Semantic Intellectual System Development

Igor Boyko, Victor Martynov[1]
Publishing Systems and Solutions Laboratory
HP Laboratories Palo Alto
HPL-2001-219
September 10th , 2001*

E-mail: igor_boyko@hp.com

knowledge representation, semantics, SIS

Knowledge representation (KR) is one of the most important sub-fields of the Artificial Intelligence field. This article considers the perspective of using the USC (Universal semantic code) KR model for semantic intellectual systems (SIS) development and demonstrates main principles of using USC for KR. Besides, since the USC gives a possibility of inventive problem solving by semantic tools we compare USC tools with tools from the Theory of inventive problem solving (TIPS) and explain of its combination profit.

* Internal Accession Date Only Approved for External Publication?
[1] European Humanity University, 24 F. Skorina Ave., Minsk, Belarus
To be presented at the Seventh International Conference on Information Networks and Systems, Technologies, October 2-4, 2001, Minsk, Belarus

SEMANTIC INTELLECTUAL SYSTEM DEVELOPMENT

Igor Boyko Ph.D.

Hewlett-Packard Labs
1501 Page Mill Road
M/S 1L-15
Palo Alto, CA 94304, USA
phone (650) 857-5295
e-mail: igor_boyko@hp.com

Prof. Victor Martynov

European Humanity University
24 F.Skorina ave.
Minsk, Belarus
phone (375 17) 234-4614

ABSTRACT

Knowledge representation (KR) is one of the most important sub-fields of Artificial Intelligence field.
This article considers perspective of using of the USC (Universal semantic code) KR model for semantic intellectual systems (SIS) development and demonstrates main principles of USC for KR.
Besides, since the USC gives a possibility of inventive problems solving by semantic tools we compare USC tools with tools of Theory of inventive problem solving (TIPS) and explain of its combination profit.

Key Words: Knowledge representation, semantics, universal semantic code, inventive problem

USC CONCEPT

The USC is a model of knowledge representation and conversion where knowledge representation is realized through knowledge conversion. Such property is only of the USC property that gives the possibility to obtain knowledge from knowledge.
This property is supported by next factor: all conversions in the USC are syntactic, i.e. depends on positions of variables in USC formula [4; 6].
If a, b, c are positions of X, Y, Z variables then conversion:

(X * Y) * Z ==> X * (Y * Z)

where:
* -- is a binary operation of coincidence on the given set of variables and it does not mean of its positions changing because their order is kept:

(a * b) * c ==> a * (b * c)

although consecution of operations is changed.

At the same time the conversion:

(X * Y) * Z ==> (Y * Y) * Z

demonstrates of changing of position of variable

or

(a * b) * c ==> (b * b) * c

the consecution of operations is kept.

You can see that USC has own algebra that uses the given set of variables.

The type of the USC algebra of formulas representation and conversion has been determined:

A = < **M** , * , - >

Where:

- **M** -- is a set of elements;
- * -- is a binary operation on the given set (the operation of coincidence);
- - -- is an unary operation on the given set the (operation of junction).

The main unit of Knowledge Representation in the USC is a complex formula that represents of variables interaction. Semantic relations in complex formulas are represented in the following way:

((X*Y)*Z)*(Z*(Y*Z)) -- X by means of Y acts on Z and as a result Z keeps Y in itself (in Z).

We would like to draw attention to the fact that the left part in this type of formulas ((X*Y)*Z) causes potential action in right part (Z*(Y*Z)). Thus complex formula reflects the following situation:

Stimulus → Reaction.

Each complex USC formula has one and only one meaning and each conversion of USC formula into another one has one and only one meaning conversion. The rules according to which meaning conversion is done are contained in the USC algebra axioms.

USC formulas represent actions that you can meet in any domain. The USC allows to structure for example the technical actions database insert it into Knowledge Base and enclose semantics in representation formalisms.

Technical part of the USC-classifier contains technical actions list and USC formulas classified in a certain way. In the USC-classifier all technical actions are analogues. Each main technical action has own USC formula that determines this technical action and verbal interpretation demonstrating the action at issue. Each action-analogue has own main action.

