in Quaife's book. Finally, we also recorded Ve for each empirical theorem according to the information provided by FreeEnCal.

The case studies in NBG set theory and Peano's arithmetic

TABLE XVIII
THE VALUE OF PARAMETER ABOUT DEGREE OF $\vee$

Logical connective	Degree	Value
$\vee$	0	1
$\vee$	1	1
$\vee$	2	1/2
$\vee$	3	1/3
$\vee$	n	1/n

TABLE XIX
THE VALUE OF PARAMETER ABOUT DEGREE OF $\neg$

Logical connective	Degree	Value
$\neg$	0	1
$\neg$	1	1
$\neg$	2	1/2
$\neg$	3	1/3
$\neg$	n	1/n

TABLE XX
THE VALUE OF PARAMETER ABOUT PROPOSITIONAL SCHEMATA OF FORMULA

Propositional schema	Value
A	3
$A \Rightarrow B$	3
$\neg A$	2
$\neg(A_1 \wedge ... \wedge A_n)$	2
$A_1 \vee ... \vee A_n$	2
$(A_1 \wedge ... \wedge A_n) \Rightarrow (B_1 \vee ... \vee B_n)$	2
Infrequent propositional schema	1
Propositional schema containing tautology	0

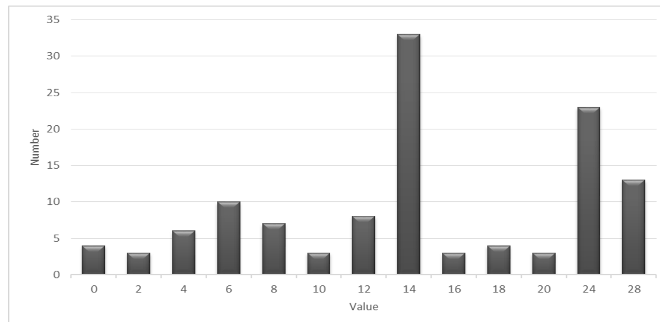

Fig. 1. The number of empirical theorems of Tarski's geometry on each value

showed that the combination $Vd{*}Vp{*}Va{*}Ve$ [9], [12] is well, because range of values is wide and deviation is obvious such that we can easily distinguish the weight of interestingness for empirical theorems. Therefore, in the case study, we also use the combination $Vd*Vp*Va*Ve$ as metric, and use it to measure the interestingness of Tarski's geometry. Then we investigated how many empirical theorems on each value for the combination $Vd*Vp*Va*Ve$ and showed the results in Fig. 1.

Comparing the investigated results with the results in case studies of NBG set theory and Peano's arithmetic, we found the following facts. First, our metrics can generally filter uninteresting theorems from all of obtained empirical theorems in different mathematical fields (the case study of NBG set theory, Peano's arithmetic and Tarski's geometry), Second, the empirical theorems whose values are in middle part are most, and the empirical theorems which hold lower value are few. However, in the case study of NBG set theory and Peano's arithmetic, the empirical theorems which hold higher values are few, but in the case study of Tarski's geometry, the empirical theorems which hold higher value are many. We consider that the reason is that the range of value of interestingness is 0-28 in the case study, but 0-28 is only a part in the last two case studies. Maybe the empirical theorems which hold higher value will be diminishing, if the maximum value is extended.

VI. CONCLUDING REMARKS

We have presented a case study in Tarski's geometry, in which we used the proposed metrics to measure the interestingness of empirical theorems reasoned out by forward

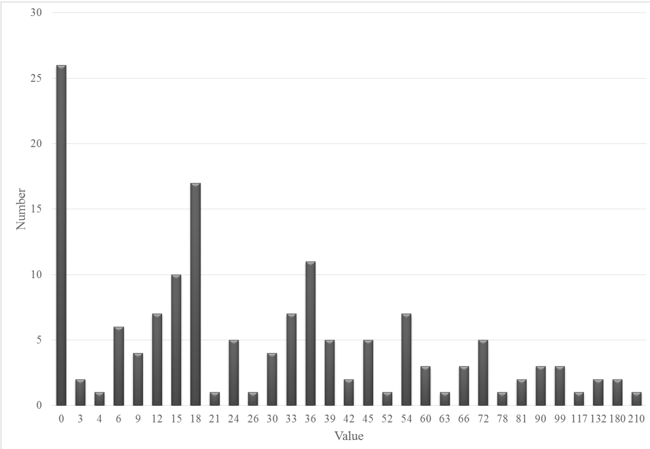

Fig. 2. The number of empirical theorems of NBG set theory on each value

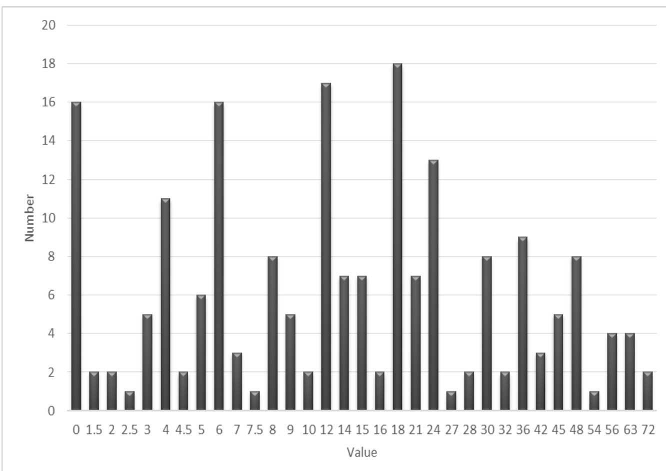

Fig. 3. The number of empirical theorems of Peano's arithmetic on each value

reasoning approach. The result of the case study showed that our metrics can be used in ATF of different mathematical fields.

