Communicability Criteria of Law Equations Discovery

Takashi Washio and Hiroshi Motoda

Institute of Scientific and Industrial Research, Osaka University,
8-1, Mihogaoka, Ibarakishi, Osaka, 567-0047, Japan
Phone: +81-6-6879-8541 Fax: +81-6-6879-8544
washio@sanken.osaka-u.ac.jp

Abstract. The "laws" in science are not the relations established by only the objective features of the nature. They have to be consistent with the assumptions and the operations commonly used in the study of scientists identifying these relations. Upon this consistency, they become communicable among the scientists. The objectives of this literature are to discuss a mathematical foundation of the communicability of the "scientific law equation" and to demonstrate "Smart Discovery System (SDS)" to discover the law equations based on the foundation. First, the studies of the scientific law equation discovery are briefly reviewed, and the need to introduce an important communicability criterion called "Mathematical Admissibility" is pointed out. Second, the axiomatic foundation of the mathematical admissibility in terms of measurement processes and quantity scale-types are discussed. Third, the strong constraints on the admissible formulae of the law equations are shown based on the criterion. Forth, the SDS is demonstrated to discover law equations by successively composing the relations that are derived from the criterion and the experimental data. Fifth, the generic criteria to discover communicable law equations for scientists are discussed in wider view, and the consideration of these criteria in the SDS is reviewed.

1 Introduction

Various relations among objects, events and/or quantity values are observed in natural and social behaviors. Especially, scientists call the relation as a "law" if it is commonly observed over the wide range of the behaviors in a domain. When the relation of the law can be represented in form of mathematical formulae constraining the values of some quantities characterizing the behaviors, the relation is called "law equations". In popular understanding, the relations of the laws and the law equations are considered to be objective in the sense that they are embedded in the behaviors independent of our processes of observation, experiment and interpretation. However, the definition of the laws and the law equations as communicable knowledge shared by scientists must be more carefully investigated. Indeed, they are not the relations established by only the objective features of the nature as discussed in this chapter. They have to be consistent with the assumptions and the operations commonly used in the study of scientists identifying these relations. Upon this consistency, they become communicable among the scientists.

On the other hand, the studies to develop automated or semi-automated systems to discover scientific law equations have been performed in the last two decades. As the main goal of the studies is to discover law equations representing meaningful relations among quantities for scientists, i.e., communicable with scientists, the systems must take into account the communicability criteria to some extent. The objectives of this chapter are to discuss a mathematical foundation of the communicability of the "scientific law equation" and to demonstrate "Smart Discovery System (SDS)" to discover the law equations based on the foundation. Through the demonstration of the scientific law equation discovery and the subsequent discussion, the communicability criteria of law equation discovery are clarified.

2 Study of Law Equation Discovery

First, we briefly review the past studies of the scientific law equation discovery from the view point of the equation formulae having the communicability in science. The most well known pioneering system to discover scientific law equations under the condition where some quantities are actively controlled in a laboratory experiment is BACON [13]. FAHRENHEIT [9] and ABACUS [6] are successors that basically use similar algorithms to BACON to discover law equations. LAGRANGE [5] and LAGRAMGE [19] are another type of scientific law equation discovery systems based on the ILP-like generate and test reasoning to discover equations representing the dynamics of the objects.

Many of these succeeding systems introduced the constraint of the unit dimension of physical quantities to prune the search space of the equation formulae. The constraint is called "Unit Dimensional Homogeneity" [2, 3] that all additive terms in a law equation formula must have an identical unit dimension. For example, a term having a length unit $[m]$ is not additive to another term having a different unit $[kg]$ in a law equation, even if the formula including their addition well fits to given data. Though the main purpose of the use of this constraint in these systems was to reduce the ambiguity in their results under noisy measurements and the high computational cost of their algorithms, the introduction also had an effect to increase the communicability of the discovered equations with scientists because the discovered equations are limited to more meaningful formulae. A law equation discovery system COPER [10] more intensively applied the constraints deduced from the unit dimensional analysis. The limitation of the constraints is so strong that some parts of the equation formulae are almost predetermined without using the measurement data set, and the derived equations has high communicability with scientists. LAGRANGE and LAGRAMGE are also capable of introducing these constraints in principle. However, the main purpose of these works is to provide an elegant measure to implement the constraints in the scientific law equation discovery but not to propose the contents of the constraints to enhance the communicability.

A strong limitation of the use of the unit dimensional constraints is its narrow applicability only to the quantities whose units are clearly known. To overcome this drawback, a law equation discovery system named "Smart Discovery System (SDS)" has been proposed [20, 21]. It discovers scientific law equations by limiting its search space to "Mathematically Admissible" equations in terms of the constraints of "scale-type" and "identity". They represent the important assumptions and operations commonly used in measurement and modeling processes identifying the relations among quantities by scientists. Since the use of scale-types and identity is not limited by the availability of the unit dimensions, SDS is applicable to non-physical domains including biology, sociology, and economics. In the following section, the axiomatic foundation of the mathematical admissibility in terms of measurement processes and quantity scale-types are discussed.

3 Scale-types of Quantities

"Mathematical Admissibility" includes the constraints of some fundamental notions in mathematics such as arithmetic operations, but they are very weak to constrain the shape of the law equation formulae. Stronger constraints are deduced from the assumptions and operations used in measurement process. The value of a quantity is obtained through a measurement in most of the scientific domains, and some features of the quantity are characterized by the measurement process. Though the unit dimension is an example of such features, a more generic feature is called "scale-types". S.S. Stevens defined that a measurement is to assign a value to each element in a set of the objects and/or events under given rules, and claimed that the rule set defines the "scale type" of the measured quantity. He categorized the scale-types into "nominal", "ordinal", "interval" and "ratio" scales [18]. In the later study, another scale-type called absolute scale is added. Subsequently, D.H. Krantz et al. axiomatized the measurement processes and the associated scale-types [12]. In this section, their theory on the scale-types is reviewed.

Definition 1 (A Relation System) *The following series of finite length α is called "a relation system".*

$$\alpha =< A, R_1, R_2, ..., R_n >$$

where A: a non-empty set of elements, R_i: $R_i(a_1, a_2, ..., a_{mi})$ which is a relation of $a_1, a_2, ..., a_{m_i} \in A$.

Definition 2 (Type and Similarity) *Given a relation system α, when each R_i is the relation of m_i elements in A, the series of positive integers $< m_1, m_2, ...,$ $m_n >$ is called "type" of α. If the types of two relation systems α and $\beta =<$ $B, S_1, S_2, ..., S_n >$ are identical, they are "similar".*

Definition 3 (Isomorphism (Homomorphism)) *Given two relation systems α and β, if the following conditions are met, they are called "isomorphic (homomorphic)".*

(1) α and β are similar.
(2) A bijection (surjection) f from A to B exists where
$R_i(a_1, a_2, ..., a_{m_i}) \Leftrightarrow S_i(f(a_1), f(a_2), ..., f(a_{m_i}))$.

Definition 4 (Numerical and Empirical Relation Systems) *When a relation system α satisfies the following conditions, α is called a "numerical relation system".*

(1) The domain $A \subseteq \Re$.
(2) $R_i (i = 1, 2, ..., n)$ is a relation among values in $\Re$.

