Discovery of First Principle Equations
Based on Scale-Type-Based and Data-Driven Reasoning

Takashi Washio and Hiroshi Motoda

Institute of Scientific and Industrial Research (I.S.I.R.), Osaka University

8-1 Mihogaoka, Ibaraki, Osaka, 567, Japan

phone: +81-6-879-8541, fax: +81-6-879-8544

e-mail: {washio, motoda}@sanken.osaka-u.ac.jp

abstract

We propose a novel approach to automatically discover formulae of first principles from the measurement data. The formulae obtained by our approach are ensured to reflect the first principles despite the fact that we try to use as littel knowledge of the target relation as possible to make it applicable to various domains not limited to physics problem. The basic idea is the combined use of deductive "*scale-type-based reasoning*" and "*data-driven reasoning*". The former is based on the scale-type information of quantities different from the quantity dimension. The features of our approach are demonstrated and discussed through its applications to many classes of examples. This approach is expected to provide a basis to discover first principles of various domains such as physics, biology, psychology, economics and social science.

1   Introduction

One of the early work to automatically discover formulae of physical first principles is a method called dimensional analysis that was based on the product theorem [Bridgman22].

Product Theorem *Assuming absolute significance of relative magnitudes of physical quantities,*

the function ρ relating a secondary quantity, Π, to the appropriate primary quantities, $x, y, ...$

has the form: $\Pi = \rho(x, y, z, ...) = Cx^a y^b z^c ....$, where $C, a, b, c, ...$ are constants.

There is another important theorem that is called Buckingham Π-theorem [Buckingham14].

Buckingham Π-theorem *If $\phi(x, y, ...) = 0$ is a complete equation, then the solution can be written in the form $F(\Pi_1, \Pi_2, ..., \Pi_{n-r}) = 0$, where n is the number of arguments of ϕ, and r is the number of basic dimensions in $x, y, z .....$ For all i, Π_i is a dimensionless number.*

Basic dimensions are such dimensions as length $[L]$, mass $[M]$ and time $[T]$, scaling quantities independently of other dimensions in the given ϕ. For example, this theorem can be used together with the product theorem to obtain the oscillation period $t[T]$ of a simple pendulum from its stick length $l[L]$, gravity acceleration $g[LT^{-2}]$ and deviation angle θ [no dimension] (see Fig. 1.). We can find two dimensionless quantities $\Pi_1 = t(g/l)^{1/2}$ and $\Pi_2 = \theta$, and derive the basic formula of the solution as $F(\Pi_1, \Pi_2) = F(t(g/l)^{1/2}, \theta) = 0$ based on the theorem. Using this dimensional analysis technique, Bhaskar and Nigam introduced a concept "*regime*" which is a formula $\Pi_i = \rho_i(x, y, ...)$ defining a dimensionless quantity Π_i[Bhaskar90]. In the above example, $t(g/l)^{1/2}$ and θ are the regimes. Further, they defined an "*ensemble*" which stands for the combination of a complete equation $F(\Pi_1, \Pi_2, ..., \Pi_{n-r}) = 0$ and the set of regimes contained in the equation. The equation $F(\Pi_1, \Pi_2, ..., \Pi_{n-r}) = 0$ is called an "*ensemble equation*". In the current example, $F(\Pi_1, \Pi_2) = 0$ is an ensemble equation, and it composes an ensemble together with the regimes $t(g/l)^{1/2}$ and θ. A regime refers to a decomposable subprocess. An ensemble stands for a complete physical process in the system. The dimensional analysis utilizes the physical insights into the objective system that is described by the unit dimension of each quantity. Another method to automatically derive physical formulae is the symmetry-based approach which was proposed by Ishida[Ishida95]. He applied the principle of symmetry to physical domains. The system he developed heuristically searches for the invariant physical formula under a given set of

isomorphic mappings. In the example of the simple pendulum, we can apply further constrains of symmetry on the form $F(t(g/l)^{1/2}, \theta) = 0$. One is phase translatory symmetry on t, and this gives $F(t(g/l)^{1/2}, \theta) = F((t(g/l)^{1/2} + 2m\pi, \theta)$, where m is an integer. Another is mirror symmetry of the time t and angle θ giving $F(t(g/l)^{1/2}, \theta) = F(-t(g/l)^{1/2}, -\theta)$. A formula satisfying these constraints is $\theta = \sin(t(g/l)^{1/2})$. As is evident, this method also has to utilize the explicit knowledge of physical characteristics of the object in terms of symmetry.

In contrast with these methods that are grounded on physical knowledge, some challenges have been made by using data-driven approaches to discover the formulae of the first principles that govern the objective system. Langley and others' BACON systems [Langley89] are most well known as the pioneering work. They founded the succeeding BACON family. FAHRENHEIT [Koehn86], IDS [Nordhausen90] and KEPLER [Wu89] are such successors that basically use similar algorithms to BACON in search for a complete equation governing the data measured in a continuous process. The target objects in various domains can be modeled because they do not use any a priori knowledge of the objects. However, the heuristics in the search algorithm have no firm theoretical bases, and the search is limited to enumerating only polynomial and meromorphic formulae. Accordingly, the set of the solutions obtained is not ensured to be sound and complete.

To alleviate this drawback, some members of the BACON family, *e.g.*, ABACUS [Falkenhainer86], have a mode to utilize the information of the quantity dimension to prune the meaningless terms based on the principle of dimensional homogeneity. However, this heuristic still limits the search to polynomial and meromorphic formulae, and leaves too many types of equations in candidates. COPER [Kokar86], another type of equation finding systems based on the principles of dimensional analysis can significantly reduce the candidate generation by explicit use of the information about the quantity dimension. Its another significant advantage is higher credibility

of the solution that it is not merely an experimental equation but is indeed a first principle equation. However, these approaches are not applicable when the information of the quantity dimension is not available. This fact strongly limits their applicability to non-physics domains.

All the previous approaches employ the following assumptions.

Assumption 1 *The relation among quantities is represented by a complete equation in a continuous process under consideration.*

Though Bhaskar and Nigam demonstrated the applicability of their dimension-based method to the systems consisting of multiple complete equations, the specification of a set of quantities is required for each complete equation in the framework, and thus this assumption must be maintained[Bhaskar90]. Their dimension-based approach also requires the following assumptions.

