ATLANTIS PRESS
Advances in Computer Science Research, volume 93
2nd International Conference on Mathematics, Modeling and Simulation Technologies and Applications (MMSTA 2019)

Knowledge Discovery in Evolution of the Structure and Form of Scientific Model: J-System Theory

Qi Qian[1], Qianhui Zhao[2], Yu Chen[3] and Zhaohua Jiang[4]
[1]Johns Hopkins University, Baltimore, USA
[2] Dongbei University of Finance and Economics, Dalian, 116021, China
[3]Neusoft group, Shenyang, 110000, China
[4]Institute of Science and Technology Management, Dalian University of Technology, Dalian, 116023, China

***Abstract*—First we establish the definition of J-System, and by some actual case according to this definition to illustrate that many scientific knowledge system are J-System which is consist of three concept sets by some cognitive pattern. This paper discusses that the J system is composed of the paradigm set and three basic concept sets. On this basis, the paradigm module and three basic concept modules are formed. The formal relationship between the four modules is revealed, which provides a new systematic method and text mining method for the model construction and knowledge discovery of scientific system and Visualization method.**

Keywords—knowledge discovery; the structure; form of scientific model

I. INTRODUCTION

A. Consistency of Scientific Models in Form of Structure

In the form of structure, it is exactly the same or similar between scientific models in many different fields, although their scientific principle, scientific thinking may be fundamentally different. For example, the spring-damper system, the mechanical rotation system, the circuit system are constituted by the inductors, resistors and capacitors .Although they are described as different systems, their model is identical that they share the same structural form $a\frac{d^2y}{dt^2}+b\frac{dy}{dt}+cy=F$, as shown in the table 1.

TABLE I. THE SIMILAR VARIABLES OF THE ELECTRO- MECHANICAL SYSTEM MODEL[1]

model type	Similar parameters						model
Mechanical translation system	Quality M	Damp N	Stiffness K	Displacement Y	speed V	Force F	$M\frac{d^2Y}{dt^2}+N\frac{dY}{dt}+KY=F$
Mechanical rotation system	Moment of inertia J	Damp β	Stiffness K	Angular displacement α	Angular velocity ω	Torque T	$J\frac{d^2\alpha}{dt^2}+\beta\frac{d\alpha}{dt}+K\alpha=T$
$R-L-C$ system	Inductance L	Resistance R	Capacitance $1/C$	Power Q	Current intensity I	Voltage U	$L\frac{d^2Q}{dt^2}+R\frac{dQ}{dt}+\frac{Q}{C}=U$

B. The Proposal of General Systems Theory and Its Problems

The American scholars R.L. Cansey who are engaged in a long-term research of scientific unity once said[2], people have been interested in the unified science which may provide a general description for a long time[3][4]:

(1) For the natural and the social science, we need to propose a system theory towards integration.

(2) This theory and its methods may become an important aid in solving some important issues in the scientific field.

(3) This theory has developed the unified theory which vertically through the individual fields of science that it get closer to the goal of the unified science[5] [6]..

C. J-system Theory and Knowledge Discovery

Jiang Zhaohua put forward the J-system theory which based on the research of scientific model's consistency, similarity and uniformity in structure and form [7] .

J-system consists of three set of concepts [A], [B] and [C] in accordance with cognitive pattern [P] ,which is not only the J-system theory's formal definition on the system, but also the starting point of this article. This paper firstly describes the definition and explains that many scientific knowledge systems are J-system according to this definition.

A knowledge system constituted by the three set of concepts according to some set of cognitive patterns; Then putting forward the concept of the module, and illustrating the relationship among the module sets, three concepts set, and the

cognitive patterns; And then summing up five formal structures and five evolutionary forms of scientific model, and discussing the mechanism of knowledge discovery and the problems in knowledge production main process at scientific modeling on this basis.