See classifier structure:

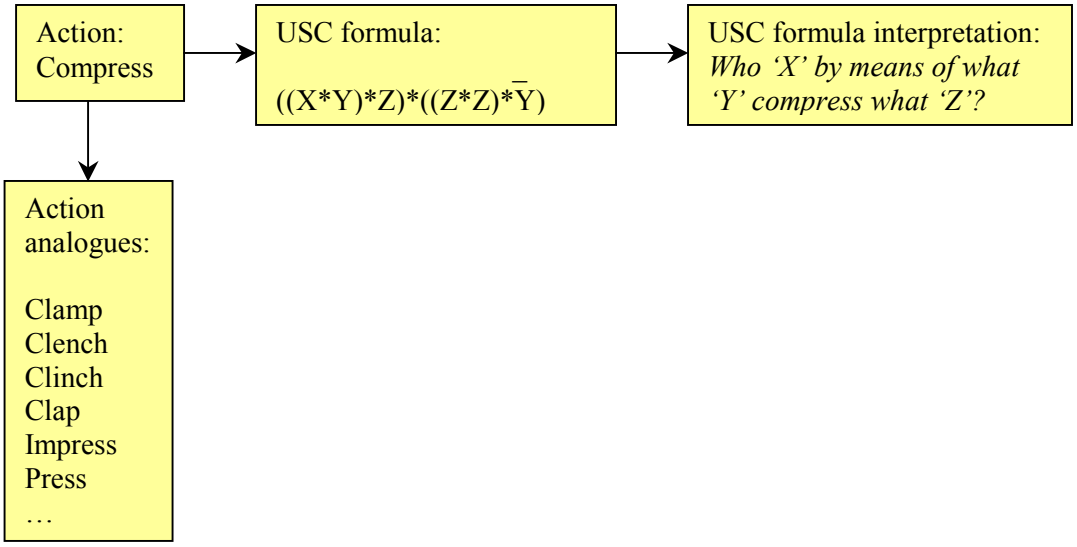


Fig.1

Practically, within the limits of the given classifier the non-traditional system of technical actions-analogues is built according to the **function they perform but not according to their achievement technology.**

In computer intellectual systems for description of problem conditions necessary to use strictly structured of KR units.

USC formulas have single sense, strictly given length and limited set variables filled by specific dates. Involuntary sequences of USC formulas define the problem solving variants.

In the USC verbs classifier used in technology verbs are presented and own interpretations in the USC formulas manner are gave. Since each verb expresses the relation appearing between subject and object of action in the USC classifier structures of proposed relations are presented in evident type. Such structures are knowledge representation units.

In USC classifier all verbs are main or inherit. Every USC formula corresponds to main verb and each USC formula has own verbal interpretation. Each inherit verb corresponds to the main verb. Practically main verb is the most abstract presentation for a certain technical action.

For example, the verb 'heat' is the main and is presented by USC formula:

heat → ((X*Y)*Z)*((Z*Z)*Y) -- 'X' by means of 'Y' heat 'Z'.

The verb 'measure out' is related with the verb 'measure' that is the main and is presented by USC formula:

measure → ((X*Y)*X)*((X*Y)*Z) -- 'X' by means of 'Y' measure 'Z'.

The verbs ‘clamp’, ‘clap’, ‘press’ and so on are inherit verbs of verb ‘compress’ (see Fig.1).

Practically, within the framework of given verbs classifier is built off-center technical actions-analogues classifier.

Syntactic diagnostics of conversion of variables allows formalizing by tools of USC algebra.

For comparing, the programming language PROLOG has the predicate logic in the basis but in the USC all conversions for variables are axioms of the USC algebra.

Besides for object-oriented technology USC axioms could be represented as objects of classes.

2. **SIS Development**

All problem-oriented conversions within specific domain are some kinds of effects:

- Physical
- Chemical
- Geometrical
- Informational
- Economical
- Psychological

and so on

All specific effects can be summarized in abstract effects. Semantic intellectual system (SIS) can use such effects and USC formulas for its representation. Actually the SIS is empty and opened system.
The SIS has unlimited possibilities for accumulating abstract and specific effects for using in absolutely different domains [1; 4; 6; 7; 8; 9;10].
We can propose of several stages of SIS development.

Stage 1.
Empty SIS (Fig.2) is oriented graph. Notes of graph are USC formulas. Links between notes are USC axioms.
USC algebra operates with a set of axioms that present conversion of elements of natural language (especially verbs) from one semantic meaning to another [4].