Assumption 2 *The type of scales of physical quantities is limited to ratio scale.*

Assumption 3 *The information on unit dimension of every quantity is available.*

Assumption 4 *Given a regime $\rho_i(\Pi_i, x, y, ...) = 0$, either one of the following conditions holds.*

1) It is a unique regime of a complete equation.

2) For each quantity x in ρ_i, any other regimes do not contain x, or any other regimes are related in such a way that x does not change Π_i from outside of ρ_i.

COPER also requires the assumptions 2 and 3 , and ABACUS the assumption 3 if it tries to apply its dimension based heuristics. The assumption 2 is the restatement of the assumption in the product theorem, *i.e.*, the absolute significance of relative magnitudes of physical quantities. Although the quantities of the ratio scale are quite common in various domains, there is another

scale-type of quantities called "*interval scale*" which is also widely encountered as explained later. The assumption 3 is needed for any dimension based approach, and this limits the applicability of the approach to the domains where the relations among measurement processes of quantities are well-defined. The assumption 4 was recently pointed out by Kalagnanam and Henrion[Kalagnanam94]. Every data-driven approach in BACON family assumes the following environment of data measurement.

Assumption 5 *The measurements on the relation in any subset of quantities in a complete equation can be made while holding the other variables constant under an experimental environment, and the measured data can be subsequently used to reason the formula relating the variables.*

The objective of this paper is to propose a new approach to automatically discover formulae of first principles. Our approach lies in between the deductive approaches as represented by dimension-based reasoning and the empirical approaches as represented by BACON family. The former systems such as ABACUS (in case that it uses the dimension based heuristics) and COPER also take the position between the deductive and empirical frameworks. However, their applicability is quite limited due to the strong assumptions 2 and 3 as explained earlier. In contrast, our approach removes the limitations of the assumptions 2 , 3 and 4 , i.e., it covers the quantities of both ratio and interval scales, does not require the information on unit dimension but only scale-types, and covers the mutually dependent regimes. The following theoretical aspects are newly introduced in this paper to achieve these advantages.

1) The sound relations among quantities of ratio and interval scales within a regime are characterized, and the product theorem is extended to reflect the relations.

2) The structure of dimension is characterized for both ratio and interval scale quantities, and the Buckingham Π-theorem is extended for the structure.

3) The condition that each regime contains a set of quantities that are mutually exclusive in an ensemble is characterized (*i.e.*, the condition that a set of regimes forms a partition of the quantities), and an algorithm to identify such regimes is proposed.

4) A new approach to automatically discover formulae of first principles is proposed based on scale-type-based and data-driven reasoning. It is applicable to the target formula containing both of ratio and interval scale quantities without using any information of their unit dimensions.

The other two assumptions 1 and 5 remain the same in our approach at present. On these grounds, our approach can derive a sound set of solutions of the formulae for each regime similarly to the dimension-based approach while still maintaining the advantage of the data-driven approach like BACON, *i.e.*, it does not require a priori insights into the objective system but only data that can be obtained by measurements. Thus, it is expected to be applied to various domains and not limited to physics. The subsequent section explains the basic principle of scale-type based reasoning. The section 3 proposes and demonstrates our new approach, and the section 4 discusses the features of our approach through its applications to many classes of examples.

2   Basic Principle of Scale-Type-Based Reasoning

Since after Helmholtz originated a research field of "*measurement theory*" [Helmholtz87], many literatures have been published on this topic in the 20th century. Stevens defined the measurement process as "*the assignment of numerals to object or events according to rules*" [Stevens46]. He claimed that different kinds of scale-types and different kinds of measurement are leaded, if numerals can be assigned under different rules. Table 1 shows three scale-types, where the last two of them, *e.g.*, ratio and interval scales, are defined by him. The interval scale and the ratio scale are the majorities of quantities in physical, psychological, economical and sociological

domains. Examples of the ratio scale quantities are physical mass, absolute temperature, pressure, time interval, frequency, and currency value. Each has an absolute origin, and the ratio of two different values x_1/x_2 is invariant against their unit change. Examples of the interval scale quantities are temperature in Celsius and Fahrenheit, energy, entropy, time (not time interval), and sound tone (proportional to the order of white keys of a piano). The origin of their scales are not absolute, and are changeable by human's definitions, whereas the ratio of two different intervals $(x_1 - x_2)/(x_3 - x_4)$ is invariant against their unit change. The scale-type is different from the dimension. The scale-type just represents the definition of the measurement rule. Accordingly, the use of the information of the scale-types does not require the assumption 3 . Another important scale-type we should mention is "*absolute scale*" as indicated in Table 1 where dimensionless quantities belong to this type. Any change of units is not defined for absolute scale quantities, and hence any unit conversion is not admissible.

Luce claimed that the basic formula of the functional relation between any two quantities of ratio and interval scales can be determined by their scale-type features, if the quantities are not coupled through any dimensionless quantities[Luce59]. In this case, the two quantities should have common basic dimensions, and consequently a unit change of quantity affects the value of another quantity. Suppose x and y are both ratio scale quantities, and y is defined by x through a continuous functional relation $y = u(x)$. Suppose the form of $u(x)$ is logarithmic, *i.e.*, $y = \log x$, and we multiply a positive constant k to x, *i.e.*, a change of unit, without violating the group structure of the ratio scale quantity x shown in Table 1, then this leads $u(kx) = \log k + \log x$. This fact causes the shift of the origin of y by $\log k$, and violates the group structure of y which is the ratio scale quantity. Hence, the direct functional relation from x to y must not be logarithmic. As the admissible transformations of x and y in their group structures are $x' = kx$ and $y' = Ky$ respectively, the relation of $y = u(x)$ becomes as $y' = u(x') \leftrightarrow Ky = u(kx)$. The factor K of the changed unit of y may depend on k, but it shall not depend upon x, so we denote

it by $K(k)$. Consequently, we obtain the following constraints on the continuous function $u(x)$.

$$u(kx) = K(k)u(x),$$

where $k > 0$ and $K(k) > 0$, as these are the factors of the changed units. The constraints for the different combinations of the scale-types are summarized in the fourth column in Table 2 [Luce59]. Luce derived each solution of $u(x)$ under the condition of $x \geq 0$ and $u(x) \geq 0$. We extended his theorems to cover the negative values of x and $u(x)$, and the result is summarized in the last column in Table 2. Because of the space limitation, we omit the proof for the results. For the detail, see the references[Washio96b], [Washio96a].

