INFORMATION SCIENCE & COMPUTING

International Book Series Number 3

Decision Making and Business Intelligence Strategies and Techniques

Supplement to
International Journal "Information Technologies and Knowledge" Volume 2 / 2008

ITHEA

INFORMATION SCIENCE & COMPUTING

International Book Series Number 3

Decision Making and Business Intelligence Strategies and Techniques

Supplement to
International Journal "Information Technologies and Knowledge" Volume 2 / 2008

ITHEA
SOFIA, 2008

Krassimir Markov, Krassimira Ivanova, Ilia Mitov (ed.)

Decision Making and Business Intelligence Strategies and Techniques

International Book Series "INFORMATION SCIENCE & COMPUTING", Number 3

Supplement to the International Journal "INFORMATION TECHNOLOGIES & KNOWLEDGE" Volume 2 / 2008

Institute of Information Theories and Applications FOI ITHEA

Sofia, Bulgaria, 2008

This issue contains a collection of papers in the field of Decision Making and Business Intelligence. Papers are selected from the International Conferences of the Joint International Events of Informatics "ITA 2008", Varna, Bulgaria.

International Book Series "INFORMATION SCIENCE & COMPUTING", Number 3
Supplement to the International Journal "INFORMATION TECHNOLOGIES & KNOWLEDGE" Volume 2, 2008

Edited by **Institute of Information Theories and Applications FOI ITHEA**, Bulgaria,
in collaboration with
- **V.M.Glushkov Institute of Cybernetics of NAS**, Ukraine,
- **Institute of Mathematics and Informatics, BAS**, Bulgaria,
- **Institute of Information Technologies, BAS**, Bulgaria.

Publisher: Institute of Information Theories and Applications FOI ITHEA, Sofia, 1000, P.O.B. 775, Bulgaria.
Издател: Институт по информационни теории и приложения ФОИ ИТЕА, София, 1000, п.к. 775, България
www.ithea.org, www.foibg.com, e-mail: info@foibg.com

General Sponsor: **Consortium FOI Bulgaria** (www.foibg.com).

Printed in Bulgaria

ISSN 1313-0455 (printed) ***ISSN 1313-048X (online)*** ***ISSN 1313-0501 (CD/DVD)***

TABLE OF CONTENTS

INDEX OF AUTHORS

METHOD OF ELIMINATION DATA SOURCES CONFLICTS IN INFORMATION SYSTEM

Viktor Levykin, Maksim Evlanov

Abstract: *Use of modern object-oriented methods of designing of information systems (IS) both descriptions of interrelations IS and automated with its help business-processes of the enterprises leads to necessity of construction uniform complete IS on the basis of set of local models of such system. As a result of use of such approach there are the contradictions caused by inconsistency of actions of separate developers IS with each other and that is much more important, inconsistency of the points of view of separate users IS. Besides similar contradictions arise while in service IS at the enterprise because of constant change separate business-processes of the enterprise. It is necessary to note also, that now overwhelming majority IS is developed and maintained as set of separate functional modules. Each of such modules can function as independent IS. However the problem of integration of separate functional modules in uniform system can lead to a lot of problems. Among these problems it is possible to specify, for example, presence in modules of functions which are not used by the enterprise to destination, to complexity of information and program integration of modules of various manufacturers, etc. In most cases these contradictions and the reasons, their caused, are consequence of primary representation IS as equilibrium steady system. In work [1] representation IS as dynamic multistable system which is capable to carry out following actions has been considered:*

Keywords: *information system, dynamic system, information gene, metamodel, data sources conflict, synergetic, conflicts elimination*

ACM Classification Keywords: *H.3.4 Systems and Software*

Conference: *The paper is selected from Sixth International Conference on Information Research and Applications – i.Tech 2008, Varna, Bulgaria, June-July 2008*

Introduction

Use of modern object-oriented methods of designing of information systems (IS) both descriptions of interrelations IS and automated with its help business-processes of the enterprises leads to necessity of construction uniform complete IS on the basis of set of local models of such system. As a result of use of such approach there are the contradictions caused by inconsistency of actions of separate developers IS with each other and that is much more important, inconsistency of the points of view of separate users IS. Besides similar contradictions arise while in service IS at the enterprise because of constant change separate business-processes of the enterprise. It is necessary to note also, that now overwhelming majority IS is developed and maintained as set of separate functional modules. Each of such modules can function as independent IS. However the problem of integration of separate functional modules in uniform system can lead to a lot of problems. Among these problems it is possible to specify, for example, presence in modules of functions which are not used by the enterprise to destination, to complexity of information and program integration of modules of various manufacturers, etc. In most cases these contradictions and the reasons, their caused, are consequence of primary representation IS as equilibrium steady system. In work [1] representation IS as dynamic multistable system which is capable to carry out following actions has been considered:

- to perceive the information;
- to remember the information;
- to generate the information.