A relation system which is not a numerical relation system is called an "empirical system".

Next, the "extensional measurement" and the "scale-types" are defined.

Definition 5 (Extensional Measurement) *"Extensional measurement" is that an empirical relation system consisting of some operations and relations among elements in the objective behaviors is isomorphic or homomorphic with a numerical relation system consisting of arbitrary chosen operations and relations among numerals and/or symbols.*

Extensional measurement is the measurement done by using a facility to directly map each element in an empirical relation system to a numeral in a numerical relation system while preserving the relations in each system. For example, the measurement of length and the measurement of weight are the extensional measurements respectively since a ruler and a balance map the lengths and the weights of objects to numerals directly.

Definition 6 (Scale-type in Extensional Measurement) $< \alpha, \beta_f, f >$ *is called a "scale-type" where*

α: an empirical relation system,
β_f: a full numerical relation system,
f : isomorphic or homomorphic mapping from α to the subsystem of β_f.

Here, the "full numerical relation system" is the system which domain is the entire $\Re$, and the "subsystem" is the system where its domain is the sub-domain of the original system, and all relations of the subsystem have one-to-one correspondence to the relations of the original system.

When an empirical relation system $\alpha =< A, I >$ is the following classification system, the measurement by "nominal scale" is applicable.

Definition 7 (Classification System) *Given $\alpha =< A, I >$, if I is a binary relation on A, α is called a "binary system". Furthermore, if the following three axioms hold for I, I is called an "equivalence relation", α a "classification system" and the set of elements where I holds "I-equivalence class".*

Reflexive law: $\forall a \in A, I(a, a)$
Symmetric law: $\forall a, \forall b \in A, I(a, b) \Rightarrow I(b, a)$
Transitive law: $\forall a, \forall b, \forall c \in A, I(a, b) \wedge I(b, c) \Rightarrow I(a, c)$

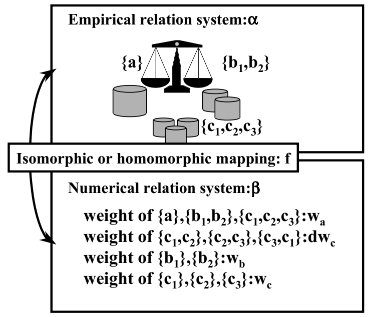

Fig. 1. Example of extensional measurement.

An example of the classification system and the measurement in the nominal scale is explained through the empirical relation system α depicted in Fig. 1. The domain A is the power set of the set of 6 weights $\{a, b_1, b_2, c_1, c_2, c_3\}$. The equivalence relation I is that two sets of the weights are balanced. The reflexive law holds since two identical weight sets balance. The symmetric law also holds because the balance of the following pair wise sets is invariant for the exchange of their positions between the left dish and the right dish.

$$\begin{aligned}&(\{b_1\}, \{b_2\}), (\{c_1\}, \{c_2\}), (\{c_2\}, \{c_3\}), (\{c_3\}, \{c_1\}),\\ &(\{c_1, c_2\}, \{c_2, c_3\}), (\{c_2, c_3\}, \{c_3, c_1\}), (\{c_3, c_1\}, \{c_1, c_2\}),\\ &(\{a\}, \{b_1, b_2\}), (\{b_1, b_2\}, \{c_1, c_2, c_3\}), (\{c_1, c_2, c_3\}, \{a\})\end{aligned}$$

In addition, if a pair of sets in the following combinations balance, then the rest pairs also balance. Thus, the transitive law holds.

$$(\{c_1\}, \{c_2\}, \{c_3\}), (\{c_1, c_2\}, \{c_2, c_3\}, \{c_3, c_1\}), (\{a\}, \{b_1, b_2\}, \{c_1, c_2, c_3\})$$

Accordingly, this empirical relation system α is a classification system. For this α, given a numerical relation system β and its domain $B \subseteq \Re$, a surjection f which maps any weight sets in I-equivalence class to an identical number on B is introduced as follows.

$$\begin{aligned}&w_a = f(\{a\}) = f(\{b_1, b_2\}) = f(\{c_1, c_2, c_3\}),\\ &w_b = f(\{b_1\}) = f(\{b_2\}),\\ &w_c = f(\{c_1\}) = f(\{c_2\}) = f(\{c_3\}),\\ &dw_c = f(\{c_1, c_2\}) = f(\{c_1, c_3\}) = f(\{c_2, c_3\})\\ &\text{where } w_a, w_b, w_c, dw_c \in B.\end{aligned}$$

If a relation of β, which is the equality of two numbers, is considered, the reflexive law holds as the equality of identical numbers is trivial. Also, the symmetric law and the transitive law hold for the equality of w_a, w_b, w_c, dw_c. Hence, β is a classification system. Therefore, the homomorphic mapping f to assign an identical number to balanced weights is a measurement in nominal scale.

When an empirical relation system $\alpha =< A, P >$ is the following series, the measurement in "ordinal scale" is possible.

Definition 8 (Series) *Given a binary system $\alpha =< A, P >$, if the following three axioms hold for P, P is called "inequivalence relation" and α "series".*

Asymmetric law: $\forall a, \forall b \in A, P(a, b) \Rightarrow \neg P(b, a)$

Transitive law: $\forall a, \forall b, \forall c \in A, P(a, b) \wedge P(b, c) \Rightarrow P(a, c)$

Law of the excluded middle: $\forall a, \forall b \in A,$ *one of* $P(a, b)$ *and* $P(b, a)$ *holds.*

In case that the relation I holds on some of the elements in A, *i.e.*, $\alpha =< A, I, P >$, the elements in each I-equivalence class are grouped, and A is replaced by the "quotient set" A/I. Then, $\alpha/I =< A/I, P >$ is a series, and it can be measured in ordinal scale. In the example of Fig. 1, the domain of $\alpha/I =< A/I, P >$ is A/I where the sets of the weights which mutually balance are grouped. Then, given two elements r and s in A/I, the binary relation $P(r, s)$ is defined that the dish on which s is put comes down. This $P(r, s)$ satisfies the conditions of the aforementioned series. On the other hand, given the binary relation $P(w_r, w_s)$ which is inequality $w_r < w_s$ between two numbers in the domain B of a numerical relation system β. This also satisfies the conditions of the series. Then we define a surjection f which assigns real numbers w_a, w_b, w_c, dw_c to the elements in A/I respectively where $w_r < w_s$ holds for r and s under $P(r, s)$. This definition of f which holds $w_c < w_b < dw_c < w_a$ is the measurement in ordinal scale.

Furthermore, when an empirical relation system $\alpha =< A, D >$ is the following "difference system", the measurement in "interval scale" is possible.