The indefinability of a ratio scale from an interval scale in Table 2 is because the ratio scale involves the information of an absolute origin, but the interval scale does not. Luce gave some examples of the equations in Table 2 [Luce59]. The quantities entering into Coulomb's law, Ohm's law and Newton's gravitation law are all ratio scales, and the formula of each law is a power function which follows the formula of No.1 for the pairwise relation in the table. We see examples of formulae No.2.1 and No.2.2 for laws associated with energy and entropy. The total energy U of a body having a constant mass m and moving at velocity v is $U = mv^2/2 + P$, where P is the potential energy. If the temperature of a perfect gas is constant, then the entropy E of the gas as a function of the pressure p is of the form $E = -R \log p + E'$, where R and E' are Boltzmann's constant and a reference value of entropy respectively. In psychophysics, Fechner's law states that the sound tone s of human sensing is proportional to the logarithm of the sound frequency f, *i.e.*, $s = \alpha \log f + \beta$, where s is an interval scale, and f is a ratio scale. As an example of formula No.4, there is the relation $x = vt + x_0$ for a particle moving at its constant velocity v, where x is the position at the present time t and x_0 is the initial position.

Finally, the following important consequence should be indicated.

Theorem 1 *An absolute scale quantity can have functional relations of any continuous formulae with other absolute scale quantities.*

For example, the behavior solution of the simple pendulum of Fig. 1 is $\theta = \sin(t(g/l)^{1/2})$. The triangular function sin which does not belong to Table 2 can hold, because θ and $t(g/l)^{1/2}$ are dimensionless, *i.e.*, absolute scale.

3   Theory and Method to Derive Possible Equations of First Principles

3.1   Scale-Type-Based Reasoning within a Regime

This subsection describes the theory and the method of scale-type-based reasoning which uses a priori knowledge of quantity scale-types (but not the knowledge of physics) to derive the formula of ρ_i for a given regime. First, the product theorem is extended to include interval scale quantities.

Theorem 2 *Assuming primary quantities in a set R are ratio scale-type, and those in another set I are interval scale-type, the function ρ relating a secondary quantity Π to $x_i \in R \cup I$ has one of the forms:*

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

$$\Pi = \sum_{x_i \in R} a_i \log |x_i| + \sum_{I_k \in C_{\bar{g}}} a_k \log(\sum_{x_j \in I_k} b_{kj}|x_j| + c_k) \qquad (ii)$$

$$+ \sum_{x_\ell \in I_g} b_{g\ell}|x_\ell| + c_g$$

where all coefficients except Π *are constants, and* C *is a covering of* I, $C_{\bar{g}}$ *a covering of* $I - I_g$ *(*$I_g \subseteq I$*).*

Theorem 3 *If* $R = \phi$ *in Theorem 2, the function* ρ *relating a secondary quantity* Π *to* $x_i \in I$ *has one of the forms:*

$$\Pi = \prod_{I_k \in C} (\sum_{x_j \in I_k} b_{kj} |x_j| + c_k)^{a_k} \qquad (i)$$

$$\Pi = \sum_{I_k \in C_{\bar{g}}} a_k \log(\sum_{x_j \in I_k} b_{kj} |x_j| + c_k) \qquad (ii)$$

$$+ \sum_{x_\ell \in I_g} b_{g\ell} |x_\ell| + c_g$$

where all coefficients except Π *are constants, and* C *is a covering of* I, $C_{\bar{g}}$ *a covering of* $I - I_g$ *(*$I_g \subseteq I$*).*

Because of the space limitation, the proofs of theorems 2 and 3 are omitted. The principle of the extension is a certain symmetry that the relations given in Table 2 must hold for each pair of quantities sharing some basic dimensions in a regime. For example, if we look at the relation on a pair of x and $y(= u(x))$ in R while regarding the rest of quantities as constants, both forms (i) and (ii) in theorem 2 reduce to the realtion of $y = u(x) = \alpha_* |x|^\beta$ which is the first relation in Table 2. Thus, both formulae are the extension of the product theorem. Furthermore, if we look at the relation on a pair of x in R and $y(= u(x))$ in I, both formulae reduce to the relations 2.1 and 2.2 in the table. The similar discussion holds for the relation on a pair of two interval scale quantities. Accordingly, these theorems present the generic relations among ratio and interval scale quantities which mutually share some basic dimensions.

Now, we consider to derive the candidates of formula $\Pi = \rho(x_1, ..., x_m)$ of a given regime ρ, where the set of quantities except derivative Π is $Q = \{x_1, ..., x_m\}$. The following algorithm based on the product theorem and the above theorems 2 and 3 derives the candidates.

Algorithm 1

(Step 1) $Q = \{x_1, ..., x_m\}$, $S = \phi$. Let R be a set of all quantities of ratio scale in Q. Let I be a

set of all quantities of interval scale in Q. $(Q = R + I)$

(Step 2) If $I = \phi$ {

Based on the product theorem, push the following to S.

$\Pi = \rho(x_1, x_2, ..., x_m) = \prod_{x_i \in R} |x_i|^{a_i}$. }

else if $R = \phi$ {

Push the relation of theorem 3 to S.}

else {

Enumerate candidate relations of (i) in theorem 2 for all coverings C of I, and push

those candidates to S. Enumerate candidate relations of (ii) in theorem 2 for all

coverings $C_{\bar{g}}$ of I, and push those candidates to S.}

The candidates are rested in the list S. The result of S is sound, since the soundness is ensured

by the product theorem, the theorems 2 and 3 . The complexity of this algorithm is not low,

since many coverings of I must be tested in the theorems. However, it is not very problematic,

because the size of a regime is generally quite limited.

This algorithm have been tested by various physical laws. The following is an example of the

ideal gas equation which forms a unique regime. A regime involving four quantities of pressure

p, volume v, mass m and temperature t is given, thus $Q = \{p, v, m, t\}$. The quantities p, v and

m are ratio scales, while only t is an interval scale unless it is absolute temperature. We assumed

the positive sign of p, v and m in advance, hence the solutions for their negative values were

omitted. The algorithm figured out the three candidate relations. The first one comes from (i)

of theorem 2 :

$$\Pi = p^{a_1} v^{a_2} m^{a_3} (b_1 t + c_1),$$

and the second and third ones come from (ii) of theorem 2 :

$$\Pi = a_1 \log p + a_2 \log v + a_3 \log m + a_4 \log(b_1 t + c_1) + c_2.$$

$$\Pi = a_1 \log p + a_2 \log v + a_3 \log m + b_1 * t + c_1.$$

The first two formulae are essentially identical, and thus the following candidates were obtained.

$$p^{-a_1} v^{-a_2} = \Pi^{-1} m^{a_3} (b_1 t + c_1).$$

$$p^{a_1} v^{a_2} = m^{-a_3} \exp(\Pi - b_1 t - c_1).$$

The former solution reflects the right formula of the ideal gas equation, when the temperature is not absolute one. Once the candidate formulae of a regime are determined, the correct formula and the values of its coefficients must be specified in data-driven manner.