Knowledge discovery, also named knowledge discovery in databases (KDD) by some scholars[8], taps implicit, useful and previously unsuspected knowledge and information from a lot of initial data. Fayyad defined database knowledge discovery as "KDD is a non-trivial process which recognizes effective, novel, potentially useful and ultimately understandable patterns from the data sets[9]. In 1986[10], professor Swanson firstly found the existence of recessive association in medical literature, and he speculated the recessive association between A and C by means of the complementarity. For example, A represents the intake of a substance that can cause a physiological change B, and the physiological change B caused a particular organ disease C, and then gets an useful information that A has an effect on C, which cannot be found in a separate literature, by contacting B, we are able to dig out of this implicit association.

Mapping knowledge domain becomes an important method of the J-system theory. The software of CiteSpace which developed by Chen Chaomei of institute of Information Science and Technology in Drexel University[11], is a knowledge visualization tool based on the JAVA computer language and can be used to analyze model evolution and the structure of information, adopted a new generation of information visualization techniques, it fits for multiple, time-sharing, dynamic and complex network analysis. CiteSpace software provides key words co-occurrence analysis[12],author co-citation analysis, and literature co-citation analysis etc. visualization capabilities. In large numbers of scientific literatures, CiteSpace software provide possibility for scientists to find out three set of concepts on a specific subject (cognitive pattern), and provide an intuitive visual way for scientist to understand the locus of model evolution.

II. PRELIMINARIES

A. The Definition and Formal Structure of J-system

A general system is formed by sets $[A]$,$[B]$ and$[C]$ rather than one set. According to cognitive pattern [P], J-system consists of three set of concepts $[A]$,$[B]$ and $[C]$ as for different systems, because of their different cognitive pattern [P], hence the three set of concepts are different.

The specific [P] marks a specific system. For example, in Newton's system, the three set of concepts are $(m, f, (r, v, a))$, and the carrier of [P] is a particle, [P] regulates the contact of these three concepts is $f = ma$ (Newton's second law).

As for the set of concepts $[A]$,$[B]$ and $[C]$, the component elements possessed a certain relationship, sequence and structure. In general, the set of concepts are provided with dynamic, integral and orderly nature.

Thus, a J-system, if it is defined as (1) the concept set [A], the subset, main elements, elements etc. are written as a_i , $[A] = \{a_1, a_2, ..., a_n\}$; (2) the concept set [B], the subset, main elements, elements etc. are written as b_i , $[B] = \{b_1, b_2, ..., b_n\}$; (3) the concept set [C], the subset, main elements, elements etc. are written as c_k , $[C] = \{c_1, c_2, ..., c_n\}$; (4) some cognitive pattern [P] makes the three set of concepts contact as a system , and it consists of some set of concepts described as [P] $= \{p_1, p_2, ..., p_n\}$;

This is a completely formal definition, and "some " conveys the information that there may be various sorts of [P], each of the specific [P] corresponds to the concrete three set of concepts thereby corresponds to a sort of system.

Thus, the relationship between the formal system J and the concrete system J_0 can be seen as $[P]: J \to J_0$ (abbreviated as $J :\to J_0$)

B. The Case of J-system

Many scientific and technological knowledge systems are 4 sets, such as

1) Newton's knowledge

$$N = ([m, I], [f, M], [x, v, a]) \tag{1}$$

$[A] \to [m, I]$ is the inertia variable, and $[B] \to [f, M]$ is the dynamic variable, $[C] \to [x, v, a]$ is the kinematic variable.

2) Mechanical system

$$\mathrm{M} = ([M, N, K, J, \beta], [F, T], [y, \alpha, V, \omega]) \tag{2}$$

Where $[\mathrm{A}] \to [M, N, K, J, \beta]$ is the inertia variable. and $[\mathrm{B}] \to [F, T]$ is the dynamic variable. $[C] \to [y, \alpha, V, \omega]$, is the kinematic variable.

3) Ideal gas theory

Ideal gas theory model is $\mathrm{p} = \mathrm{nRT} / V$,$[A] \to \mathrm{nR}$ is Molar gas mass, and $[B] \to T$ is the temperature variables, $[C] \to 1/V$ is volume variable.