Definition 9 (Difference System) *Given $\alpha =< A, D >$, if the relation D is a quadruple relation on A, α is called a "quadruple system". Moreover, α is called a "difference system" if the following axioms holds for $\{a, b, c, d, e, f\} \subseteq A$.*

$P(a, b) \not\Leftrightarrow D(a, b, a, a),$
$I(a, b) \Leftrightarrow D(a, b, b, a) \wedge D(b, a, a, b),$
$D(a, b, c, d) \wedge D(c, d, e, f) \Rightarrow D(a, b, e, f),$
One of $D(a, b, c, d)$ *and* $D(c, d, a, b)$ *holds,*
$D(a, b, c, d) \Rightarrow D(a, c, b, d),$
$D(a, b, c, d) \Rightarrow D(d, c, b, a),$
$\exists c \in A, D(a, c, c, b) \wedge D(c, b, a, c),$
$P(a, b) \wedge \neg D(a, b, c, d) \Rightarrow \exists e \in A, P(a, e) \wedge P(e, b) \wedge D(c, d, a, e),$
$\exists e, \exists f \in A, \exists$ *an integer* $n, P(a, b) \wedge D(a, b, c, d) \Rightarrow M_n(c, e, f, d).$

Here, M_n is the relation to locate e and f between c and d in A where the distance between c and e and that between f and d are identical and one n-th of the distance between c and f. Even if some elements in A satisfy the equivalence relation I, $\alpha/I =< A/I, D >$ is a difference system, and α/I can be measured by an interval scale quantity. In the example of Fig. 1, let the relation $D(r, s, t, u)$ on A/I be that the left dish comes down when two sets of weights r and u are put on the left dish and s and t on the right dish. Then α/I is a difference system. Let $D(w_r, w_s, w_t, w_u)$ on the domain B of a numerical relation system β be $(w_s - w_r) \leq (w_u - w_t)$, and let a surjection f be the assignment of numerals w_a, w_b, w_c, dw_c to the sets of weights in such a way that $D(w_r, w_s, w_t, w_u)$ holds in β when $D(r, s, t, u)$ holds in α/I. In this example, $w_a = w_c + 4(w_b - w_c)$ and $dw_c = w_c + 2(w_b - w_c)$ are obtained, and the numerals mapped by the surjection f are interval scale. The f which satisfies this relation is not unique. The different mappings f_1 and f_2 for two numerical relation systems β_1 and β_2 which are homomorphic with α/I respectively have a linear relation $f_2 = k \cdot f_1 + c$ where k and c are constants, and this is the admissible unit conversion. The interval scale quantities follow the axioms of the classification system, the series and the difference system, but do not have any absolute origins. The examples are position, time and musical sound pitch since the origins of their coordinate systems are arbitrarily introduced.

The quantities of ratio scale are derived by the extension of the difference system. Given two difference systems α/I and β, define a surjection f from $A/I \times A/I$ to $B \times B$ satisfying $f(r, s) = w_s - w_r$ and $f(\phi, r) = w_r$. Under this mapping, α/I is measured by a ratio scale quantity. In the example of Fig. 1, the two weights c_1 on the left dish and the weights c_2 and c_3 on the right dish balance in α/I, and this is homomorphic with the following relation in β.

$$f(\phi, \{c_1\}) = f(\{c_1\}, \{c_2, c_3\}),$$

where $f(\phi, \{c_1\}) = w_c$ and $f(\{c_1\}, \{c_2, c_3\}) = dw_c - w_c$. This deduces the relation $dw_c = 2w_c$. By substituting this relation to the aforementioned $w_a = w_c + 4(w_b - w_c)$ and $dw_c = w_c + 2(w_b - w_c)$, $2w_b = 3w_c$ and $w_a = 3w_c$ are deduced, and the ratio scale of weight is derived. f satisfying these relations are not unique. Given two numerical relation systems β_1 and β_2 which are homomorphic with α/I, the corresponding f_1 and f_2 have a similarity relation $f_2 = k \cdot f_1$, *i.e.*, the admissible unit conversion. The ratio scale quantities have absolute origins. The examples are distance, elapsed time and physical mass.

Besides the quantities defined in the extensional measurement, another sort of quantities which can not be directly measured by any facilities but indirectly measured by functions of the other quantities exit. The process of this indirect measurement is called "intentional measurement", and the quantities measured through this measurement are "derivative quantities" obtained from the other quantities. The nature of each scale-type in the intentional measurement is very similar to that of the extensional measurement, and the admissible unit conversion is identical for each scale-type. The descriptions on the rigorous definitions of this measurement and its scale-types are omitted due to the space limitation. An example of the derivative quantity obtained through the intentional measurement is the temperature. It can be measured only through some other measured quantities such as the expansion length of mercury. The other representative derivative quantities are density, energy and entropy.

An important scale-type which is defined by the intentional measurement is absolute scale-type. Given a quantity g defined by the other ratio scale quantities

Table 1. Admissible relations between two quantities.

No.	scale-types independent quantity x	dependent quantity $y(x)$	admissible relation
1	ratio	ratio	$y(x) = \alpha x^{\beta}$
2.1	ratio	interval	$y(x) = \alpha x^{\beta} + \delta$
2.2			$y(x) = \alpha \log x + \beta$
3	interval	ratio	impossible
4	interval	interval	$y(x) = \alpha x + \beta$

$f_1, f_2, ..., f_n$ through $g = \prod_{i=1}^{n} f_i^{\gamma_i}$, when the relation $\prod_{i=1}^{n} k_i^{\gamma_i} = 1$ holds under any unit conversions of $f_1, f_2, ..., f_n$, the scale-type of g is called "absolute scale". Because the value of g is invariant for any unit conversions, it is uniquely defined and called "dimensionless number". Its admissible unit conversion follows the identity group $g'_2 = g_1$. The examples are the ratio of two masses and angle in radian.

4 Admissible Formulae of Law Equations

In this section, we review some important theorems on the relations among observed quantities, and show their extension for the discovery of law equations as communicable knowledge.

R.D. Luce claimed that the group structure of each scale-type is conserved through the unit transformation, and this fact strongly limits the mathematically admissible relations among quantities having interval and ratio scale-types [14]. For example, when x and y are ratio scale, the admissible unit conversions are $x' = kx$ and $y' = Ky$ respectively. When we assume the relation between x and y to be $y = \log x$, and apply a unit conversion on x, then the unit of y should be also converted as $y' = \log x' = \log kx = \log x + \log k$. This consequence that the origin of y is changed is contradictory to the above admissible unit conversions of y. Thus, the logarithmic relation between two ratio scale quantities is not admissible. R.D. Luce further proceeded this discussion, and derived the admissible binary relations between ratio and interval scale quantities depicted in Table 1.

On the other hand, an important theorem called "Product Theorem" on the relation formula among multiple measured quantities had been presented in the unit dimensional analysis which was independently studied by old scientists [2]. However, this theorem addresses on the relation among ratio scale quantities only. We derived the following "Extended Product Theorem" [20] to the case where the quantities of ratio, interval and absolute scales are included in the formula by introducing the consequences of R.D. Luce.

Theorem 1 (Extended Product Theorem). *Given a set of ratio scale quantities R and a set of interval scale quantities I, a derivative quantity Π is related* *with each $x_i \in R \cup I$ through one of the following formulae.*

$$\Pi = (\prod_{x_i \in R} |x_i|^{a_i})(\prod_{I_k \in C} (\sum_{x_j \in I_k} b_{kj}|x_j| + c_k)^{a_k}),$$

$$\Pi = \sum_{x_i \in R} a_i \log |x_i| + \sum_{I_k \in C_{\bar{g}}} a_k \log(\sum_{x_j \in I_k} b_{kj}|x_j| + c_k) + \sum_{x_\ell \in I_g} b_{g\ell}|x_\ell| + c_g,$$

where R or I can be empty, and C is a covering of I, $C_{\bar{g}}$ a covering of $I - I_g$ ($I_g \subseteq I$). Π can be any of interval, ratio and absolute scale, and each coefficient is constant.