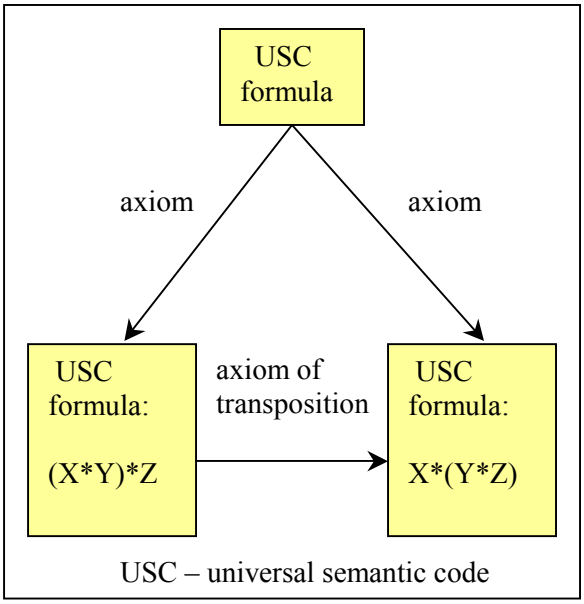


Fig.2

Stage 2.
Unit of abstract effects and processing rules is added (Fig 3).

For example, we can use next USC formula for presenting one of the types of substance phase changing:

((X * Y) * Z) * (Z * (W * Z)) -- X changes the phase of Z (transforms in a gas)

In the chapter Abstract Effects this approach will be considered more detail.

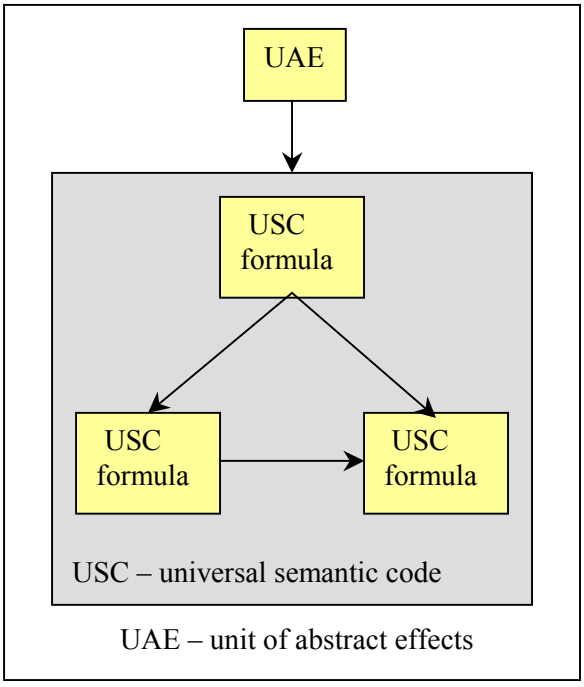
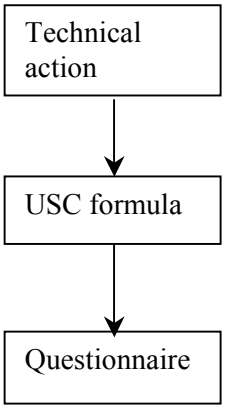


Fig.3

Stage 3.
The SIS comprises specific effects and classifier of technical actions (Fig.4).

Every technical action can be represented by next way:



Example of USC formula for verb ‘separate’:

((X * Y) * Z) * (Z * (W * Y))

Example of questionnaire for verb ‘separate’:

Who?	By means of what?	Separate what?	From what?
X	Y	Z	W

Every position of the table is a name of real object of action. The positions can be filled by user or automatically by some tagging and parsing system on basis of user’s problem state.
The solution of intellectual problem can be realized as a way in the view of connected links.

The first link is connected with the USC formula of initial state of problem. The last link is connected with the USC formula of target state of problem.
The way of inventive problem solving consists of USC formulas are presented by verbs of natural language.

Now the SIS has three levels.
On the first level the USC verbs classifier is placed.
On the second level we placed the unit of abstract effects that connected with the unit of specific effects.
On the third level the unit of formulas and axioms of the USC is placed.
Solving of inventive problem begins with the initial state of problem on the first level. On the third level the SIS processes the target state of problem.