3.2   Data-Driven Reasoning on Ensemble

Before proposing the data-driven reasoning on an ensemble, the definition of dimension, regime, ensemble and the consequence of Buckingham Π-theorem need to be re-examined. Within the conventional view, a basic dimension, D, of a quantity is defined by a basic unit, U, of the quantity, *i.e.*, the basic dimension is represented as D(U). The value of a measured quantity is represented by the numeric value of units equivalent to the amount of the quantity. For example, if the mass of an object is equal to the mass of seven basic units where each mass unit is named as $1kg$, then the mass of the object is said to be $7kg$. When we use another basic unit such as $1g$, then the measured value of the same object will be 1000 times larger. This definition is valid only for basic quantities of ratio scale. Basic dimension of interval scale has a bit more complex structure. A basic unit and a *basic origin* are required to determine the value of temperature in Celsius. These unit and origin are different for temperature in Fahrenheit. Consequently, we need to extend the notion of a basic dimension as D(U, O) where O stands for a basic origin. In

case of a basic dimension of ratio scale, O is constrained to be 0, while O can take an arbitrarily value for a basic dimension of interval scale.

Based on the above discussion, the notion of the structure of a general dimension consisting of multiple basic dimensions is also extended. For example, let us consider the total energy of a particle. The particle having a certain temperature is moving at a certain height in a certain velocity. Thus, its total energy E_t is:

$$E_t = E_h + E_p + E_m,$$

where E_h, E_p and E_m are contained heat energy, potential energy and kinematic energy respectively. Every energy has an identical unit of $[M][L]^2[T]^{-2}$ where $[M], [L]$ and $[T]$ are basic units of mass, length and time. On the other hand, E_h, E_p and E_m are interval scale, since E_h is defined by temperature measured on a reference temperature point $[T_o]$, E_p by a reference height level $[H_o]$ and E_m by the kinematic energy of a reference coordinate $[E_o]$. Every energy component has its own reference point, *i.e.* the basic origin. Because E_t is the summation of E_h, E_p and E_m, the dependency of the value of E_t to these basic origins has a linear form of $mc[T_o] + mg[H_o] + [E_o]$ where m is the mass of this particle, c the specific heat coefficient and g the gravitational acceleration. Accordingly, the structure of the dimension of E_t is represented as $D([M][L]^2[T]^{-2}, mc[T_o] + mg[H_o] + [E_o])$. The structure of a general dimension can be represented as $D(U_1^{\alpha_1} U_2^{\alpha_2} \cdots, \beta_1 O_1 + \beta_2 O_2 + \cdots)$ where $U_1, U_2, \cdots$ are basic units, and $O_1, O_2, \cdots$ are basic origins. This is because the relation among origins follows the theorem 3 , while the relation among units follows the product theorem.

This extended structure of the dimension requires to extend the Buckingham Π-theorem as follows.

Theorem 4 (Extended Buckingham Π-theorem) *If $\phi(x, y, z \cdots) = 0$ is a complete equa-*

tion, and if each argument is one of interval, ratio and absolute scale-types, then the solution can be written in the form

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

where n is the number of arguments of ϕ, and r and s are the number of basic units and that of basic origins of the dimensions in $x, y, z \cdots$. For all i, Π_i is a dimensionless quantity.

The basic background of the original Buckingham Π-theorem is the fact that the intrinsic degree of freedom of n ratio scale quantities of $x, y, z \cdots$ is $n - r$, because the values of r quantities are determined by the specification of r basic units in the measurement. As an interval scale quantity has an extra factor, *i.e.*, basic origin, beside the basic unit, the values of $r + s$ quantities are determined by the specification of the measurement, and the intrinsic degree of freedom reduces to $n - r - s$. This theorem ensures that we can remove the limitation of the assumption 2 .

Another important subject is to overcome the limitation of the assumption 4 that was pointed out by Kalagnanam and Henrion[Kalagnanam94]. The process of identifying the formula of each regime $\rho_i(x, y \cdots) = \Pi_i$ is basically to find a relation among a subset of quantities in a complete equation while holding the condition that must hold outside of $\rho_i = \Pi$ constant as stated in the assumption 5 . If the assumption 4 does not hold, this operation is impossible because of the dependency among multiple regimes. This difficulty can be removed by introducing an idea of *pseudo-regime.*

Definition of pseudo-regime

A pseudo-regime is a subsystem, where the relation among quantities follows one of the product theorem, theorem 2 and 3 , of a given complete equation.

Consider the example of convection heat transfer depicted in Fig. 2 given by Kalagnanam et al.[Kalagnanam94] The heat transfer from a fluid to a pipe wall takes place through convection

when a fluid is forced through a pipe. A complete equation for this phenomenon under the turbulent flow is known as follows.

$$\Pi_1 = 0.023\Pi_2{}^{0.8}\Pi_3{}^{0.4},$$

where $\Pi_1 = \frac{hd}{k}, \Pi_2 = \frac{dv\rho}{\mu}$ and $\Pi_3 = \frac{c\mu}{k}$. Here, h is the convection heat transfer coefficient dependent on the other quantities, d and v are diameter of pipe and velocity of stream, and μ, ρ, c and k are the material quantities of the fluid, *i.e.*, viscosity, density, specific heat and thermal conductivity respectively. The number of quantities and that of the basic dimensions are $n = 7, r = 4$ and $s = 0$ respectively. Hence, three $(n - r - s = 3)$ regimes exist in this ensemble, and Π_1, Π_2 and Π_3 are called as Nusselt's, Reynold's and Prandtl's numbers in the thermal hydraulics domain. This example violates the assumption 4, because d, μ and k appear in multiple regimes. However, this equation is regarded as a unique pseudo-regime, because it can be rewritten to follow the product theorem as:

$$\frac{k^{0.6}v^{0.8}\rho^{0.8}c^{0.4}}{hd^{0.2}\mu^{0.4}} = \frac{1}{0.023}.$$

Accordingly, the data-driven search of a pseudo-regime can discover this formula without facing the difficulty pointed out by Kalagnanam and Henrion. On the other hand, the relation can be decomposed into:

$$\Pi_1' = d, \quad \Pi_2' = k, \quad \Pi_3' = \mu, \quad \Pi_4' = h, \quad \Pi_5' = v\rho \quad \text{and} \quad \Pi_6' = c.$$

Because these pseudo-regimes are mutually independent, the data-driven discovery of these pseudo-regimes is also possible. Consequently, the following theorem can be stated.

