Some APL Defined Functions for Fast Analysis of Innovations in Food Science and Human Nutrition

—by B. Makeev and A. Zoueva
Moscow, Russia

THIS PAPER DESCRIBES THE METHODS of using the Expert System "Forecaster" for information intelligence, information support of managers' decisions and fast analysis of innovations.

Keywords: Analysis of situation, APL, cloud of terms, decision support system, expert system, food science & technology, forecasting, human nutrition, innovation, map of science, tendency vectors

Introduction

This paper presents some practical results of more then four years of a sophisticated work aimed at elaborating, upgrading and using the Expert system "Forecaster." Its first commercial version was developed (the system is upgraded continuously!) for supporting a number of R&D programs of Russian nuclear power industry. But at the present time we found many new applications! In particular, "information intelligence," "information support of managers' decisions" and "fast analysis of innovations" in food science and human nutrition. Figure 1 illustrates current and planning knowledge areas where we are trying to use the Expert system.

These tasks are helpful to science and technology managers if they face the following problems before them [1]:

i) what are the right ways to distribute material, financial, human and other resources?

ii) what are the right ways to define and to keep on monitoring additional branches of science, required for achieving the assigned objectives?

iii) how can we find an environment where innovation is expected?

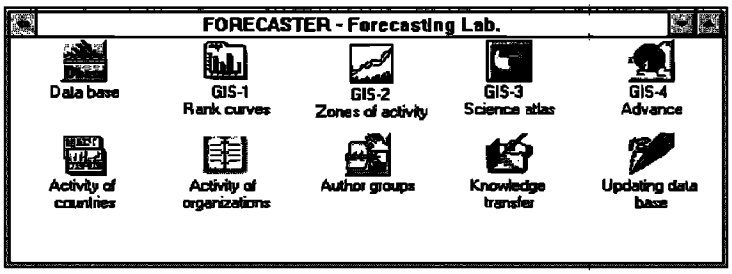

Figure 2: Functional parts and blocks of the Expert system

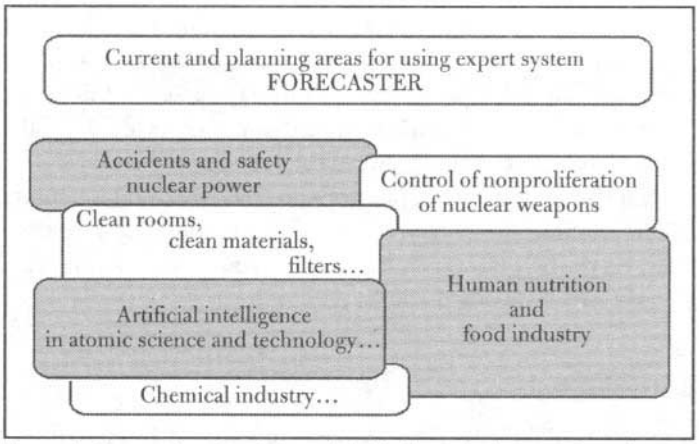

Figure 1: Current (gray) and planned (white) knowledge areas for using the Expert system

The Expert system will assist some managers in finding full or partial answers to these and other questions. Figure 2 shows the basic capabilities of the Expert system. This figure illustrates all systems' functional parts.

All these functional parts of the Expert system are written using APL★PLUS III for Windows. Each of the functional part has same functional blocks which are APL defined functions. In this paper we will discuss the methods of "knowledge interpretation."

At the origin...

Man-elaborated scientific terminology comprehensively and completely reflects physical phenomena and processes that the researcher studies and uses in his every-day practice [2]. The Scientific Community collaborates and applies universally the system of postulates that regulate conceptual, theoretical, implementable and methodological ways to reach the targets in a given branch of science. We can add that the mechanisms of the evolution of science paradigm is found in the Bible, as is shown in Figure 3.

The New American Bible.
CHAPTER 1.
The Primeval History.
1. In the beginning, God created the heavens and the earth.
5. God called the light "day," and the darkness He called "night."
8. God called the dome "the sky."
10. God called the dry land "the earth" and the basin of water he called "the sea."
19. So the Lord God formed out of the ground various wild animals and various birds of the air, and He brought them to the man to see what he would call them; whatever the man called each of them would be its name.