There are many interesting and challenging research problems in our future works. First, we will confirm the proposed metrics by measuring interestingness of known mathematical theorems in mathematical books, however current works only apply those metrics in empirical theorems obtained by forward reasoning approach. We will use our metrics to measure the interestingness of known theorems in mathematical books and sort the order from low value to high value, then we compare the sorted order with the appearing order of those known theorems in mathematical books. We expect two orders are almost same, because known theorems in mathematical books are always recorded from simple to complex. Second, we will do case studies of ATF in other fields to confirm the generality of the metrics, such as graph theory and lattice theory.

REFERENCES

[1] J. Cheng, "A Relevant Logic Approach to Automated Theorem Finding," Proceedings of the Workshop on Automated Theorem Proving attached to International Symposium on Fifth Generation Computer Systems, pp. 8-15, 1994.

[2] J. Cheng, "Entailment Calculus as the Logical Basis of Automated Theorem Finding in Scientific Discovery," Systematic Methods of Scientific Discovery: Papers from the 1995 Spring Symposium, AAAI Press - American Association for Artificial Intelligence, pp. 105-110, 1995.

[3] J. Cheng, "A Strong Relevant Logic Model of Epistemic Processes in Scientific Discovery," Frontiers in Artificial Intelligence and Applications 61, pp. 136-159, 2000.

[4] J. Cheng, S. Nara, and Y. Goto, "FreeEnCal: A Forward Reasoning Engine with General-Purpose," Proceedings of 11th International Conference on Knowledge-Based Intelligent Information and Engineering Systems, LNCS, vol. 4693, Springer, Heidelberg, pp. 444-452, 2007.

[5] S. Colton, "Automated Theorem Discovery: A Future Direction for Theorem Provers," Proceedings of 1st Automated Reasoning: International Joint Conference, Workshop on Future Directions in Automated Reasoning, pp. 38-47, 2001.

[6] S. Colton, A. Meier, V. Sorge, and R. McCasland, "Automatic Generation of Classification Theorems for Finite Algebras," Automated Reasoning, LNCS, vol. 3097, Springer, Heidelberg, pp. 400-414, 2004.

[7] G. Dalzotto and T. Recio, "On Protocols for the Automated Discovery of Theorems in Elementary Geometry," Journal of Automated Reasoning 43.2, pp. 203-236, 2009.

[8] H. Gao and J. Cheng, "Automated Theorem Finding by Forward Reasoning Based on Strong Relevant Logic: A Case Study in Tarski's Geometry," Proceedings of International Conference on Advanced Multimedia and Ubiquitous Engineering - FutureTech & MUE," LNEE, vol. 393, Springer, pp. 55-61, 2016.

[9] H. Gao, and J. Cheng, "Measuring Interestingness of Theorems in Automated Theorem Finding by Forward Reasoning: A Case Study in Peano's Arithmetic," Proceedings of 9th Asian Conference on Intelligent Information and Database Systems, LNCS, vol. 10192, pp. 115-124, Springer, 2017.

[10] H. Gao, Y. Goto, and J. Cheng, "A Systematic Methodology for Automated Theorem Finding," Theoretical Computer Science 554, Elsevier, pp. 2-21, 2014.

[11] H. Gao, Y. Goto, and J. Cheng, "Research on Automated Theorem Finding: Current State and Future Directions," Proceedings of 9th FTRA International Conference, FutureTech 2014, LNEE, vol. 309, Springer, Heidelberg, pp 105-110, 2014.

[12] H. Gao, Y. Goto, and J. Cheng, "A Set of Metrics for Measuring Interestingness of Theorems in Automated Theorem Finding by Forward Reasoning: A Case Study in NBG Set Theory," Proceedings of 5th International Conference on Intelligence Science and Big Data Engineering, LNCS, vol. 9243, Springer, pp. 508-517, 2015.

[13] R. McCasland, A. Bundy, and S. Autexier, "Automated Discovery of Inductive Theorems," Journal of Studies in Logic, Grammar and Rhetoric 10.23, pp. 135-149, 2007.

[14] A. Montes, T. Recio, "Automatic Discovery of Geometry Theorems Using Minimal Canonical Comprehensive Grobner Systems," Proceedings of 6th International Workshop on Automated Deduction in Geometry, LNCS, vol. 4869, Springer Heidelberg, pp. 113-138, 2007.

[15] Y. Puzis, Y. Gao, and G. Sutcliffe, "Automated Generation of Interesting Theorems," Proceedings of 19th International Florida Artificial Intelligence Research Society Conference, AAAI press, pp. 49-54, 2006.

[16] A. Quaife, "Automated Development of Fundamental Mathematical Theories," Kluwer Academic, 1992.

[17] A. Tarski, and S. Givant, "Tarski's System of Geometry," Bulletin of Symbolic Logic 5.2, pp. 175-214, 1999

[18] L. Wos, "Automated Reasoning: 33 Basic Research Problem," Prentice-Hall, 1988.

[19] L. Wos, "The Problem of Automated Theorem Finding," Journal of Automated Reasoning, 10(1):137-138, 1993.