Theorem 5 *Given a complete equation $\phi(x_1, x_2, ...x_n) = 0$, its decomposition into pseudo-regimes always exists where*

$$F(\Pi_1, \Pi_2, ..., \Pi_k) = 0 \quad and \quad \{\rho_i'(x_{1_i}, x_{2_i}, ...x_{m_i}) = \Pi_i | i = 1, ..., k\}.$$

Here $\{Q_i | Q_i = \{x_{1_i}, x_{2_i}, ...x_{m_i}\}, i = 1, ..., k\}$ *is a partition of entire quantity set* $Q = \{x_1, x_2, ...x_n\}$.

The quantity Π of a pseudo-regime may not be dimensionless, since a pseudo-regime does not necessarily consist of quantities that cancel out every dimensions. The Π defined in a pseudo-regime is called a "*pseudo dimensionless quantity*". Accordingly, the following assumption must hold.

Assumption 6 *The measurement for the data of the entire complete equation must follow a consistent unit system* [1]

Otherwise, the formulae of the ensemble equation may become invalid, since it is not invariant to the change of dimensions. In addition, the number of pseudo-regimes does not follow $n - r - s$.

The algorithm to discover an ensemble equation and a set of pseudo-regimes involves a stage to identify pseudo-regimes, a stage to determine the arguments Πs of the ensemble and a stage to determine the formula of the ensemble equation. The following algorithm to discover the formula of a complete equation from the given data has been constructed.

---

[1]For example, SI units system should not mixed up with CGS units system in an experiment.

Algorithm 2

(Step 1) $E = \{x_1, x_2, ...x_n\}$, $LE = \phi$ and $k = 1$.

(Step 2) Repeat until k becomes equal to n. {

    Repeat for every partition Γ_i of E where $|\Gamma_i| = k$. {

        Repeat for every $Q_{i_j} \in \Gamma_i (j = 1, ..., k)$. {

            Apply the algorithm 1 to Q_{i_j} to identify a pseudo-regime. Test each solution in S

            obtained in the algorithm 1 through the least square fitting to the measured data

            under some sets of constant values of quantities in $(E - Q_{i_j})$. If some solutions

            are accepted, substitute them to a list LQ_{i_j}.}

        If every $LQ_{i_j} \neq \phi$, let a list $L\Gamma_i = \{LQ_{i_j} | Q_{i_j} \in \Gamma_i, j = 1, ..., k\}$. Push $L\Gamma_i$ to LE.}

    If $LE \neq \phi$, go to (Step 3), else $k = k + 1$.}

(Step 3) Repeat for every $L\Gamma_i$ in LE. {

    Take Cartesian products $LP_i = LQ_{i_1} \times LQ_{i_2} \times \cdots \times LQ_{i_k}$ in $L\Gamma_i$.

    Let $\Pi(\rho_{i_j})$ be the pseudo dimensionless quantity of the regime $\rho_{i_j}(\in LQ_{i_j})$.

    Repeat for every $\{\rho_{i_1}, \rho_{i_2}, ..., \rho_{i_k}\} \in LP_i$. {

    Determine the formula $F(\Pi(\rho_{i_1}), \Pi(\rho_{i_2}), ..., \Pi(\rho_{i_k})) = 0$.}}

More concrete contents of this algorithm are demonstrated through an example of an electric circuit depicted in Fig. 3. Both V_1 and V_2 are the input voltages and V_1 is set higher than V_2. The intermediate voltage that is determined by the ratio of the resistances R_1, R_2 is provided to the base of the transistor Tr, and a capacitor C is connected to the emitter of the Tr. The line between the collector and the emitter is ON due to the voltage difference between the base and the emitter, when the C is not charged, *i.e.*, the voltage difference across the C is equal to $0V$. Consequently, the C is charged by the electric current through the collector and the emitter, and the voltage of the emitter approaches to that of the base. When the voltage difference

between the base and the emitter becomes negligible, the line between the collector and the

emitter becomes OFF, and the electric charge in the C is maintained. The dynamics of this

circuit is represented by a complete differential equation as

$$\dot{Q} \;=\; \frac{h_{FE}}{R_{BE}}(\frac{R_2}{R_1 + R_2}(V_1 - V_2) - \frac{Q}{C}),$$

where

$$\begin{array}{rcl}
Q & : & \text{electric charge of capacitor,} \\[1.2em]
\dot{Q} & : & \text{time derivative of } Q, \\[1.2em]
C & : & \text{charge capacity of capacitor,} \\[1.2em]
R_1, R_2 & : & \text{values of resistance,} \\[1.2em]
V_1, V_2 & : & \text{applied higher and lower voltage levels,} \\[1.2em]
h_{FE} & : & \text{amplifying rate of electric current of transistor,} \\[1.2em]
R_{BE} & : & \text{resistance between base and emitter of transistor.}
\end{array}$$

This equation consists of multiple elementary physical laws. The measurement data of $Q, \dot{Q}, V_1, V_2$

have been obtained for the various values of the parameters C, R_1, R_2 through the numerical

simulation of this circuit. The other parameters h_{FE} and R_{BE} are considered to be constant

throughout this example.

In (Step 1), the set of measured quantities E is set as $\{Q, \dot{Q}, V_1, V_2, C, R_1, R_2\}$. (Step 2) is the

process to enumerate all partitions of E where each subset of E is interpreted as a pseudo-regime.