Figure 3: Bible. Chapter 1. The Primeval History

The business of innovation and surfing the data streams

Paul M. Cook, the founder and CEO of Raychem Corporation, says that there is no secret in being an innovative company. "You assemble a group of talented people who are eager to do new things and put them in an environment where innovation is expected. It's that simple—and that hard." He thinks that any innovation is much about sales or service or information systems as it is about products. Too many companies are only immersing in their markets. They bring along whatever technology they think is necessary to satisfy a market need. Then they fall flat on their faces because the technology they deliver isn't sophisticated enough or because they don't know what alternatives the competition can deliver. There are at least three principles of innovation. First, intellectual property is the absolute key. Second, technologies are becoming more complex and interdependent. Third, innovation is a global game—both on the supply side and the demand side. Mr. Cook thinks that every company is innovative—otherwise it isn't successful. It's just a question of degree [3].

But how any small or big company can become innovative? Howard Rheingold, a writer and an editor based in Sausalito, near San Francisco, has written such books as "The Virtual Community." He stresses that there are going to be a lot of advantages in having access to the information networks and through them to the different databases.

"Clearly there are going to be class divisions in that. And we're seeing the beginnings of "symbolic analysts"—people like journalists who manipulate ideas, who need a lot of streams of ideas, date and information sources, and develop skills for deciding what's useful [4]. Most importantly, with global information networks and different databases, the finding of the projects, patents, technical reports, articles with old and new ideas, conference reports, etc. becomes easier and becomes imperative if one wants to join the community of the networks [5].

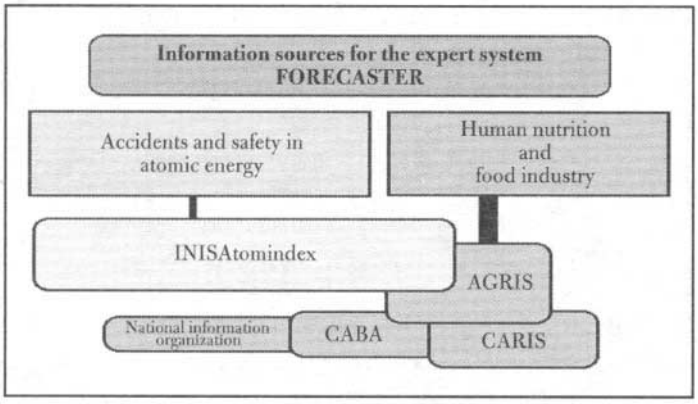

Figure 4: Information sources of the Expert system

The Expert system under discussion can use any information sources in different branches of science. As Figure 4 shows we are using INIS Atomindex (any flows of documentary information on atomic energy), AGRIS & CARIS & CABA (any flows of documentary information on agricultural, food science & technology and human nutrition) information sources and other CD-ROM databases by SilverPlatter.

Information intelligence

It is evident from the foregoing, that "any managers and any symbolic analysts of any innovative companies" need data streams for retrieval, selection and determination of useful ideas, etc. To do and to make the Expert system work properly, they usually form the initial terminological formulas that can be presented as follows [6]:

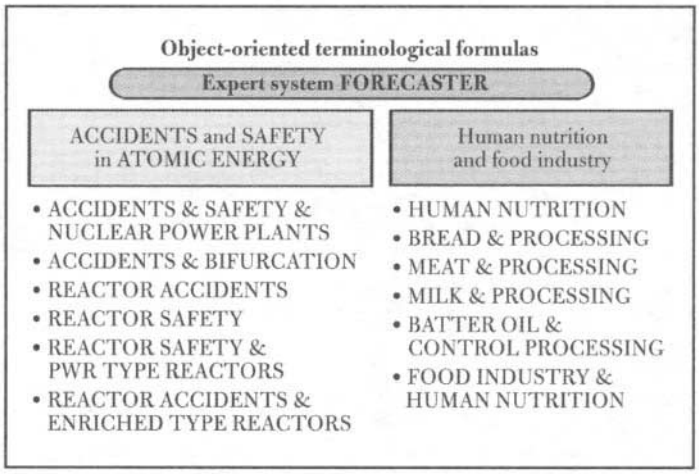

Figure 5: Examples of some terminological formulas used in the Expert system

$$Tf = Fun\ (T_1, T_2, \ldots, T_n, AND, OR, NAND, NOR)$$

where $T_1, T_2, \ldots, T_n$ — scientific terms.