Here, a "covering" C of a set I is a set of finite subsets Is_is of I where $I = \cup_i Is_i$. The same definition applies to $C_{\bar{g}}$ for $I - I_g$. When the argument quantities appearing in a law equation are ratio scale and/or interval scale, the relation among the quantities sharing arbitrary unit dimensions has one of the above formulae.

Another major theorem called "Buckingham Π-theorem" on the structure of a law equation consisting of ratio scale quantities only had also been presented in the old work in the unit dimensional analysis [3]. We further extended this theorem to include the interval, ratio and absolute scale quantities in the argument [20].

Theorem 2 (Extended Buckingham Π-theorem). *Given a complete equation $\phi(x, y, z, ...) = 0$, if every argument of this equation is either of interval, ratio and absolute scales, then the equation can be rewritten in the following form.*

$$F(\Pi_1, \Pi_2, ..., \Pi_{n-r-s}) = 0$$

where n is the number of the arguments of ϕ, r and s are the numbers of the basic unit and the basic origin contained in $x, y, ...$, and Π_i is absolute scale for all i and represented by the formulae of the regime defined by Extended Product Theorem.

Here, the basic unit is the unit dimension which defines the scaling independent of the other unit in ϕ as length $[L]$, mass $[M]$ and time $[T]$, and the basic origin is the origin which is artificially chosen in the measurement of an interval scale quantity, for example, the origin of temperature in Celsius defined as the melting point of water under the standard atmosphere pressure. Each $\Pi_i = \rho_i(x, y, ...)$ defining Π_i is called a "*regime*" and $F(\Pi_1, \Pi_2, ..., \Pi_{n-r}) = 0$ an "*ensemble*". Because all arguments of $F = 0$ are absolute scale, *i.e.*, dimensionless, the shape of the formula does not constrained by the theorem 1, and the arbitrary formula is admissible for $F = 0$ in terms of the scale-type.

The following example of the nuclear decay of a radioactive element is an example of the theorem 1 and the theorem 2.

$$N = N_0 \exp[-\lambda(t - t_0)] \tag{1}$$

where $t[s]$: time, $t_0[s]$: time origin, $\lambda[s^{-1}]$: decay speed constant,
$N[kg]$: current element mass, $N_0[kg]$: $t_0$original element mass

t and t_0 are interval scale, and λ, N and N_0 are ratio scale. By introducing dimensionless Π_1 and Π_2, the equation can be rewritten as

$$\Pi_1 = \exp(-\Pi_2), \quad (2)$$
$$\Pi_1 = N/N_0, \quad (3)$$
$$\Pi_2 = \lambda(t - t_0), \quad (4)$$

which are an *ensemble* and two *regime*s. The *regime*s (3) and (4) follow the first formula in the theorem 1. The number of the original arguments n is 5. r is equal to 2 because t, t_0 and λ share a basic unit of time $[s]$ and N and N_0 share the basic unit of mass $[kg]$. s is equal to 1 since t and t_0 share a basic origin of time. Thus $n - r - s = 2$ holds, and this satisfies the theorem 2. As indicated in the above example, the scale-type of measurement quantities strongly constrains the formulae of the law equations which are communicable among scientists. Empirical equations which relate the measurement quantities in arbitrary formulae do not provide excellent knowledge representation for the understanding and the communication among domain experts.

5 Algorithm of Smart Discovery System (SDS)

In this section, an algorithm of our "Smart Discovery System (SDS)" to discover a law equation based on the mathematical admissibility and the experiments on the objective behaviors is explained. An important point to perform these procedures is to establish a method to check if an equation holds for all behaviors which can be occurred in the experiments on the objective behaviors. A natural approach is to collect all possible combinations of the values of the controllable quantities in experiments and to fit the various candidate equations to the collected data. However, this generate and test approach faces the combinatorial explosion in the data collection and the candidate equation generation. To avoid this difficulty, we introduce the following assumptions.

(a) The objective behaviors are represented by a complete equation, and all quantities except one dependent quantity are controllable at least.
(b) The objective behaviors are static, or the time derivatives of some quantities are directly observable if the behaviors are dynamic.
(c) Given a pair of any quantities observed in the objective behaviors, the bi-variate relation on the pair can be identified while fixing the values of the other quantities in experiments.

5.1 Discovery of regime equations

bi-variate fitting: If the objective behaviors and the experimental conditions satisfy these assumptions, "*bi-variate fitting*" which searches a pair wise relation of two observed quantities can be applied to reduce the data for the search. In addition, the mathematical admissibility criterion on the scale-type is used to limit the equation formula to be fitted to the observed data. Initially, for a pair of interval scale quantities $\{x_i, x_j\}$, a linear relation

$$b_{ij}x_i + x_j = d_{ij}$$

is searched in the fitting based on the constraints in the table 1 where b_{ij} should be a constant coefficient. For a pair of ratio scale quantities, a power relation

$${x_i}^{a_{ij}} x_j = d_{ij}$$

is searched. In case of a pair of an interval scale x_i and ratio scale x_j, the following two candidate relations are searched.

$$b_{ij}x_i + x_j^{a_{ij}} = d_{ij},\ b_{ij}x_i + \log x_j = d_{ij}.$$

The goodness of fitting is checked in every fitting by the statistical F-test [1]. The same experiments are repeated $m = 10$ times. Then the bi-variate fittings to the data obtained in each experiment are conducted to check the reproducibility of the coefficient b_{ij}, *i.e.*, its constancy, through χ^2-test, and the effect of noise and error on their values are reduced by averaging the coefficients over $m = 10$ results. By applying these fitting to every pair of quantities in the data, all bi-variate relations satisfying the constraints in the table 1 are identified. The mathematical complexity of the bi-variate fitting is $O(mn^2)$ where n is the total number of the quantities in the given data.

triplet test: In the next step, the mutually consistent bi-variate relations are composed to multiple regime equations shown in the theorem 1. Each regime equation is composed in bottom up manner which searches the equation relating less number of quantities in the data. The consistent composition is made through the following "triplet test". The consistency among the values of the constant coefficients in a triplet of the bi-variate relations for three observed quantities is checked under the assumption of a linear relation among the interval scale quantities as indicated in the theorem 1.

For example, given a set of three interval scale quantities $\{x_i, x_j, x_k\}$, if the following three bi-variate relations among them are mutually consistent,

$$b_{ij}x_i + x_j = d_{ij},\ b_{jk}x_j + x_k = d_{jk},\ b_{ki}x_k + x_i = d_{ki}$$

the following relation holds among the coefficients.

$$1 = b_{ij}b_{jk}b_{ki}.$$

This condition can be tested by the normal distribution test considering the error bounds of the coefficients. The error bounds of b_{ij}, b_{jk} and b_{ki}, *i.e.*, $\Delta b_{ij}, \Delta b_{jk}$ and Δb_{ki}, can be statistically evaluated based on the errors of the m least square fittings of each relation. Then the total error bound Δb_{rhs} of the right hand side of the above relation is derived by the following formula of error propagation.

$$\Delta b_{rhs} = \sqrt{(b_{jk}b_{ki}\Delta b_{ij})^2 + (b_{ij}b_{ki}\Delta b_{jk})^2 + (b_{ij}b_{jk}\Delta b_{ki})^2}$$

This standard deviation error bound is used to judge if the value of the product of the three coefficients are sufficiently close to 1 under the normal distribution test.