The goodness of the least square fitting of each candidate solution derived by the algorithm 1 is

checked through the F-test which is a statistical hypothesis test to judge if the measured data

follows the solution. If a parameter is close enough to an integer value, then the parameter

is forced to be the integer, because the parameter having an integer value is quite common

in various domains. Every partition Γ_i, where all of its components are judged to be pseudo-regimes, is searched in an ascending order of the cardinal number k of the partition. Once such partitions are found at the level of the certain cardinal number, (Step 2) is finished at that level to obtain the solutions of the ensembles involving the least number of pseudo-regimes. This criterion decreases the ambiguity of the formula of the ensemble equation by reducing the number of the absolute scale quantities in it. In the current example, only the partition of $\{\{R_1, R_2\}, \{\dot{Q}\}, \{Q, C\}, \{V_1, V_2\}\}$ is accepted at the least cardinal number, $k = 4$. All quantities in the last pseudo-regime are of interval scale, and those in the others are of ratio scale, and thus the formula of these pseudo-regimes are enumerated as:

$$\Pi_1 = R_1^{a_1} R_2^{a_2}, \qquad \Pi_2 = \dot{Q}, \qquad \Pi_3 = Q^{a_3} C^{a_4}, \qquad \Pi_4 = b_1 V_1 + b_2 V_2 + c_1.$$

Notice that the number of quantities n is 7 while $r = 3$ and $s = 1$, since these quantities have three basic units of $[Q], [V] and [s]$ and one basic origin of voltage level. Therefore, the number of regime must be $n - r - s = 3$ according to the extended Buckingham Π-theorem, whereas the number of the pseudo-regimes is 4. In fact, Π_2, Π_3 and Π_4 are not dimensionless. Their more specific formulae are identified through the data fitting as follows.

$$\Pi_1 = R_1/R_2, \qquad \Pi_2 = \dot{Q}, \qquad \Pi_3 = Q/C, \qquad \Pi_4 = V_1 - V_2.$$

(Step 3) searches the ensemble equation, *i.e.*, the relation among Πs figured out in the preceding steps. This step utilizes the BACON like algorithm, where the basic approach is to generate bivariate equations based on data fitting and compose them in a hierarchical manner by repeatedly applying hypothesis generation rules. First, the polynomial and meromorphic relations between each pair of $\{\Pi_1, \Pi_2, \Pi_3, \Pi_4\}$ are tested, and a linear relation between Π_2 and Π_3 is found under constant Π_1 and Π_4. Thus, we obtain

$$d_1 = \Pi_2 + d_3 \Pi_3.$$

Next, the pairwise relations among $\{d_1, \Pi_1, \Pi_4\}$ are tested, and another linear relation between d_1 and Π_4 is found under constant Π_1.

$$d_1 = e_1 \Pi_4.$$

Also, it can find that d_3 is independent of the other quantities. Finally, the following relation between e_1 and Π_1 is searched.

$$e_1 = d_3 1/(\Pi_1 + 1).$$

By combining above formulae, the original equation to represent this system is reconstructed. This example demonstrates that our approach is applicable to systems consisting of multiple elementary laws and having dynamic and nonlinear characteristics.

4   Discussion

This section discusses the characteristics of the proposed method in several aspects. The assumptions 1, 5 and 6 are retained in our method. The first and the second assumptions are common in the conventional data-driven approach, and do not limit the applicability of our method in the same sense with the conventional approaches, when the experimental environment is available. The last assumption is acceptable, if the measurements of quantities are appropriately arranged to reflect the features of the objective system, even when the dimensions of quantities are not known. Accordingly, our approach has the wide applicability to various domains.

The overcoming of the assumption 2 and 3 in our framework is clear. The examples of the ideal gas law and the convection heat transfer contains only ratio scale quantities. On the other hand, the example of the electric circuit contains interval scale quantities. In each case, our approach discovers the formula of the first principles governing the target system without using

the information of quantity dimension. Another example to show the wide applicability of our approach with respect to the assumption 3 is the following equation known to be the law of spaciousness of a room in psychophysics[M. Kan.72].

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

where S_p, R, L_i and W_i are average spaciousness of a room, room capacity, average light intensity and solid angle of window at the location i in the room. Though the dimension of S_p is unclear, its scale-type is known to be ratio scale based on its definition. L and R are ratio scale, and W is absolute scale. Our approach can derive the above expression under given measurement data without any difficulty. This ability of our approach comes from the theory of scale-types, the extended Product Theorem and Buckingham Π-theorem.

The limitation stated in the assumption 4 has also been removed in our approach by introducing the idea of the pseudo-regime. The ideal gas law aforementioned follows the first condition of the assumption. An example that follows the second condition is the well-known Black's specific heat law. This relates the initial temperatures of two substances T_1 and T_2 with their temperature T_f after they have been combined. This law has the formula of

$$T_f = \frac{M_1}{M_1 + M_2} T_1 + \frac{M_2}{M_1 + M_2} T_2$$

where M_1 and M_2 are the masses of the two substances. This contains the following two regimes that do not mutually share any common quantities and an ensemble equation.

$$\Pi_1 = \frac{T_f - T_1}{T_f - T_2}, \Pi_2 = \frac{M_2}{M_1} \text{ and } \Pi_1 + \Pi_2 = 0.$$

The formulae of the ideal gas law and the Black's specific heat law are easily discovered in our approach. On the other hand, the assumption does not hold in the aforementioned examples of the convection heat transfer and the electric circuit. In spite of this more complex condition, our approach can reconstruct the target formulae through the identification of pseudo-regimes.

Another important feature of the objective system which has not been addressed earlier is if the system is represented by a *read-once formula*. A read-once formulae is a formula where each quantity appears at most once in it. The formulae of the ideal gas law, the convection heat transfer and the electric circuit are such examples. In contrast, the formula of the Black's specific heat law is not a read-once formula. Some of T_f, T_1 and T_2 in the first regime or one of M_1 and M_2 in the second appears more than once in the same regime under any conformation of the formula. Because the type of the formula of a regime is strongly limited, the search of the regime which is not read-once is not a very expensive process. However, suppose the case where a quantity appears more than once in different regimes. The following formula of the decay speed of a radioactive element is such an example.

$$\dot{C} = C_0 \lambda \exp(\lambda t),$$

where $\dot{C}$ is the decay speed, C_0 the initial amount of the element, λ the decay constant and t time. This has two regimes

$$\Pi_1 = \frac{\dot{C}}{C_0 \lambda} \text{ and } \Pi_2 = \lambda t,$$

where λ appears at each regime. In this case, our approach identifies the following three pseudo-regimes and an ensemble equation.

$$\Pi_1 = \frac{\dot{C}}{C_0}, \Pi_2 = t \text{ and } \Pi_3 = \lambda \text{ and } \Pi_1 = \Pi_3 \exp(\Pi_2 \Pi_3).$$

Then, the vast number of candidate formulae must be searched under the assumption that any Π_i can appear more than once in an ensemble equation. This becomes expensive, when the formula of the objective system is complex. This difficulty must be alleviated in the future work.