Figure 5 gives some examples of terminological formulas used in the Expert system. These formulas are object-oriented.

In this case we can do "the information intelligence," i.e., we can find the relevant science and technology subbranches as well countries and companies. Figure 6 and Figure 7 illustrate the interface of the Expert system for forming terminological formulas, and APL defined function supports this interface.

As it can be seen from Figure 7, the explicit result "VEC" of this APL defined function "Z4" is the terminological formula. It may be useful to other APL defined functions like the ones for the creation of "maps of science," "vectors of tendency," "advanced diagrams," "Zipf's diagrams" [7].

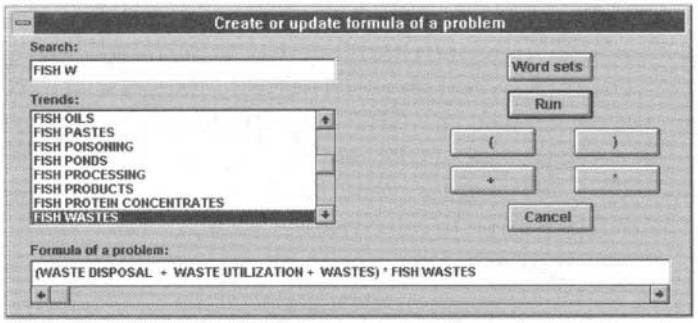

Figure 6: Interface of the Expert system for forming terminological formulas

```
    ∇ Z4;ALF;NK;I;na;n;a1;f;f1;ff;MAT;k;t;MM;W;a
[1]   ⍝ programma working formula of a problem
[2]     ALF←'ABCDEFGHIJKLMNOPQRSTUVWXYZ'⋄ ⎕PP←10 ⋄ NN←0
[3]     NK←I←0 ⋄ MM←0 ⋄ na←↓(⎕FSIZE N)[2]+31 ⋄ MAT←0 41⍴' '
[4]   CL1: t←1↑⍴(⎕FREAD N, (5+NK)) ⋄ MM←MM+t ⋄ I←I+1 ⋄ NK←31×I ⋄
          →CL1×⍳I<na
[5]     a1←⍴MM←MM⍴0 ⋄ ⎕ELX←'→CL10'
[6]   :if ((+/Tf='(')≠(+/Tf=')'))
[7]     CL10: Mes1
[8]   :else
[9]     LIST←W←0⍴' ' ⋄  I←1 ⋄ k←0
[10] AAA: →CL3×⍳(Tf[I]='(')=1 ⋄ →CL4×⍳(Tf[I]=')')=1
[11]    →CL5×⍳(Tf[I]='+')=1 ⋄ →CL6×⍳(Tf[I]='*')=1
[12]    →CL7×⍳(Tf[I]=' ')=1
[13] CL: W←W,Tf[I] ⋄  I←I+1 ⋄ →CL9×⍳I>⍴Tf
[14]    →CL×⍳(((Tf[I]='+')∨(Tf[I]='*')∨(Tf[I]='(')∨(Tf[I]=')'))=0)
[15] CL9: →CL10×⍳(⍴W)>31 ⋄ →CL13×⍳(⍴W)≤30 ⋄ W←30↑W
[16] CL13: W←W,((30-⍴W)⍴' ') ⋄ t←1↑⍸,(ALF∘.=1↑W)
[17]    →CL11×⍳((+/((⎕FREAD(300+N),t)∧.=W))=0)
[18]    n←1↑⍸((⎕FREAD(300+N),t)∧.=W)⋄ n←(t×100000)+n ⋄ NN←NN,n
[19]    n←1↑⍸(⎕FREAD (400+N),4)∘.=n ⋄ n←,(⎕FREAD(1100+N),n) ⋄ MM[n]←1
[20]    LIST←LIST,(⍒MM) ⋄ W←((8-(⍴(⍒+/MM)))⍴' '),(⍒(+/MM),('- ')),W ⋄
        MAT←MAT,[1]W
[21] CL12: f ⎕wi 'list' (MAT) ⋄ f ⎕wi 'font' 'Arial Cyr' 0.9 1 ''
[22]    ff←'fmMes' ⎕wi 'Show' 0  →CL8
[23] CL11: W←(⍉(⎕fread 55 4)[6;])⍴W ⋄ MAT←MAT,[1]W ⋄ LIST←LIST,(⍒MM) ⋄
          →CL12
[24] CL3: LIST←LIST,('(') ⋄ →CL7
[25] CL4: LIST←LIST,(')') ⋄ →CL7
[26] CL5: LIST←LIST,('v') ⋄ →CL7
[27] CL6: LIST←LIST,('*')
[28] CL7: I←I+1
[29] CL8: MM←a1⍴0 ⋄ W←0⍴' ' ⋄ →AAA×⍳I≤(⍴Tf) ⋄ ⎕ELX←'→CL10' ⋄ t←+/⍋LIST
[30]    'fmMes1' ⎕wi 'Delete'
[31]    f1 ⎕wi 'translate' 1 ⋄ f1 ⎕wi 'font' 'Arial Cyr' 0.8 1 ''
[32]    f1 ⎕wi 'caption' ((20↑((⎕fread 55 1)[10;])),(⍒t))
[33]    f1 ⎕wi 'where' 11 7 1 35 ⋄ f1 ⎕wi 'color' 255 255 255 255 0 0
[34]    ⎕wself←'fmMes.bn1' ⎕wi 'New' 'Button' ('where' 13 10 1.5 8)
        ('caption''OK')
[35]    ⎕wi 'style' 1 ⋄ ⎕wi 'Focus' ⋄ ⎕wi 'onClick' 'Exit_Save'
[36]    ⎕wself←'fmMes.bn2' ⎕wi 'New' 'Button' ('where' 13 30 1.5 8)
        ('caption''Cancel')
[37]    ⎕wi 'onClick' 'Exit_mes1' ⋄ f←'fmMes' ⎕wi 'Wait'
[38] :end
    ∇
```