4) Control System

The general control system $= (x, [m, u, r], y)$,

where $[A] \to x, [B] \to (m, u, r), [C] \to y$,and

$$dx / dt = mx + u, \quad y = rx \tag{3}$$

To draw a conclusion, these are many cases and, in fact, the main models in all major field of science can be involved in the analyses of the J-System.

III. The Formalized Model in J-System

Scientific models in different fields are often exactly the same or similar in the form or structure, so beginning at a number of concrete scientific knowledge system on the structure and analyzing the form of model, J-system theory further abstracts the various types of formal models and sums them up into five kinds of simple formal models. In addition, any model can be regarded as one made up of three modules.

Figure 1 shows the modeling logic of Newtonian Mechanics.

A. Module

In general, modules A, B, C and P are codetermined by $[A],[B],[C]$ and $[P]$, and there is no simple corresponding relation between them[13]. However, in usual, A is determined by $[A]$ largely and is generated based on $[A]$. Similarly, B is determined by $[B]$ largely and is generated based on $[B]$; C is determined by $[C]$ largely and is generated based on $[C]$; P is determined by $[P]$ largely and is generated based on $[P]$.

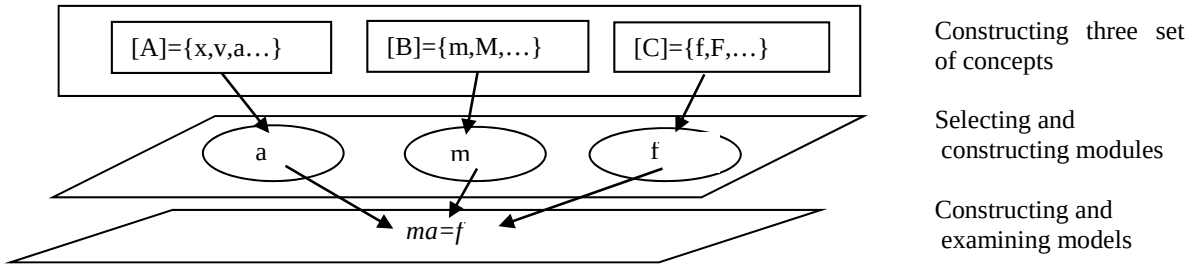

FIGURE I. THE MODELING LOGIC OF NEWTONIAN MECHANICS

B. The Wholes and Parts Model

There are totally 5 categories: the wholes and parts model, causal model, control model, optimization model and evolutionary model[14].

The wholes and parts model, its formal model is:

$$J = P(A, B, C) \tag{4}$$

Wherein (4), three more specific form is

$$J = Ao_1 Bo_2 C \tag{5}$$

$$J = A \pm B \pm C \tag{6}$$

$$J = A \bullet B \bullet C \tag{7}$$

From the perspective of natural science, the water flow model is a typical model of the wholes and parts:

$$V = (1/n) R^{2/3} S_o^{1/2} \tag{8}$$

Where V ——velocity of water, n ——Manning coefficient, R ——hydraulic radius, S_o ——Slope of tunnel foundation. These three modules are respectively $(1/n)$, $R^{2/3}$ and S [15].

Causal model

Causal model reflects the relationship between the three sets which constitute the system. Its formal model is:

$$BOA \Rightarrow C \tag{9}$$

where O is the mutual action between A and B.

Depending on different explanations for "O" and "$\Rightarrow$", formula $BOA \Rightarrow C$ has different forms, for instance:

$$A + B = C \quad \text{or} \quad A - B = C \tag{10}$$

In the general theory of relativity:

$$R_{ik} + uT_{ik} = \left(\frac{1}{2}\right) g_{ik} R \tag{11}$$

where $A \to R_{ik}, B \to uT_{ik}, C \to \left(\frac{1}{2}\right) g_{ik} R.$

For thermodynamic law

$$\mathrm{d}U + \mathrm{d}A = \mathrm{d}Q, A \to \mathrm{d}U, B \to \mathrm{d}A, C \to \mathrm{d}Q$$

$$B \cdot A = C \quad or \quad A = C / B. \tag{12}$$

The typical form for this is $ma = f$ (Newton's second law of motion).

$$B : A \to C \tag{13}$$

Specially, function is a useful mapping of number sets to number sets.