5   Conclusion

In this paper we have proposed a new method that enables us to automatically discover formulae of first principles. Our approach makes the best use of both the deductive and the empirical data-driven approaches. The major characteristics of our approach are summarized as follows.

1) The sound solutions of basic formulae of law equations within a regime are provided by using only the knowledge of quantity scale-types (not the knowledge of unit dimensions).

2) An ensemble equation and its pseudo-regimes in the objective system are identified from the experimental data without any strong limitations. The identified pseudo regimes are not necessarily the real regimes but they are ensured to hold independently of other regimes, thus considered to be informative decomposable subprocesses.

3) The applicability is not limited to well-defined domains such as physics, because the method does not require a priori knowledge such as unit dimensions except the knowledge of quantity scale-types as mentioned in 1).

The scale-type-based reasoning may provide a basis to discover qualitative models of various domains such as biology, psychology, economics and social science.

References

[Bhaskar90]   R. Bhaskar and Anil Nigam. Qualitative Physics Using Dimentional Analysis. *Artificial Intelligece*, Vol. 45, pp. 73–111, 1990.

[Bridgman22]   P.W. Bridgman. *Dimensional Analysis*. Yale University Press, New Haven, CT, 1922.

[Buckingham14] E. Buckingham. On physically similar systems; Illustrations of the use of dimensional equations. *Phys. rev.*, Vol. IV, No. 4, pp. 345–376, 1914.

[Falkenhainer86] B.C. Falkenhainer and R.S. Michalski. Integrating Quantitative and Qualitative Discovery: The ABACUS System. *Machine Learning*, Vol. 1, pp. 367–401, 1986.

[Helmholtz87] H.V. Helmholtz. *Numbering and Measuring from Epistemological Viewpoint.* Philosophische Aufsatze, 1887.

[Ishida95] Y. Ishida. Symmetry-Based Reasoning about Equations of Physical Laws. In *Working Papers of Nith International Workshop on Qualitative Reasoning, Amsterdam, the Netherlands*, pp. 84–93, 1995.

[Kalagnanam94] J. Kalagnanam, M. Henrion, and E. Subrahmanian. The Scope of Dimensional Analysis in Qualitative Reasoning. *Computational Intelligence*, Vol. 10, No. 2, pp. 117–133, 1994.

[Koehn86] B. Koehn and J.M. Zytkow. Experimeting and theorizing in theory formation. In *Proc. of the International Symposium on Methodologies for Intelligent Systems*, pp. 296–307. ACM SIGART Press, 1986.

[Kokar86] M.M. Kokar. Determining Arguments of Invariant Functional Descriptions. *Machine Learning*, Vol. 1, pp. 403–422, 1986.

[Langley89] P. Langley and J.M. Zytkow. Data-Driven Approaches to Empirical Discovery. *Artificial Intelligece*, Vol. 40, pp. 283–312, 1989.

[Luce59] R.D. Luce. On the Possible Psychological Laws. *The Psychological Review*, Vol. 66, No. 2, pp. 81–95, 1959.

[M. Kan.72] N. Miyata M. Kan. and K. Watanabe. Research on Spaciousness. *Japanese Journal of Architechture*, No. 193, pp. 51–57, 1972.

[Nordhausen90]   B. Nordhausen and P.W. Langlay. *An Integrated Approach to Empirical Discovery, Computational Models of Scientific Discovery and Theory Formation.* Morgan Kaufman Publishers, Inc, San Mateo, California, 1990.

[Stevens46]   S.S. Stevens. On the Theory of Scales of Measurement. *Science*, Vol. 103, No. 2684, pp. 677–680, 1946.

[Washio96a]   T. Washio and H. Motoda. Discovery of Possible Law Equations by Combined Use of Scale-Based and Data-Driven Reasoning. In *Working Papers of PKAW'96: The 1996 Pacific Rim Knowledge Aquisition Workshop*, pp. 130–149, 1996.

[Washio96b]   T. Washio and H. Motoda. Scale-Based Reasoning on Possible Law Equations. In Y. Iwasaki and A. Farquhar ed., editors, *AAAI Technical Report WS-96-01: Tenth Int. Workshop on Qualitative Reasonin*, Stanford Sierra Camp, CA, pp. 255–264, 1996.

[Wu89]   Y. Wu and S. Wang. Discovering Knowledge from Observational Data. In G. ed. Piatetsky-Shapiro, editor, *Knowledge Discovery in Database, IJCAI-89 Workshop Proceedings, Detroit, MI*, pp. 369–377, 1989.

Figure 1 A simple pendulum.

Figure 2 Convection heat transfer.

Figure 3 An electric circuit.

Table 1: Scale-types

Scale-type	Basic Empirical Operation	Mathematical Group Structure
Absolute	Determination of equality of values	Identity group $x' = x$
Ratio	Determination of equality of ratios	Similarity group $x' = kx$
Interval	Determination of equality of intervals or differences	Generic linear group $x' = kx + c$

Table 2: Constraints and their possible equations satisfying the scale-type characteristics

| No. | Scale-types | | Constraints | Possible Relations |
	Independent variable	Dependent (Defined) variable		
1	ratio	ratio	$u(kx) = K(k)u(x)$	$u(x) = \alpha_* \lvert x \rvert^\beta$
2.1	ratio	interval	$u(kx) = K(k)u(x) + C(k)$	$u(x) = \alpha \log \lvert x \rvert + \beta_*$
2.2				$u(x) = \alpha_* \lvert x \rvert^\beta + \delta$
3	interval	ratio	$u(kx + c) = K(k,c)u(x)$	impossible
4	interval	interval	$u(kx + c) = K(k,c)u(x) + C(k,c)$	$u(x) = \alpha_* \lvert x \rvert + \beta$

1) $k > 0, K(k) > 0, K(k,c) > 0, c$ and C can be any real numbers.