Figure 7: APL defined function for supporting the formation terminological formulae interface

Manager's decision information support

The Expert system is based on a familiar fact that between citation frequencies of science terms and processes of science information rising, developing and aging it take place certain correlation. Figure 8 shows three of the most common correlation curves. The first curve shows the typical process of information aging without any events. The second curve demonstrates "the information explosion" and then "drastically decrease" of publications. The third curve illustrates "the unexpected termination" of publications.

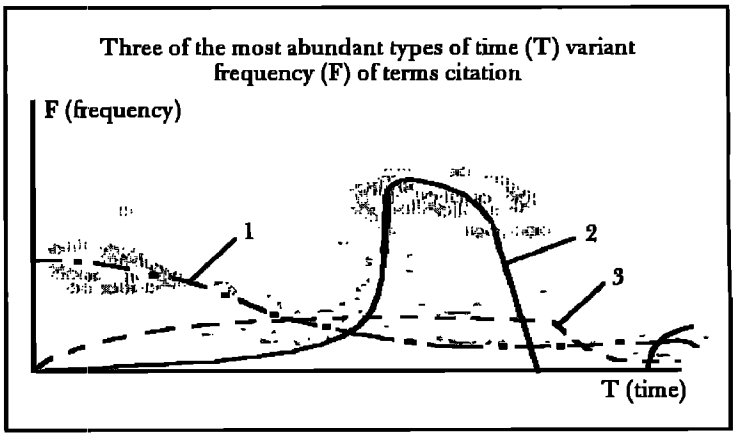

Figure 8: Curves are showing the different information processes

Relying on the knowledge of data streams behaviour, all science notations of "cloud of terms" may be classified in six zones by their vectors of tendency. To make such classification, we need to compare Zipf's (rank) curves of the current interval of time and the past intervals of time as well between frequency of terms' citation at the present time and in the past. Figure 9 shows an example of Zipf's (rank) curves. The methods of Zipf's curves plotting have been detailed in [8]. Table 1 gives some relations between the ranks and frequencies of notations as well the vectors of tendency' classifications.

When the classification of notations from data substreams is made by APL defined function "*VECTOR*" presented in Figure 10, the expert system initiates a program for counting the intensity of those vectors. The explicit result "*V*" of this function is the value of the vectors of tendency.

We can see the results of this function's work on the screen as shown on Figure 11, for different problems of food industry.