Control Model

The general control model is the following nonlinear system:

$$\frac{dx}{dt} = f(x,u,t)$$

$$y = g(x,u,t) \tag{14}$$

Further specific form of (16) is

$$\frac{dx}{dt} = f(x) + g(x)u$$

$$y = h(x) \tag{15}$$

C. Optimization Model

The general form of the optimization model can be written as

$$\begin{aligned} &\min A \\ &s.t. B \le C \end{aligned} \tag{16}$$

Where A is the optimization function, B and C are the constraints. Fuzzy logic systems [16], neural networks [17], genetic algorithms [18], particle swarm optimization, support vector machines, and rough sets are all based on the optimization model[19] [20].

D. Evolutionary Model

The general form of the evolutionary model can be written as

$$\frac{dA}{dt} = f(A,B,C) - g(A,B,C) \tag{17}$$

"f" represents the promotion and recognition to evolution, and "g" represents the suppression and denial to evolution.

Take the concentration of pollutants in environmental media model as example [21]. There are three main types of movement patterns after the pollutants entering into the environment. The three main types are the plug-flow migration along with the environmental media, dispersion of the pollutants particle, transformation and attenuation of the pollutants. If we treat the researched environment as a unit with boundary, Q represents the flow of environmental media, and reflecting the effect of plug-flow; S represents the total number entering into the environment; C_0 represents the original concentration of pollutants in environmental media; C represents the concentration of pollutants in environmental media; V represents the flow rate of environmental media. The Complete mixed model can be written as:

$$V\frac{dC}{dt} = QC_0 - QC + S + rV \tag{18}$$

Of the model above, $QC_0 + S + rV$ represents the promotion to evolutionary concentration of pollutants in environmental media, and $-QC$ represents the suppression to evolutionary concentration of pollutants in environmental media.

Summarizing our preceding discussion, the J-system model can be divided into three levels, the first level is the formal model, for example, $BOA \Rightarrow C$ and so on; the second level is the concrete of the formal model (explained model), in fact it's a second-rate or a third- rate formal model, such as $B \cdot A = C$ and $A + B = C$ or $A - B = C$ and so on; the third level is the case model, those who applied in various scientific models specifically and successfully, such as $ma = f$ (Newton's second law of motion), $\mathrm{d}Xu/dt = \lambda uXu + hu(Xu, Xs) + fu(t)$. (Haken's equation), and so on.

IV. Evolutionary Ways of Model

By means of the research towards a large number of evolutionary case of scientific models, many ways of evolutionary scientific models is summarized as following:

A. Analog Modeling

Analog modeling helps researchers with limited experience construct a reasonable model and solve the problem more quickly.

B. Inclusive Modeling

For example, the model of the ideal gas law:

$$PV = NKT \tag{19}$$

wherein, p-system pressure; V-volume; N-molar concentration; K-gas constant; T- temperature. and when founded later that the state equation of real gas:

$$(P + a/V^2)(V - b) = RT \tag{20}$$

That's state equation of Van der Waals, for the thin air under high-temperature and low pressure, as the comparison between a/V^2 and b, P and V can be neglected, so it closes to the ideal gas.

C. Transformation Modeling

For example, causal model changed into the control model. Considering the single-joint robot system [22], the dynamic system model is

$$J_m \frac{d^2 q(t)}{dt^2} + sg\sin(q(t)) = f(t) \tag{21}$$

Where $f(t)$ is the torque acting on the node, g is the acceleration of gravity, $q(t)$ is the angle of arm rotation,

let $$q = x_1, \frac{dq}{dt} = x_2, u(t) = f(t)$$

The system can be described as

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -J_m^{-1} sg\sin x_1 \end{pmatrix} + \begin{pmatrix} 0 \\ J_m^{-1} \end{pmatrix} u \tag{22}$$

we can see that the causal model changed into control mode by transformation.