The principle of this test can be applied to the other triplets containing of ratio and/or interval scale quantities. If the consistency is confirmed, they can be merged into a relation. In the above example, they are merged into

$$x_i + b_{jk}b_{ki}x_j + b_{ki}x_k = \pi_{ijk},$$

where π_{ijk} is an intermediate derivative quantities composed by b_{jk} and b_{ki} which are known to be dependent of the other quantities, i.e., constants.

This procedure is continued for another quantity x_ℓ and any two quantities in $\{x_i, x_j, x_k\}$. If every triplet among the bi-variate relations of $\{x_h, x_i, x_j, x_\ell\}$ is consistent, they can be merged to a relation among the four quantities since all constant coefficients in a linear formulae are mutually consistent. In this case, the following linear relation is obtained.

$$x_i + b_{j\ell}b_{\ell j}x_j + b_{k\ell}b_{\ell k}x_k + b_{\ell i}x_\ell = \pi_{ijk\ell},$$

This procedure further repeated until no larger sets of quantities having consistency are found. This is similar to the generalization of bi-variate relations to multi-variate relations in BACON [13]. However, the computational complexity of the triplet test $O(n^3)$ is lower than the conventional approach. This is because of the use of the mathematical admissibility constraints and the systematic triplet consistency test. Through this procedure, the set of regime equations relating the many original quantities with less number of dimensionless quantities $\{\Pi_i | i = 1, ..., n - r - s\}$ can be discovered, and this efficiently reduces the computational cost for the discovery of a complete law equation.

5.2 Discovery of ensemble equation

term merge: Once all regimes to define Π_is are discovered, an ensemble equation among Π_is is searched. Because the ensemble equation does not follow the scale-type constraints, it can take any arbitrary formula. Accordingly, we introduce an assumption that the ensemble equation consists of only the arithmetic operators and elementary functions among Π_is to limit the search space of the formula. The most of the law equations follows this assumption, and it is widely used in the other equation discovery approaches [6].

In our approach, a set CE of candidate binary relations such as addition, multiplication, linear, exponential and logarithmic relations is given. Then by the technique of the bi-variate fitting, each relation in CE is applied to the data of Π_is calculated by the regime equations. For example, the following bi-variate product form and linear form are applied.

$$\Pi_i^{a_{ij}}\Pi_j = b_{ij} \text{ (product form) and}, a_{ij}\Pi_i + \Pi_j = b_{ij}, \text{ (linear form)}.$$

First, the former product form is adopted to the least square fitting to every pair of Π_i and $\Pi_j (i, j = 1, ..., n - r - s)$. Then, the statistical F-tests mentioned earlier are applied.

This process is repeated over the $k = 10$ different data sets obtained in the random experiments. The bi-variate equations passed all these tests are stored, and the invariance of the exponent a_{ij} of each bi-variate relation against the value changes of any other quantities are checked by examining the $k = 10$ values of a_{ij} obtained in the experiments through χ^2-test. If a_{ij} is invariant, we observe a high possibility that a_{ij} is a constant characterizing the nature of the objective system within the scope of the experiment. The relations having the invariant a_{ij}s are marked, and every maximal convex set MCS of quantities is searched where all pairs of quantities in MCS are related by the bi-variate relations marked as having the invariant a_{ij}. Then the quantities in every MCS are merged into the following term.

$$\Theta_i = \prod_{x_j \in MCS_i} {x_j}^{a_j}.$$

Similar procedure is applied to the linear bi-variate form, in which case the merged term of an MCS is as follows.

$$\Theta_i = \sum_{x_j \in MCS_i} a_j x_j.$$

This procedure is recursively repeated for all bi-variate relations in CE among Π_is and Θ_is until no new term becomes available. A Θ_i is a unique derivative term in each relation which is dependent of the values of the other Π_is outside the relation.

identity constraint: If all terms are merged into one in the above term merge process, the relation is the ensemble equation. Otherwise the following procedure to merge the Θ_is further continues by applying an extra mathematical constraint based on the "identity" of the relations. The basic principle of the identity constraints comes by answering the question that "what is the relation among Θ_h, Θ_i and Θ_j, if $\Theta_i = f_{\Theta_j}(\Theta_h)$ and $\Theta_j = f_{\Theta_i}(\Theta_h)$ are known?" For example, if $a(\Theta_j)\Theta_h + \Theta_i = b(\Theta_j)$ and $a(\Theta_i)\Theta_h + \Theta_j = b(\Theta_i)$ are given, the following identity equation is obtained by solving each for Θ_h.

$$\Theta_h \equiv -\frac{\Theta_i}{a(\Theta_j)} + \frac{b(\Theta_j)}{a(\Theta_j)} \equiv -\frac{\Theta_j}{a(\Theta_i)} + \frac{b(\Theta_i)}{a(\Theta_i)}$$

Because the third expression is linear with Θ_j for any Θ_i, the second must be so. Accordingly, the following must hold.

$$1/a(\Theta_j) = \alpha_1\Theta_j + \beta_1,\ b(\Theta_j)/a(\Theta_j) = -\alpha_2\Theta_j - \beta_2.$$

By substituting these to the second expression,

$$\Theta_h + \alpha_1\Theta_i\Theta_j + \beta_1\Theta_i + \alpha_2\Theta_j + \beta_2 = 0$$

is obtained.

This principle is generalized to various relations among multiple terms. Table 2 shows such relations for multiple linear relations and multiple product relations. The relation is used to fit to the data and to merge Θ_is further into another new term Θ which is a coefficient of the relation dependent of the values of the other Θ_is outside the relation. Similarly to the bi-variate fitting, the goodness of fitting is checked by the statistical F-test. These merging operations are repeated until a complete ensemble equation among the terms is obtained where all coefficients are constant in a relation.

Table 2. Identity constraints

bi-variate relation	general relation
$ax + y = b$	$\sum_{(A_i \in 2^{LQ})\&(p \not\subseteq A_i \forall p \in L)} a_i \prod_{x_j \in A_i} x_j = 0$
$x^a y = b$	$\prod_{(A_i \in 2^{PQ})\&(p \not\subseteq A_i \forall p \in P)} \exp(a_i \prod_{x_j \in A_i} \log x_j) = 0$

L is a set of pair wise terms having a bi-variate linear relation and $LQ = \cup_{p \in L} p$. P is a set of pair wise terms having a bi-variate product relation and $PQ = \cup_{p \in P} p$.

6 Application to Law Equation Discovery

6.1 Discovery of law based models

The aforementioned principles have been implemented to "Smart Discovery System (SDS)" [20]. SDS receives the data and the scale-type information of the quantities observed in model simulations, and tries to discover a complete law equation governing the simulation without knowing the model.