In addition to the specified properties, IS according to similar representation should: to use the information for achievement of the purpose; processing the information to take from it valuable. Thus it is necessary to note, that

here and in the further it is necessary to understand the term "macroinformation" as the term "information" - initially casual, and then the remembered choice of one or several carried out variants from all set possible and equal in rights [2]. In this connection actual there are problems of development of methodology, methods and models of designing and operation IS in which basis representation IS as dynamic multistable system lies.

Information gene concept

The decision specified above problems demands development of special models which allow to describe IS and business-processes of the enterprises at the general level. As such description authors have offered the concept of an information gene (IG). Under IG the compressed sequence of knowledge (rules) of construction IS as a whole and its separate elements [3] is understood generally ordered and in a strong degree. Inherently IG is metamodel IS - the formalized representation which defines syntax and semantics of concrete realizations IS and its components [4]. Though the term "metamodel" in this treatment has got narrower sense, its essence has remained the same, as in the traditional theory of systems where existence of the metadescription and a metamodel is both necessary and constructive [5]. In the most comprehensible way of the organization of a metamodel realization of some mechanism providing the separate organization of work of the mechanism of a logic conclusion and the mechanism of interpretation of results of a conclusion [6] is represented.

The analysis of the basic approaches to modeling IS has shown, that static models of information structures are local, can display only separate fragments studied IS and change during generation of the new information. For creation and modernization IS as basic it is necessary to consider models of operations which should carry out IS. Such operations, as follows from representation IS in the form of dynamic multistable system, concern to three basic types:

- a subset of new information generation attitudes;
- a subset of information perception attitudes;
- a subset of information storage attitudes.

It is necessary to note, that the operations realizing given subsets, are carried out only above separate elements of models of information structures (attributes) or, accordingly, above separate information structures (clusters information space IS). Therefore in the further such operations we shall name operations of a microlevel.

Defined allows to allocate the uniform alphabet of modeling of operations above the information, saving up all variety of an embodiment of these operations in concrete models IS and its providing complexes.

Realization of laws of composition IS from separate functional modules which consist of clusters set information space IS and the operations of a microlevel certain on these structures, demands allocation of a class of operations which possess following features:

- the given operations are carried out above all or a part of attributes, clusters information space IS and operations of a microlevel of all IS, or the concrete functional module;

  the given operations do not depend on a concrete kind of descriptions of attributes, clusters information space IS and operations of a microlevel of all IS, or the concrete functional module;
- the given operations are constant within the limits of all IS, or the concrete functional module.

Such operations in the further we shall name operations of a macrolevel [7]. Then formation of metamodel IS on the basis of the concept of IG will be carried out on the basis of the following interconnected formalized representations:

- the description of separate clusters information space IS;
- the description of operations of a microlevel which reflect interrelations between separate elements clusters information space;
- the description of a subgroup of operations of the macrolevel providing integration separate of clusters information space and microlevel operations, certain on these clusters, in the form of transformations of metamodel IS during which its structure and the maintenance changes;

- the description of a subgroup of operations of a macrolevel in the form of displays of metamodel IS in a metamodel and model of providing complexes IS.