D. Restructuring Modeling

For example [23], the original model of financial accounting in assets management is(24)

$$\begin{cases} x_{i+1} = s + pz_i \\ y_{i+1} = qx_i + ry_i z_i \\ z_{i+1} = 1 - x_i - y_i + z_i \end{cases} \tag{23}$$

In formula (24), s, u, q and r are parameters. s represents the sales base, u represents stock funds transfer rate, q represents product resources rate, r represents inventory efficient rate, x_i represents sales of resources in stage i, y_i represents the number of customers in stage i, z_i represents the company's stock capital in stage i.

$$\begin{cases} x'_{i+1} = s + pz'_i \\ y'_{i+1} = qx'_i + ry'_i z'_i - a(y_i - y'_i) \\ z'_{i+1} = 1 - x'_i - y'_i + z'_i \end{cases} \tag{24}$$

Model (23) was improved by (24), where a is a gain coefficient. In the effect of feedback, from the situation of Lyapunov exponent in chaos model of asset management changing with a, we can see when a is more than a certain value, the largest Lyapunov exponent is less than zero, the driven asset management chaos systems (23) and the respondent asset management chaos system (24) come to synchronization.

Further improved (25) as:

$$\begin{cases} x'_{i+1} = s + 2z'_i / 1 + \tanh(z'_i - g^{-1}) \\ y'_{i+1} = \tanh(qx'_i) + \frac{1}{2}[1 + \tanh(z'_i - g^{-1}]\tanh(qx'_{i-1}) - a(y_i - y'_i) \\ z'_{i+1} = 1 - x'_i - y'_i + z'_i \end{cases} \tag{25}$$

All actions observed in Model (25) are almost similar with the former model (24) of which with no transformation. We can also get the fold bifurcation will go to chaos in a very small parameter space.

E. Integration Modeling

This paper will define the evolutionary form of the model as integration modeling.

For example [24], the integration of control model and causal model, Study the seven degree of freedom of the whole vehicle model. Firstly, establish the motion equation of passive suspension and active suspension, then study the control of the active suspension and establish control model. Active control is based on the passive control adding actuator; the form of the equation is

$$M\ddot{x} + C'\dot{x} + Kx = B_1 u + B_2 \omega \tag{26}$$

Where M、C' and K is respectively the mass matrix, damping matrix and stiffness matrix of suspension system, u and ω is respectively the control vector and input vector, B_1 and B_2 is the matrix respectively related to controlling force and road inputs.

V. Concluding Remarks

In this paper, we were devoted to study evolution of the structure and form of scientific model. The major results obtained can be concluded as following four aspects:

(1)J-system theory has taken the measure of case analysis to study the composition of scientific knowledge systems thoroughly, and has found that many knowledge systems consist of three concepts sets in accordance with some cognitive patterns.

(2)It is often exactly the same or similar in the form or structure between Scientific models in different fields. And these models could be divided to three or four simply modules.

(3)Compared to some of the existing knowledge discovery methods, the advantage of knowledge discovery which is based on J-system is that the selection of experimental model (combined with various scientific functions and expressions) is not limited to polynomial function.

(4)Mapping knowledge domain offers a strong analysis tool for “four sets analysis” which is put forward by J-system methodology. J-system provides a new way to construct the Scientific Knowledge Discovery System.

Acknowledgments

Acknowledgments for the guidance and support of the following organizations. This paper is the result of a project of

the Henan Department of Transportation, a project of the State Intellectual Property Office of China, a key project of the Chinese Academy of Engineering, a project of the National Natural Science Foundation of China. Thanks for Waseda University professor Watada Junzo guidance.

REFERENCE

[1] Guo Qisheng Guo Xiaojun Li Guanghui Xue Qing. modeling principles and methods[M]. National Defense University Press ,2003

[2] Rudolf Seising.Cybernetics,system theory,information theory and Fuzzy Sets since 1950s and 1960s[J]. Information Sciences, 2010 (180):4459–4476.

[3] Haken, H. Advanced Synergetics: Instability Hierarchies of Self-Organizing Systems and Devices[M]. Heidelberg, New York: Springer-Verlag, 1993.