First, the application of SDS to a circuit depicted in Fig. 2 is demonstrated. This is a circuit of photometer to measure the rate of increase of photo intensity within a certain time period. This is represented by the following complete equation containing 18 quantities.

$$(\frac{R_3 h_{fe_2}}{R_3 h_{fe_2} + h_{ie_2}} \frac{R_2 h_{fe_1}}{R_2 h_{fe_1} + h_{ie_1}} \frac{rL^2}{rL^2 + R_1})(V_1 - V_0) - \frac{Q}{C} - \frac{K h_{ie_3} X}{B h_{fe_3}} = 0 \qquad (5)$$

Here, L and r are photo intensity and sensitivity of the Csd device which is one of popular optical sensors. X, K and B are the position of indicator, spring constant and the intensity of magnetic field of the current meter respectively. h_{ie_i} is the input impedance of the base of the i-th transistor. h_{fe_i} is the gain ratio of the currents at the base and the collector of the i-th transistor. The definitions of the other quantities follow the standard symbolic representations in the electric circuit domain.

The electric voltage levels V_1 and V_2 are interval scale and h_{fe_i}s absolute scale. Thus, the set of interval scale quantities is $IQ = \{V_1, V_2\}$, that of ratio scale quantities $RQ = \{L, r, R_1, R_2, R_3, h_{ie_1}, h_{ie_2}, h_{ie_3}, Q, C, X, K, B\}$ and that of absolute scale quantities $AQ = \{h_{fe_1}, h_{fe_2}, h_{fe_3}\}$. In the following equation fitting, the value of each coefficient is rounded into the nearest integer or the nearest inverse of integer, if the value is close enough to it within the error bound. This is due to the empirical observation that the coefficients are often the integers or their inverses in a law equation.

Initially, the bi-variate fitting was applied to IQ, and a binary relation $\Pi_1 = V_1 - V_0$ was obtained. Since IQ includes only the two quantities, the search for Πs in IQ was stopped. In the next step, the bi-variate fitting was applied to the quantities in RQ and Π_1. Because the basic origin of the voltage level has been cancel out between V_1 and V_0, Π_1 became a ratio scale quantity. The resultant binary relations were as follows.

$$L^2 r = b_1, L^{-2} R_1 = b_2, r^{-1} R_1 = b_3, R_2^{-1} h_{ie_1} = b_4, R_3^{-1} h_{ie_2} = b_5, Q^{-1} C = b_6,$$
$$h_{ie_3} X = b_7, h_{ie_3} K = b_8, h_{ie_3}^{-1} B = b_9, XK = b_{10}, X^{-1} B = b_{11}, K^{-1} B = b_{12}$$

Subsequently, the triplet tests were applied to these relations, and the following regime equations were obtained.

$$\Pi_1 = V_1 - V_0, \Pi_2 = R_1 r^{-1.0} L^{-2.0}, \Pi_3 = h_{ie_1} R_2^{-1.0},$$
$$\Pi_4 = h_{ie_2} R_3^{-1.0}, \Pi_5 = h_{ie_3} XKB^{-1.0}, \Pi_6 = QC^{-1.0}$$

Then, the merge of these Πs and the quantities in AQ was performed by applying the binary relations in CE, and the following new terms were derived.

$$\begin{aligned}
&\Theta_1 = \Pi_2 h_{fe_1} = R_1 r^{-1.0} L^{-2.0} h_{fe_1},\\
&\Theta_2 = \Pi_3 h_{fe_2} = h_{ie_1} R_2^{-1.0} h_{fe_2},\\
&\Theta_3 = \Pi_4 h_{fe_3} = h_{ie_2} R_3^{-1.0} h_{fe_3},\\
&\Theta_4 = \Pi_5 + \Pi_6 = h_{ie_3} XKB^{-1.0} + QC^{-1.0},\\
&\Theta_5 = \Pi_1 \Theta_4^{-1.0} = (V_1 - V_0)(h_{ie_3} XKB^{-1.0} + QC^{-1.0})^{-1.0}
\end{aligned}$$

Thus, the quantities were merged into five terms $\{\Theta_1, \Theta_2, \Theta_3, \Theta_5\}$.

Furthermore, the identity constraint was applied to these terms since the binary linear relations were found in the combinations of $\{\Theta_1, \Theta_5\}, \{\Theta_2, \Theta_5\}$ and $\{\Theta_3, \Theta_5\}$. This derived the following multi-linear formula.

$$\Theta_1\Theta_2\Theta_3 + \Theta_1\Theta_2 + \Theta_2\Theta_3 + \Theta_1\Theta_3 + \Theta_1 + \Theta_2 + \Theta_3 + \Theta_5 + 1 = 0$$

Because every coefficient is independent of any terms, this is considered to be the ensemble equation. The equivalence of this result to Eq.(5) is easily checked by substituting the intermediate terms to this ensemble equation.

SDS has been also applied to non-physics domain. For example, given a sound frequency f and a musical sound pitch I where the former is ratio scale and the

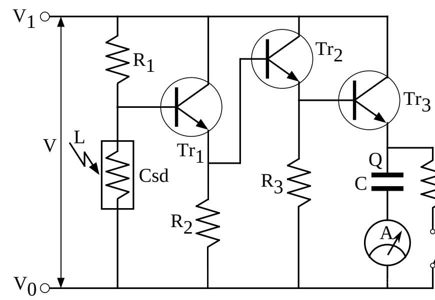

Fig. 2. A circuit of photometer.

latter interval scale, the following two candidate relations have been derived by SDS.

$$I = \alpha f^{\beta} + \gamma, \text{ or } I = \alpha \log f + \beta$$

Because both equations show similar accuracy, and the latter contains less parameters, SDS prefers the latter by following the criterion of parsimony which will be discussed later. This equation has been called "*Fechner's Law*" in psychophysics. Another example is the law of spaciousness of a room in psychophysics [8].

$$S_p = c \sum_{i=1}^{n} R L_i^{0.3} W_i^{0.3},$$

where S_p, R, L_i and W_i are average spaciousness of a room, room capacity, light intensity and solid angle of window at the location i in the room. Though the unit dimension of S_p is unclear, its scale-type is known to be ratio scale since it was evaluated through the method of magnitude estimation which is a popular method to derive a ratio scale quantity in psychophysics. L and R are ratio scale, and W is absolute scale. SDS easily obtained the above expression.

6.2 Basic Performance of SDS

Table 3 shows the performance of SDS to discover various physical law equations. The relative CPU time of SDS normalized by the first case shows that its computational time is nearly proportional to n^2. For reference, the relative CPU time of ABACUS is indicated for the same cases except for the circuit examples of this paper [6]. Though ABACUS applies various heuristics including the information of unit dimension, its computational time is non-polynomial, and it could not derive the law equations for the complicated circuits within a tractable time.

The robustness of SDS against the noisy experimental environment has been also evaluated. The upper limitation of the noise level to obtain the correct result in the cases of more than 80% of 10 trials was investigated for each physical law,

Table 3. Performance of SDS and ABACUS in reconstructing physical laws.