Relations between the ranks and frequencies of science notations and the 'vectors of tendency' classifications

Relations *	Zone of tendency vectors
Rc − Rp < 0	
Fc − Fp > 0	interest growth
Rc − Rp = 0	
Fc − Fp > 0	interest growth
Rc − Rp > 0	
Fc − Fp > 0	interest decay
Rc − Rp > 0	
Fc − Fp = 0	interest decay
Rc − Rp = 0	
Fc − Fp = 0	instability
Fc − Fp = 0	new interests

* Rc, Rp, Fc, Fp — the ranks and the frequencies of the current and the past time interval

Table 1

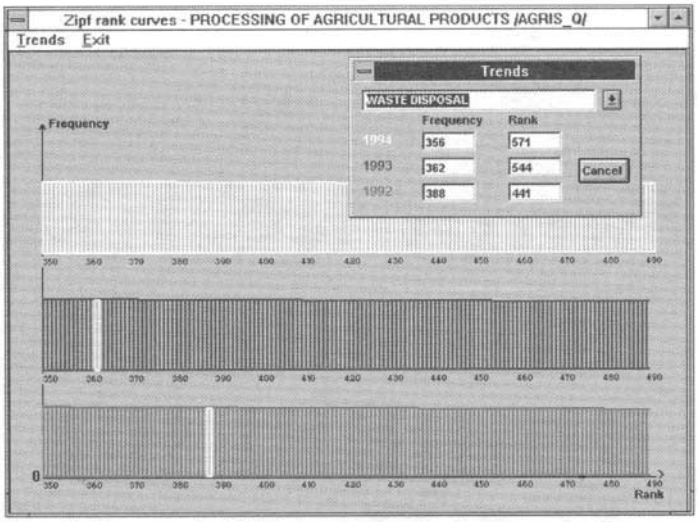

Figure 9: The example of Zipf's (rank) curves for three intervals of time

This function and such pictures of the expert system allow the users to obtain some answers to two questions: "What places do my scientific problems occupy among others in the knowledge area?" and "What is the intensity and the tendency of my scientific problems?"

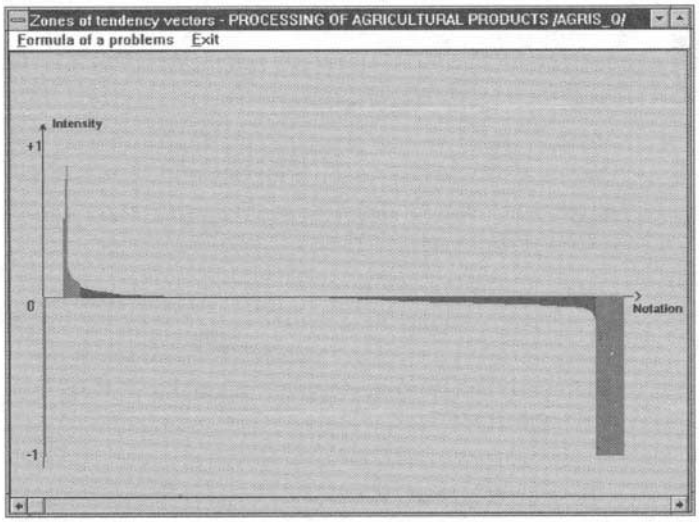

Figure 11: Example of the picture of tendency vectors for different problems of food industry

Fast analysis of innovations

When the user wants to find an answer about the right ways to distribute material, financial, human and other resources, the right ways to define the science branches which need support and how to find an environment where the innovation is expected, he may use the so-called "maps of science" [2].