2) The notations α_*, β_* are α_+, β_+ for $x \geq 0$ and α_-, β_- for $x < 0$, respectively.
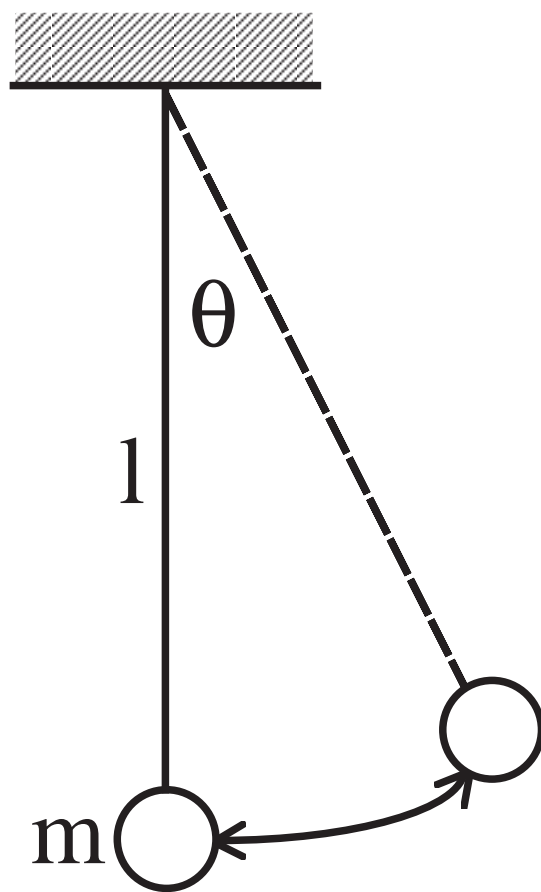
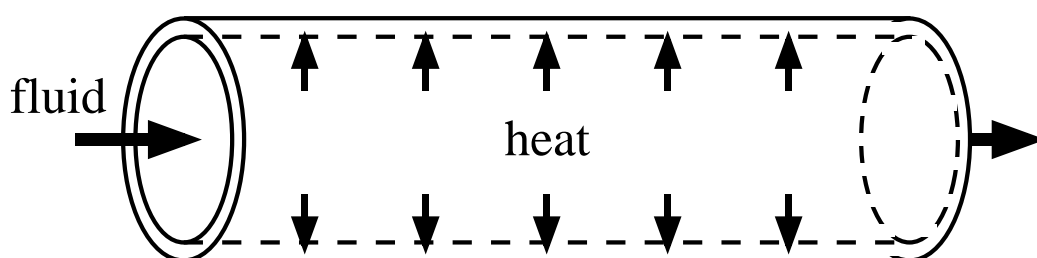Figure 1: Knowledge-Based Systems, Takashi Washio and Hiroshi Motoda

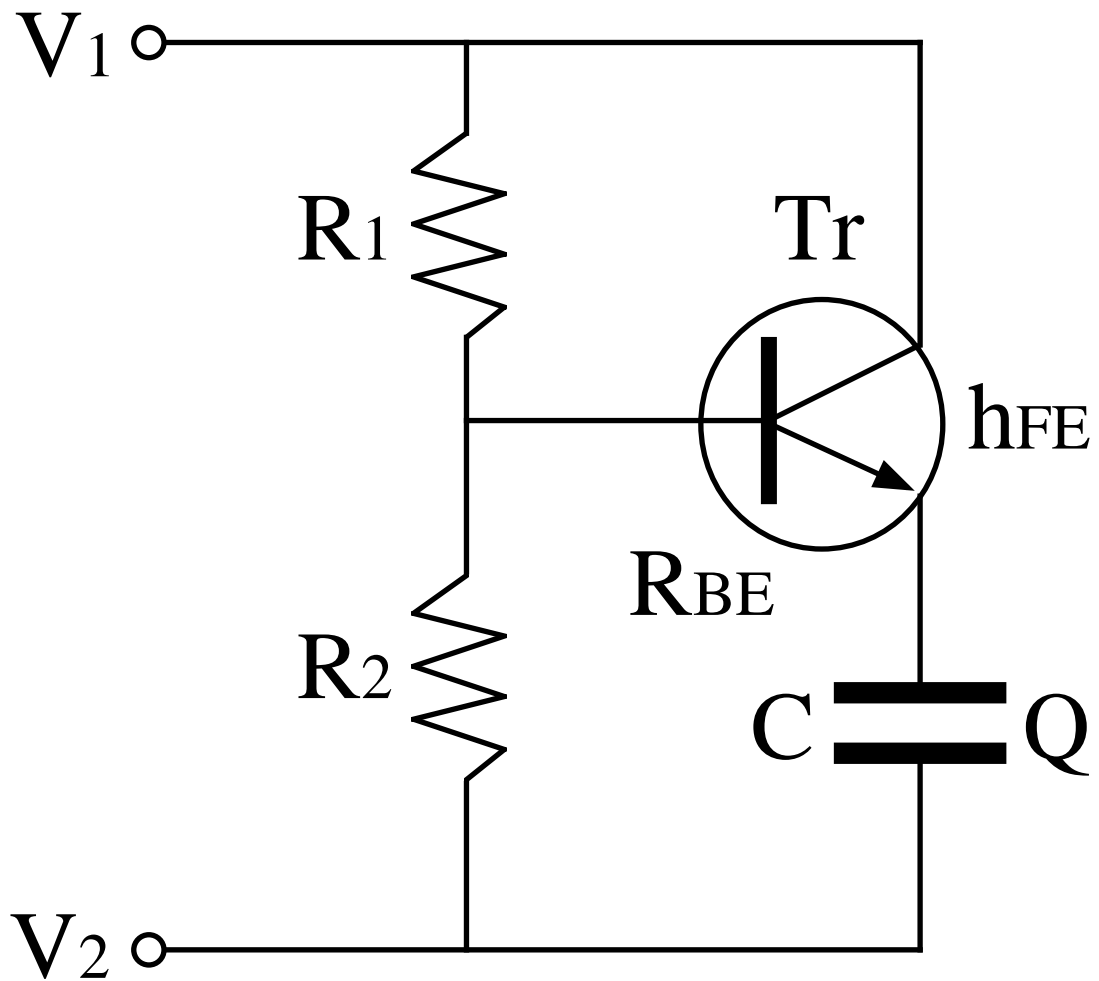Figure 2: Knowledge-Based Systems, Takashi Washio and Hiroshi Motoda

Figure 3: Knowledge-Based Systems, Takashi Washio and Hiroshi Motoda