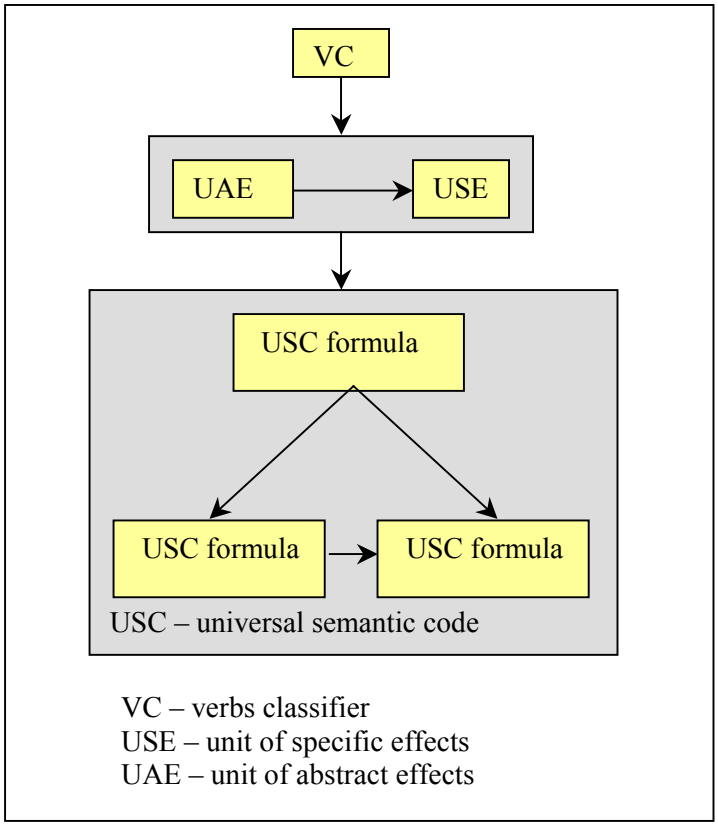


Fig.4

At last the **simple solution** of inventive problem is the way:

initial state → specific effect → target state.

Stage 4.
Using of graphics (Fig.5).
The SIS comprises nouns that represent objects of actions. Every object has relation with it graphical image.

When user will write names of objects in correspondence with positions in USC formula the computer will combine images of objects in correspondence of combining rules of objects for every specific action.

The object-image base must comprise rules for limiting of voluntary placing of objects.
For example, if hammer drives nail then in a space the wide part of hammer must be combined with the nail-head.

As a result the user must look both language and graphical representation of problem solving.

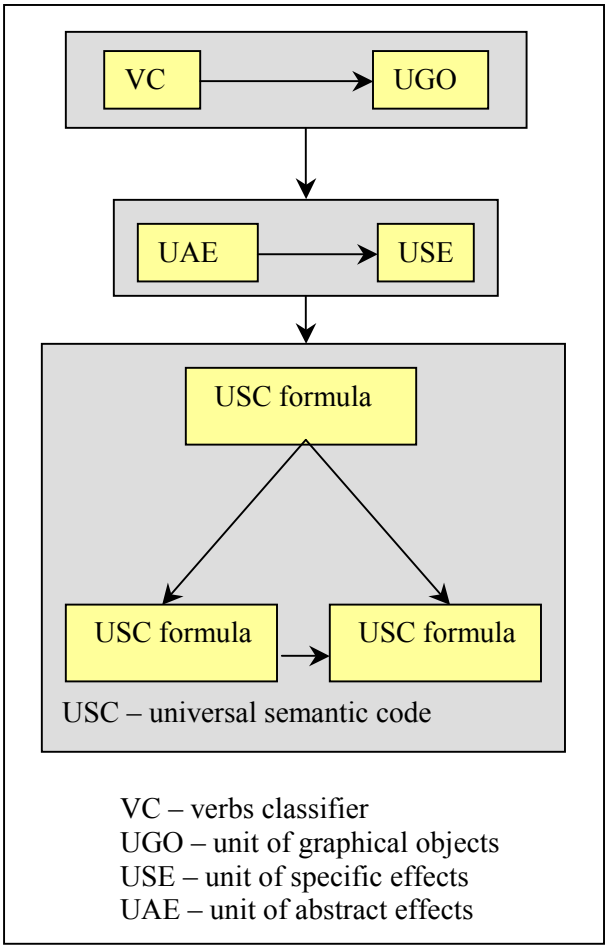


Fig.5

ABSTRACT EFFECTS

Effect representing class of specific effects named abstract effect (AE).
All kinds of specific effects (SE) can be collected into classes of AE.
It is true not only for effects but for different kinds of information from literature in general. Any types of documents from different fields can be classified by extracting of different kinds of AE from them.

Let's analyze effect of change of phase: 'the water boils under 100 degree C'. AE for it is the change of phase or in other words the change of aggregate state or of molecular structure.

The empty SIS where links represent USC axioms can consist of abstract effects only. Every AE can comprise set of specific effects.

USC AND TIPS

Theory of inventive problem solving (TIPS) [3] has system of standards of inventive problem solving. The system USC formulas represents AE and correlates with the system of standards in the TIPS.

Their differences are:

1. The system of standards in the TIPS was developed empirically on the basis a lot of patents. Such system cannot be completed. The USC is a deductive system that allows avoiding incompleteness in other words has a property of **completeness** of elements.