With a view of demonstrating the methods of "maps" application we will use them for solving the above problem, that is, the problem of waste disposal and utilization in food industry [9].

```
    ∇ v←VECTOR;f;CC;sh;vp;vm;RR;FF;NN;R5;R12;R34;RVSE;N5;N12;N34;G...
[1]   ⍝ function for classification of notations from data substrems
[2]   ⍝ and counting the intensity of tendency vectors
[3]      vp←vm←RR←0 ⋄ FF←NN←0 ⋄ R5←R12←R34←RVSE←0 ⋄ N5←N12←N34←NVSE←0
[4]      G←⎕FREAD(N+400),1 ⋄ RM←⎕FREAD(N+400),2 ⋄ VV←⎕FREAD(N+400),4
[5]      C12←C34←C5←C6←COL←0 3⍴0 ⋄ →KKK×⍳(+/+/(⎕TS-2)[1]≤G)=0
[6]   ⍝ new
[7]      CC←A←,(RM[(⍳1-1);]=0) ⋄ sh←⍕((⍴⎕fread (400+N),4)-(+/A))
[8]      sh←((⍳((⎕fread 55 1)[39;])),⍕II),(' ( ',sh,' )')
[9]      'fmOpen.lab2' ⎕wi 'caption' sh
[10]     →MMM×⍳(+/A)=0 ⋄ Ft←A/[1],(RM⍳i;]) ⋄ Nn←A/[1]VV ⋄ I←1 ⋄ V6←0
[11]  CL:→CCC×⍳Ft[I]=0 ⋄ V←(Ft[I]-1)+Ft[I]
[12]  CC2:V6←V6,V ⋄ C6←C6,[1](0 128 0) ⋄ I←I+1 ⋄ →CL×⍳I≤(⍴Ft)
[13]     CC←Nn←Nn[⍋V6] ⋄ V6←V6[⍋V6] ⋄ RVSE←RVSE,V6 ⋄ NVSE←NVSE,Nn ⋄
         COL←COL,[1]C6
[14]  ⍝ 1.1  1.2  1.3  1.4  1.5
[15]  MMM:X←1 ⋄ Nn←(A←(|(VV∘.CC)-1))/[1]VV ⋄ Ft←A/[1],RM[i;] ⋄
         Fp←A/[1],RM[(i-1);]
[16]     V←⍒Ft ⋄ V1←⍒Fp ⋄ Nn←Nn[V] ⋄ Fp←Fp[V] ⋄ Ft←Ft[V] ⋄ ⎕PP←3
[17]  CL3:→CCC3×⍳(⍴V)=0 ⋄ Rp←(1↑⍒(V[X]∘.=V1)) ⋄ F←X-Rp ⋄ A←Ft[X]-Fp[X]
[18]     →((F<0)∧(A>0))/CL2 ⍝ 1.1.
[19]     →((F=0)∧(A>0))/CL4 ⍝ 1.2.
[20]     →((F>0)∧(A>0))/CL1 ⍝ 1.3.
[21]     →((F>0)∧(A=0))/CL8 ⍝ 1.4.
[22]     →((F=0)∧(A=0))/CL9 ⍝ 1.5.
[23]  CL5: 'fmOpen.ed2' ⎕wi 'text' (⍕X) ⋄  mu←⎕DL 0.001 ⋄ X←X+1 ⋄
         →CL3×⍳X≤pV
[24]     Z←0 ⋄ →CCC1×⍳(⍴R12)=0 ⋄  A←R12[(1↑⍒(vp∘.=(⌈/vp)))]
[25]     C12←C12[(⍒R12);] ⋄ N12←N12[⍒R12] ⋄ R12←R12[⍒R12]
[26]  CCC1: →CCC2×⍳(⍴R34)=0 ⋄ Z←R34[(1↑⍒(vm∘.=(⌈/vm)))]
[27]     C34←C34[(⍒R34);] ⋄ N34←N34[⍒R34] ⋄ R34←R34[⍒R34]
[28]  CCC2: →CCC3×⍳(⍴R5)=0 ⋄ C5←C5[(⍒R5);] ⋄ N5←N5[⍒R5] ⋄ R5←R5[⍒R5]
[29]  CCC3:RVSE←RVSE,R12,R34,R5 ⋄ NVSE←NVSE,N12,N34,N5
[30]     RVSE ⎕FREPLACE(400+N),6 ⋄ NVSE ⎕FREPLACE(400+N),7
[31]     →CCC4×⍳(⍴R12)=0 ⋄ C←(R12≤A)∧(R12≥0) ⋄ C←(+/C)↑(⍒(C=1)) ⋄ I←1
[32]  CL6:C12[(C[I]);]←(128 0 0) ⋄ I←I+1 ⋄ →CL6×⍳I≤(⍴C)
[33]  CCC4:→CCC5×⍳(⍴R34)=0 ⋄ C←(R34≤0)∧(R34≥Z) ⋄ C←(+/C)↑(⍒(C=1)) ⋄ I←1
[34]  CL7:C34[(C[I]);]←(255 0 128) ⋄ I←I+1 ⋄ →CL7×⍳I≤(⍴C)
[35]  CCC5:COL←COL,[1]C12,[1]C34,[1]C5 ⋄ COL ⎕FREPLACE(400+N),10 ⋄
         II←II+1 ⋄ →0
[36]  CL8:v←(Rp-X)+X ⋄ R34←R34,v ⋄ N34←N34,(Nn[X])
[37]  MM1:C34←C34,[1](64 0 128)
[38]  MM2:F←Rp←X ⋄ vm←vm,A ⋄ →CL5
[39]  CL9: v←¯1×(1+Ft[X])
[40]     R5←R5,v ⋄ C5←C5,[1](0 128 255) ⋄ N5←N5,(Nn[X]) ⋄ F←Rp←X ⋄ →CL5
[41]  CL1: v←((Rp-X)+X)×(1+(Ft[X]-Fp[X])) ⋄ R34←R34,v
[42]  MM3:C34←C34,[1](64 0 128)
[43]  MM4:N34←N34,(Nn[X]) ⋄ vm←vm,A ⋄ →CL5
[44]  CL2: v←((Rp-X)+Rp)×((Ft[X]-Fp[X])+Ft[X]) ⋄ R12←R12,v
[45]  MM5:C12←C12,[1](255 0 0)
[46]  MM6:N12←N12,(Nn[X]) ⋄ vp←vp,A ⋄ →CL5
[47]  CL4:v←(Ft[X]-Fp[X])+Ft[X] ⋄ R12←R12,v
[48]  MM7:C12←C12,[1](255 0 0)
[49]  MM8:N12←N12,(Nn[X]) ⋄ vp←vp,A ⋄ →CL5
[50]  CCC:V←Ft[I]-1 ⋄ →CC2
    ∇
```