Revealing of conflicts as a result of the analysis of an information gene

Use of the concept of IG allows considering processes of designing IS as synergetic interaction of separate elements of system. Such approach will allow solving a problem of revealing and elimination of the contradictions arising because of ambiguous perception of a subject domain by participants of the project. For revealing these conditions it is offered to present IG as oscillatory system. Thus fluctuations in the given system are generated by performance of operations of generation of the new information and extend in IS as a result of performing reception operations and storages of the information. In such systems without dependence from their nature Fermi-Paste-Ulam (FPU) - the phenomenon of distribution of energy of initial indignation on the maximum harmonics with the subsequent gathering in a spectrum of initial indignation is observed. Especially important property of return FPU has appeared presence of "memory" in its spectrum in relation to entry conditions of its active fashions [8]. With reference to IG supervision of return FPU means, that values of the separate attributes forming clusters information spaces IS, after unitary generating many times are used separately or in the most various combinations, and can be if necessary returned in a condition observed at performance of operation of generation of the new information. Differently, if for studied IG return FPU is observed, it is meant, that with the information, which is entered into corresponding system, can be divided into separate components and subsequently is reproduced without mistakes. Thus process of functioning IS represents sequence of operations on processing the information. In the event that for studied IG return FPU is not observed or observed with distortions, it means, that the same information is duplicated in IS and during distribution is mutually absorbed or deformed. Thus there is an opportunity in advance to reveal areas of information space IS in which conflicts concerning data will be observed.

Model of the conflict of sources of data

Elimination of conflicts of sources of data in this case is reduced to the analysis of a phase portrait of system which is described by expression [2]

$$\begin{cases} \dfrac{dat_e(t)}{dt} = \dfrac{1}{\tau_e} at_e - a_e(at_e) - b_{ei} at_e at_i - a_e(at_e)^2 \\ \dfrac{dat_i(t)}{dt} = \dfrac{1}{\tau_i} at_i - a_i(at_i) - b_{ie} at_i at_e - a_i(at_i)^2 \end{cases}$$

where at_e and at_i - attributes of data, causing conflict; τ_e and τ_i - typical time of value reproduction of attributes at_e and at_i respectively; a_e and a_i - coefficient of process of attributes value generation deceleration of at_e and at_i respectively; b_{ei} and b_{ie} - coefficient of antagonistic influence of attribute values of at_e and at_i respectively.

It should b marked that coefficients b_{ei} and b_{ie} are, as shown in [1], coefficients of antagonistic interaction of attributes at_e and at_i. Physical meaning of coefficient b_{ei} in this case is an estimation of usage in the moment of new information generated values, percepted from attribute at_i, instead of values, percepted from attribute at_e.

So, physical meaning of coefficient b_{ie} is an ability estimation of usage in moment of new information generation values, percepted from attribute at_e, instead of values, percepted from attribute at_i. Physical meaning of coefficient a_e is estimation of possible reception of information, stored in attribute at_e, instead of information, received from attribute at_e. Physical sense of coefficient a_i is estimation of reception information ability, stored in at_i instead of information, received from attribute at_e.

The physical sense of factor τ_e consists in an estimation of relative time of generation of new values of attribute at_e which analogues are not present in aggregate values of attribute at_e. The physical sense of factor τ_i consists in an estimation of relative time of generation of new values of attribute at_i which analogues are not present in aggregate values of attribute at_i.

Proceeding from offered interpretation, it is possible to draw a conclusion, that factors b_{ei} and b_{ie} define a degree of independence of a source of data at work with the macroinformation - the less value of factor b_{ei} for attribute at_e, the it is less probability of that during generation of the new information a source of data will use the another's information from other attribute at_i. The factor a_e in this case defines a degree of novelty of the information generated by a source of data, and, indirectly, a degree of stability IS on the given attribute - the it is less value of the given factor, the above probability of that the generated information will not be percepted from stored values of attribute and. Hence, that above the probability of that after generation of similar information IS will leave an equilibrium condition. The factor τ_e in this case is an estimation of time of generation by a source of data of new value of attribute at_e - the less value of the given factor, the there will be less often new values and, hence, the less probability of output IS from an equilibrium condition.

Results of such analysis allow making such conclusion:

a) If b_{ei} is equal to b_{ie}, and a_e is equal to a_i, advantage receives attribute with lesser generation time

b) If τ_e is equal to τ_i, and a_e is equal to a_i, advantage receives attribute, which has lesser value of coefficient b in expression, describing process of value generation for this attribute

c) If b_{ei} is equal to b_{ie}, and τ_e is equal to τ_i, advantage receives attribute with lesser value of coefficient a in expression, which describes process of value generation for this attribute

Method of conflicts elimination

In such case it becomes possible use a specified below **method of conflicts elimination** for resolving conflicts of data sources, which occurs when there are two or more sources of same data.

Step 1. If it is possible to determine values of all coefficients τ_e, τ_i, a_e, a_i, b_{ei} and b_{ie}, go to Step 9. Else go to step 2.