2. Standards in TIPS are represented by the non-canonized natural language with the ambiguity of its words because the ambiguity is the main property and problem of natural language. Therefore the inventions from different domain duplicate each other because they were written by different words from different specific industrial domains. In a result similar inventions in principles of action are placed in different standards. The USC system of AE is represented by canonized natural language and can not has any duplication. In other words the USC system of AE has a properties of **consistency** and **independence**.

Since the USC system of AE has all characteristics of deductive system:

- **Completeness**
- **Consistency**
- **Independence**

it can be computerized effectively.
The deficiency of such characteristics creates problems of computerizing of TIPS but evolution of the TIPS depends on possibilities of computerizing. For solving the problem necessary to combine USC and TIPS tools for development unique system of AE and semantic classification of SE.

Let's we represent some of standards of TIPS by formulas of USC:

Change of phase (change of microstructure)

a) ((X * Y) * Z) * (Z * (W * Z)) -- X changes the phase of Z (transforms in a gas)

b) ((X * Y) * Z) * (Z * (W * Z)) -- X changes the phase of Z (transforms in a liquid)

c) ((X * Y) * Z) * (Z * (W * Z)) -- X changes the phase of Z (transforms in a solid)

One USC formula represents AE that can be specified by real kinds of substances: gas, liquid or solid.

Non-phase change (change of macrostructure)

a) ((X * Y) * Z) * ((Z * W) * W) -- X changes the property of Z (the color)

b) ((X * Y) * Z) * (Z * (W * W)) -- X changes the property of Z (the form)

The presented examples are only part of one of classes of TIPS standards.
The intermediate goals of phase and non-phase changes can be considered as 'defending' from unwanted effect or 'reorientation' of unwanted effect to another object.

'Defending' indicates of changing that initiates of appearing of:

a) The filter for transmitting only needed substance.
b) The defending layer for transmitting only needed field.
c) The isolation for isolating both substance and field.

'Reorientation' indicates of changing that initiates of appearing of:

a) New objects of unwanted effect (direct action).
b) New subject of opposed effect (opposed action).
Besides the tools of 'reinforcing' of useful result can be used.

'Reinforcing' indicates of changing that initiates of introducing of additional tool-mediator.
Comparing of classification of standards in the TIPS with semantic classification in the USC demonstrates necessity of more detail investigation in this field.
Preliminary results can be represented as oriented graph of AE with classification of SE in it for using by SIS.

CONCLUSION

We have the theoretical base for development semantic intellectual systems in particular of inventive systems. But semantic Knowledge representation and management by USC tools could be used for solving of problems of different kinds:

- Inventive problem
- Natural language processing and understanding
- Documents processing and classifying
- Sense understanding and calculation
- Information search in Intranet and Internet
- Modeling spoken and written language
- Independence from kind of natural language
- Machine translation
  and more

References

1. Boyko I.M. Innovation in the Knowledge Representation Concept. // *The Second All-Union Conference "Artificial Intelligence--90". Minsk.* 1990. (in Russian).
2. Martynov V.V. Development of Languages of Semantic Representation and Conversion of Knowledge. // *The Second All-Union Conference "Artificial Intelligence--90". Minsk.* 1990. (in Russian).
3. Altshuller G.S. Small Unlimited Worlds. Standards for Inventive Problems Solving. // *Guiding Thread into Labyrinth. Petrozavodsk.* 1988. (in Russian).
4. Martynov V.V. Universal Semantic Code: USC-3. Preprint N2. Institute of Linguistics of Academy of Science of Belarus. Минск. 1984. (in Russian).
5. Martynov V.V., Boyko I.M., Gyminski A.P. Knowledge Bases Construction of Systems for Solving Intellectual Problems. // *Controlling Systems and Machines. Kiev.* 1992. N5/6.
6. Martynov V.V. Universal Semantic Code: USC-5. *Pre-print N4. Minsk State Linguistic University.* 1995.
7. Boyko I.M. Semantic Method of Knowledge Bases Development. // *Practical Aspect of Language Functioning. Association of Communicative Linguistics of Belarus. Minsk.* 1995. (in Russian).
8. Martynov V.V. Principles of Objective Semantic Classification. // *Practical Aspect of Language Functioning. Association of Communicative Linguistics of Belarus. Minsk.* 1995. (in Russian).
9. Boyko I.M. Computer Semantic Search of Inventive Solutions. // *TRIZ Journal. USA. March, 2001.*
10. Jackson P. Introduction to Expert System. Addison Wesley Longman Limited. England. 1999.