Example	n	TC(S)	TC(A)	NL(S)
Ideal Gas	4	1.00	1.00	±40%
Momentum	8	6.14	22.7	±35%
Coulomb	5	1.63	24.7	±35%
Stoke's	5	1.59	16.3	±35%
Kinetic Energy	8	6.19	285.	±30%
Circuit*1	17	21.6	-	±20%
Circuit*2	18	21.9	-	±20%

n: Number of Quantities, TC(S): Total CPU time of SDS, TC(A): Total CPU Time of ABACUS, NL(S): Limitation of Noise Level of SDS, *1: Case that electronic voltage is represented by a ratio scale V, *2: Case that electronic voltage is represented by two interval scale V_0 and V_1.

and they are indicated in the last column of Table 3. The noise levels shown here are the standard deviation of Gaussian noise relative to the real values of quantities, and were added to both controlled (input) quantities and measured (output) quantities at the same time. Thus actual noise level is higher than these levels. The results show the significant robustness of SDS. This is due to the bottom up approach of the bi-variate fitting where the fitting is generally robust because of its simplicity. SDS can provide appropriate results under any practical noise condition.

As shown in the above results, SDS can discover quite complex law based models containing more than 10 quantities under practical conditions. As the modeling of the objective behaviors represented by many quantities is a difficult and time consuming task for scientists and engineers, the approach presented in this chapter provide a significant advantage.

7 Generic Criteria to Discover Communicable Law Equations

As we have seen in the previous sections, the "Mathematical Admissibility" plays an important role to discover the law equations as communicable knowledge shared by scientists, since it is based on the assumptions and the operations commonly used in the study of scientists. However, this is merely one of the criteria for the communicability. Many other important criteria must be considered in the process of the law equation discovery, and in fact the SDS takes these criteria into account under the environment where the data are experimentally obtained. In this section, the extra and important criteria are discussed. Probably, the complete axiomatization of the definitions and the conditions of law equations without any exception may be difficult since some relations might be named as "laws" in purely empirical manner. However, the clarification of its criteria is considered to be highly important to give a firm basis of the science.

Some of the important conditions on the scientific proposition are given by R. Descartes. They are clarity, distinctness, soundness and consistency in the deduction of the proposition [4], and these conditions should be also take into account to clarify the scientific law criteria. I. Newton also proposed some conditions of the law equations [15]. The first condition is the objectiveness where the relation reflects only the causal assumptions of the nature while excluding any human's mental effects, the second the parsimony of the causal assumptions supporting the relation, the third the generality where the relation holds over the various behaviors in a domain and the forth the soundness where the relation is not violated by any experimental result performed under the environment following the causal assumptions. H.A. Simon also claimed the importance of the parsimony of the law description [17]. In the modern physics, the importance of the mathematical admissibility of the relation formulae under the nature of the time and the space also became to be stressed by some major physicists including R.P. Feynman [7].

We introduce the following definitions and propositions associated with the criteria for the law equations discovery based on the above claims.

Definition 10 (A Scientific Region) *A scientific region T is represented by the following quadruplet.*

$$\begin{aligned}
&T =< S, A, L, P > \\
&\textit{where} \\
&S = \{s_h | s_h \textit{ is a rule in syntax}, h = 1, ..., p\}, \\
&A = \{a_i | a_i \textit{ is an axiom in semantics}, i = 1, ..., q\}, \\
&L = \{\ell_j | \ell_j \textit{ is a postulate in semantics}, j = 1, ..., r\}, \\
&P = \{o_k | o_k \textit{ is an objective behavior}, k = 1, ..., s\}.
\end{aligned}$$

S is the syntax of T, and for example its elements are the coordinate system, the definitions of quantities such as velocity and energy and the definitions of the algebraic operators in physics. The axioms in A are the set of the mathematical relations independent of objective behaviors, for example, the relations of distances among points in an Euclidean space. A postulate $\ell_j (\in L)$ is a law equation where its validity is empirically believed under some conditions which will be described later. An example is the following law of gravity in physics.

$$F = G \frac{M_1 M_2}{R^2}, \tag{6}$$

where $F[kg{\cdot}m/s^2]$ is the gravity force interacting between two mass points $M_1[kg]$ and $M_2[kg]$ when their interval distance is $R[m]$. $G[m^3/(kg{\cdot}s^2)]$ is the gravity constant. A and L give the semantics of T.

In addition, the definition of T involves a set of objective behaviors P which is analyzed in the scientific domain, since the scientific domain is established for the purpose to study some limited part of the universe. In other words, S, A

and L must be valid within the analysis of P, and hence each ℓ_j is requested to satisfy the conditions of the law equations for P but not requested outside of P.

Moreover, an ℓ_j is used in the analysis of a part of P but not necessarily used for all of P. For example, the law of gravity is not necessarily used in the analysis of a spring behavior.

Definition 11 (Objective Behaviors of a Relation) *Given a mathematical relation e, if all quantities in e appear in the description of a behavior as mutually relevant quantities, the behavior is called an "objective behavior of e". A subset of P, in which the behaviors are the objective of e, is called "the set of the objective behaviors of e"* $P_e(\subseteq P)$.

For example, the gravity interaction between mass points characterized by the quantities of F, M_1, M_2 and R is an objective behavior of the aforementioned law of gravity.

Definition 12 (Satisfaction and Consistency of a Relation) *Given a mathematical relation e and its objective behavior, if the behavior is explicitly constrained by e, e is said to be "satisfactory" in the behavior. On the other hand, if the behavior does not explicitly violate e, e is said to be "consistent" with the behavior.*

When we consider the kinematic momentum conservation in the collision of two mass points, if the mass points are very heavy, this behavior is analyzed under the requirement that the law of gravity should be satisfactory. Otherwise, the law of gravity is ignored. But, it should be consistent in both cases.

Based on these definitions and the aforementioned claims of some major scientists, the criteria of a relation e to be a law equation are described as follows

(1) **Objectiveness**: All quantities appearing in e are observable directly and/or indirectly in the behaviors in P_e.

(2) **Generality**: The satisfaction of e is widely identified in the test on the behaviors included in P_e.

(3) **Reproducibility**: For every behavior in P_e, the identical result on the satisfaction and the consistency is identified in repeated tests.

(4) **Soundness**: The consistency of e is identified in the test on every behavior in P_e.

(5) **Parsimony**: e includes the least number of quantities to characterize the behaviors in P_e.

(6) **Mathematical Admissibility**: e follows the syntax S and the axioms of the semantics A.

Here, the "test" is an experiment or an observation, and the "identification" is to confirm a fact in the test while considering the uncertainty and/or the accuracy of the test. Though the objectiveness and the generality include the criteria of (3), (4) and (5) in wider sense, each criterion is more specifically defined in this literature to reduce their ambiguity.

Some widely known scientific relations are not identified as law equations among scientists. For example, given the enforced turbulence flow in a circular pipe, the heat transfer behavior from the flow liquid to the pipe wall is represented by the following Dittus-Boelter equation which is called as an "experimental equation" but not a "law equation" in thermo-hydraulics domain.

$$Nu = 0.023Re^{0.8}Pr^{0.4}, \tag{7}$$

where $Nu = hd/\lambda$, $Re = \rho ud/\eta$, and $Pr = \eta c_p/\lambda$. $h[W/(m^2 \cdot {}^\circ K)]$ is the coefficient of the heat transfer rate between the liquid and the wall, $d[m]$ the diameter of the circular pipe, $\lambda[W/(m \cdot {}^\circ K)]$, $\rho[kg/m^3]$, $u[m/s]$, $\eta[Pa \cdot s]$, $c_p[J/(kg \cdot {}^\circ K)]$ the heat conductance, density, velocity, viscosity and specific heat of the liquid under a constant pressure respectively [11].