Step 2. If it isn't possible to determine values of all coefficients τ_e, τ_i, a_e, a_i, b_{ei} and b_{ie} take a decision of lack of data about informational system and propose Developer to gather additional information in process of recurring solution task of selecting customer variant of IS configuration for attributes, involved in situation of plural generation of same information. Else go to step 3.

Step 3. In case of impossibility of determination values of coefficients τ_e, τ_i, b_{ei} and b_{ie} go to step 11. Else go to step 4.

Step 4. In case of impossibility of determination values of coefficients τ_e, τ_i, a_e and a_i go to step 12. Else go to step 5.

Step 5. In case of impossibility of determination values of coefficients b_{ei}, b_{ie}, a_e and a_i go to step 13. Else go to step 6.

Step 6. In case of impossibility of determination values of coefficients τ_e and τ_i take there values equal to 1/2 and go to step 9. Else go to step 7.

Step 7. In case of impossibility of determination values of coefficients b_{ei} and b_{ie} take there values equal to 1 and go to step 9. Else go to step 8.

Step 8. In case of impossibility of determination values of coefficients a_e and a_i take there values equal to 1/2 and go to step 9.

Step 9. Count regions of attributes stability using expressions $at_e = 0$, $at_i = \frac{1/\tau_e - a_e - a_e at_e}{b_{ei}}$ (vertical isoclines) и $at_i = 0$, $at_e = \frac{1/\tau_i - a_i - a_i at_i}{b_{ie}}$ (horizontal isoclines).

Step 10. Compare areas of calculated areas of stability and propose instead of operation of new information generation for attribute, for which area of stability is lesser, perform operation of information reception, generated by attribute, for which area of stability is bigger.

Step 11. Compare values of coefficients a_e and a_i and propose instead of operation of new information generation for attribute, with higher coefficient value, perform operation of information reception, generated by attribute, for which value of attribute is lesser.

Step 12. Compare values of coefficients b_{ei} and b_{ie} and propose instead of operation of information generation for attribute, which has higher value of this coefficient, perform reception of information, generated by attribute with lesser coefficient value.

Step 13. Compare values of coefficients τ_e and τ_i and propose instead of generating new information for attribute with higher coefficient value, perform reception of stored information generated by attribute with lesser value of coefficient.

Conclusion

Developed method could be used on different stages of IS development, beginning from process of creating and analytics requirements for IS, which is created, and its functional structure. Developed method allows to estimate and solve the task of data sources conflicts elimination in process of consulting and business-process reengineering.

Bibliography

1. Чернавский Д.С. Синергетика и информация (динамическая теория информации. М.: Едиториал УРСС, 2004. 288с.
2. Information theory in biology. /Ed. by H. Quastler. Urbana: University of Illinois Press, 1953.
3. Левыкин В.М., Евланов М.В. Выявление несоответствий в модели гена информационной системы // Proceedings of the International Conference "e-Management & Business Intelligence", Varna. Sofia: Institute of Information Theories and Applications FOI ITHEA. 2007. P.75-77.
4. Фаулер М., Скотт К. UML в кратком изложении. Применение стандартного языка объектного моделирования. М.: Мир, 1999. 191 с.
5. Месарович М., Такахара Я. Общая теория систем: математические основы. М.: Мир, 1978. 312 с.
6. Лачинов В.М., Поляков А.О. Информодинамика, или Путь к Миру открытых систем. СПб.: Издательство СПбГТУ, 1999.
7. Евланов М.В. Использование положений общей теории систем для моделирования процессов разработки информационных систем организационного управления // 2-й Международный радиоэлектронный форум «Прикладная радиоэлектроника. Состояние и перспективы развития» МРФ-2005. Сборник научных трудов. Том III. Межд. конференция «Информационные системы и технологии». Харьков: АНПРЭ, ХНУРЭ, 2005. С.67-70.
8. Гаряев П.П. Волновой геном. М.: Общественная польза, 1994. 280 с.

Authors' Information

Viktor Levykin – *Head of Information Control Systems (ICS) chair, Kharkiv National University of Radio Electronics, Lenin av. 14, Kharkiv, 61166, Ukraine; e-mail: iyc@kture.kharkov.ua*

Maksim Evlanov – *Information Control Systems (ICS) chair, Lenin av. 14, Kharkiv, 61166, Ukraine; e-mail: evlanov_max@mail.ru*