[4] Bertalanffy,L.V. General system theory--A criticial review[J]. General System,7:1-20.1962.

[5] Patricia Roma'n-Roma'n , Desir'ee Romero,Francisco Torres-Ruiz. A diffusion process to model generalized von Bertalanffy growth patterns: Fitting to real data[J]. Journal of Theoretical Biology 2010 (263) :59–69.

[6] Jiang.Z.H,From general system theory to J-system[J]. Seeking Truth, 1989,9(5):32-38.

[7] Jiang.Z.H, Jiang.Z.H, "J-Ssystem Reconstructability: A f ormalized Study," International Journal of General Systems. 2000. 29(3);, Int. J.General Systems 2000,29(3):363-371.

[8] Yang Kun, Chen Jun, Zhang Pei, Liang Yu. The knowledge discovery research of POL equipment based on regular taping s[J].Logistics Technology, 2001（4）:35-42.

[9] Bernstein A, Provost F J, Hill S. Toward intelligent assistance for a data mining p rocess: An ontology - based app roach for cost- sensitive classification[J]. IEEE Transaction on Knowledge and Data Engineering, 2005, 17 (4) : 503 – 518.

[10] Dong Fenghua Lan Xiaojun , Literature-based knowledge discovery tool- -Arrowsmith [J]. Journal of Information, 2004, （5）52-54.

[11] Chen C. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature[J]. Journal of the American Society for Information Science and Technology, 2006, 57（3）: 359-377.

[12] Chen Yue, Liu Zeyuan.The rise of mapping knowledge domain[J]. Studies In Science of Science 2005. 2(3); 99-120.

[13] Romer, P.M..Endogenous Technological Change[J]. Journal of Political Economy, 1990,98(5):71-102.

[14] Jones, C. R&D-based models of economic growth [J].Journal of Political Economy,1995,103 (4):759–784.

[15] Zhao Xiping.Research and practice of automatic Control and hydraulic simulation in the Yellow River Diversion Project [M]. Beijing: China Water and Power Press,2006.

[16] Shuming Wang, Junzo Watada.Modelling redundancy allocation for a fuzzy random parallel_series system[J].Journal of Computational and Applied Mathematics, 2009 (232) :539-557.

[17] Watada Junzo; Wang, Shuming; Pedrycz, Witold. Formulation of Fuzzy Random Regression Model[C].6th IEEE International Symposium on Intelligent Signal Processing , Budapest, HUNGARY , 2009 .

[18] Watada Junzo,Cisneros,Haydee Rocio Melo.Meta-Controlled Boltzmann Machine: Its Convergence and Application to Power System. World Automation Congress (WAC) [C].Puerto Vallarta, 2012.

[19] Junzo Watada,Ichiro Shigaki.A comparison of dispatching rules and genetic algorithms for job shop schedules of standard hydraulic cylinders[J].Soft Computing,January, 2008, 12(2): 121–128.

[20] Matsumoto Yoshiyuki, Watada Junzo. Time Series Data Analysis by Rough Set and Merging Method of Decision Rule[C].Joint 17th [1] World Congress of International-Fuzzy-Systems-Association / 9th International Conference on Soft Computing and Intelligent Systems (IFSA-SCIS) ,Otsu City, JAPAN , 2017.

[21] Cheng Shengtong.Tutorial of Environmental Systems Analysis[M].Beijing: Chemical Industry Press,2006.

[22] Yu Shaojuan Qi Xiangdong Wu Juhua, Iterative Learning Control Theory and Applications[M]. Beijing: Machinery Industry Press,2005.

[23] Song Xuefeng ,Yang Liexun,Cao Qingren,The development of Complex scientific study[M]. Beijing:CHINA SCIENCE PUBLISHING & MEDIA LTD., 2004.

[24] Yang shaopu,Shen Yongjun, Bifurcation and Singularity of hysteretic nonlinear systems[M].Beijing:CHINA SCIENCE PUBLISHING & MEDIA LTD., 2003.