This relation stands objectively independent of our interpretation. The set of the objective behaviors of the thermo-hydraulics P includes all behaviors over all value ranges of Nu, Re and Pr. Thus, according to the definition 11, P_e of the Dittus-Boelter equation is the set of all behaviors represented in some value ranges of Nu, Re and Pr in P. This equation meets the criterion of the objectiveness because Nu, Re and Pr are observable through some experiments. It is general over various enforced turbulence flows in circular pipes and reproducible for the repetition of the tests. It also has a parsimonious shape, and satisfies the unit dimensional constraint in terms of the mathematical admissibility. However, this equation is not sound in P_e, because it stands for only the value ranges of $10^4 \leq Re \leq 10^5$ and $1 \leq Pr \leq 10$, and is explicitly violated outside of these ranges. In this regard, this equation is not a law equation.

On the other hand, P of the classical mechanics includes the behaviors over all value ranges of mass, distance and force, and thus P_e of the law of gravity is the set of all behaviors represented in some value ranges of these quantities. This equation also meets the criteria of objectiveness, generality, reproducibility, parsimony and mathematical admissibility in P_e. Furthermore, as any behaviors in P_e do not violate this relation, it is sound.

Strictly speaking, the verifications of the generality and the soundness are very hard since they require the experimental knowledge on various behaviors. However, these can be checked if we relax the requirements to limit the verification within a given set of the objective behaviors. Under this premise, SDS seeks an equation having the generality to explain all behaviors shown by the combinations of the values of some quantities in the experiments on the objective behaviors. It also seeks the equation having the soundness not to contradict with all behaviors observed in the experiments. Eventually, the generality is subsumed by the soundness by limiting the behaviors for the verification. The objectiveness is ensured by seeking the relation among directly and indirectly observed quantities. The reproducibility is also ensured by checking if identical bi-variate relations are obtained multiple times in the repeated statistical tests.

The parsimony is automatically induced in the algorithm to compose the equation in bottom up manner. The mathematical admissibility is well addressed as mentioned earlier.

8 Summary

In this chapter, the criteria on the relation among quantities observed in objective behaviors to be a the law equation as the communicable knowledge among domain experts were discussed through the demonstration of a law discovery system SDS. Especially, the criterion of the mathematical admissibility has been analyzed in detail on the axiomatic basis. The definitions of scale-types of quantities and the admissibility conditions on their relations based on the characteristics of the scale-types have been introduced, and the extension of the major theorems in the unit dimensional analysis was shown. Through these analyses, the communicability criteria of the law equation have been clarified.

Moreover, the superior performance of SDS was demonstrated through some simulation experiments. In the evaluation, the validity of the presented principles has been confirmed, and its power to systematically discover candidate law equations over various domains along the communicability criteria has been shown.

In the recent study, the function and ability of SDS have been further extended. It became to discover law based models consisting of simultaneous equations [22]. Moreover, the most recent version of SDS can discover the law based models from the data which are passively observed not in the artificial experiments but the natural environment [23]. These developments extend the practical domains where communicable law equations are discovered for scientists.

References

1. Beaumont, A.P., Knowles, J. D. and Beuamont, G. P.: Statistical Tests: An Introduction With Minitab Commentary, Upper Saddle River, NJ, Prentice Hall PTR (1996)
2. Bridgman, P.W.: Dimensional Analysis, New Haven, CT, Yale University Press (1922)
3. Buckingham, E.: On physically similar systems; Illustrations of the use of dimensional equations, Physical Review, **IV, 4** (1914) 345–376
4. Descartes, R.: Discours de la Methode/Discourse on the Method, Heffernan, G. Eds., New York, NY, Univ of Notre Dame Pr; Bilingual edition (1637/1994)
5. Dzeroski, S. and Todorovski, L.: Discovering Dynamics: From Inductive Logic Programming to Machine Discovery, Journal of Intelligent Information Systems, Boston, MA, Kluwer Academic Publishers (1994) 1–20
6. Falkenhainer B.C. and Michalski R.S.: Integrating Quantitative and Qualitative Discovery: The ABACUS System, Machine Learning, Boston, MA, Kluwer Academic Publishers (1986) 367–401
7. Feynman, R.P.: The Character of Physical Law, Boston, MA, Charles E. Tuttle Co. Inc. (1965)

8. Kan, M., Miyata, N. and Watanabe, K.: Research on Spaciousness, Japanese Journal of Architecture **193** (1972) 51–57
9. Koehn B. and Zytkow J.M.: Experimeting and theorizing in theory formation, Proceedings of the International Symposium on Methodologies for Intelligent Systems, New York, NY, ACM SIGART Press (1986) 296–307
10. Kokar M.M.: Determining Arguments of Invariant Functional Descriptions, Machine Learning, Boston, MA, Kluwer Academic Publishers (1986) 403–422
11. Kouzou, T.: Material on Heat Transfer Engineering, Takataya Kouzou Eds., Tokyo, Japan, The Japan Society of Mechanical Engineers (JSME) **1**, Chap.2 (1986) 55–56
12. Krantz, D.H., Luce, R.D., Suppes, P., Tversky, A.: Fundations of Measurement, NY, Academic Press **1** (1971)
13. Langley P.W., Simon H.A., Bradshaw G. and Zytkow, J.M.: Scientific Discovery; Computational Explorations of the Creative Process, Cambridge, MA, MIT Press (1985)
14. Luce, R.D.: On the Possible Psychological Laws, The Psychological Review, **66** 2 (1959) 81–95
15. Newton, I.: Principia, Vol.II, The System of the World, Translated into English by A. Motte (1729), London, England, University of California Press, Ltd. (Copyright 1962) (1686)
16. Rissanen, J.: Modeling by shortest data description, Automatica **14** (1978) 465–471.
17. Simon, H.A.: Models of Discovery, Dordrecht, Holland, D. Reidel Publishing Company (1977)
18. Stevens, S.S.: On the Theory of Scales of Measurement, Science **103** 2684 (1946) 677–680
19. Todorovski, L. and Dzeroski, S.: Declarative Bias in Equation Discovery, Proceeding of the fourteenth International Conference on Machine Learning, San Mateo, CA, Morgan Kaufmann (1997) 376–384
20. Washio, T. and Motoda, H.: Discovering Admissible Models of Complex Systems Based on Scale-Types and Identity Constraints, Proc. of Fifteenth International Joint Conference on Artificial Intelligence, **2** (1997) 810–817
21. Washio, T., Motoda, H. and Niwa, Y.: Enhancing the Plausibility of Law Equation Discovery, Proc. of the Seventeenth International Conference on Machine Learning (2000) 1127–1134
22. Washio, T. and Motoda, H.: Discovering Admissible Simultaneous Equations of Large Scale Systems, Proc. of Fifteenth National Conference on Artificial Intelligence, (1998) 189–196
23. Washio, T., Motoda, H. and Niwa, Y.: Discovering Admissible Model Equations from Observed Data Based on Scale-types and Identity Constraints, Proc. of Sixteenth International Joint Conference on Artificial Intelligence, **2** (1999) 772–779