Figure 10: APL defined function for classification of notations from data substreams and counting the intensity of tendency vectors

First we'll build a map of "common problems of waste handling technology in food industry." The notation "WASTE DISPOSAL OR WASTE UTILIZATION OR WASTES" is suitable. The map of this notations include 1032 links. Therefore, it's very difficult to analyse. Now we shall try to build a map for narrower subjects related to this problem, namely, "(WASTE DISPOSAL OR WASTE UTILIZATION OR WASTES) AND FISH WASTES." Now we have only 83 links and can begin to interpret them. Figure 12 shows the map.

As it can be seen from the map, the common waste handling technology in fish industry is related to many other technologies and science branches as well as countries: FISH PRODUCTS, SHELLFISH, CHEMICAL COMPOSITION, STORAGE, RELATIVE HUMIDITY, ASIA, JAPAN, etc. At this step of the analysis we can evaluate an internal connection between the "core" of the map and any branches, number of links and documents.

Furthermore, we can compare the map of the international sciences and technologies branches with the national ones. Now we have only 83 links and we can begin to interpret them.

By comparing the internal structure of the fish industry international waste treatment technology problems with our own problems, we may find innovative or old solutions.

In order to find the place of these technology problems among other branches properly, we may use the APL defined function "*VECTOR*". Figure 13 shows the resulting picture of the common problems of waste treatment technology in fish industry. The arrow on this picture is pointing at the place of the problem—in this case, decay of interests (database from 1962 to 1996).

Conclusion

The practical methods, which have been discussed above, allow us to conclude that the Expert system "Forecaster" may be a good instrument of the information support of strategical and tactical manager decisions in different scientific and technological knowledge areas. But in this paper we have shown only a small part of these methods provided by the system; you may find many other methods in the bibliography given below. ■

***Boris Makeev** and **A. Zoueva** work at REDStars, Inc., RF MinAtom, in Moscow, Russia. They can be reached at "makeev@ai.ru".*

Acknowledgments

The authors are deeply indebted to Yuriy Tkachuk, Evgeniy Postnikov, Alexandr Aldoshin (Russian Ministry for Atomic Energy), Valeriy Kalinin, Alexey Samsonov, Liubov Snopkova (Central Scientific Research Institute for Atomic).

References

[1] Makeev, B., and Zoueva, A.: "APL-Graphics application for Maps of Science Construction in Expert System "Forecaster-E" used for scientific forecasting in Atomic Science and Technology (II)." FUZZY sets and systems. Special issue of*IFSA*, 1994

[2] Makeev, B., and Zoueva A.: "Necessity is the mother of invention, or APL-Graphics application for Maps of Science."*APL Quote Quad*, Vol. 25, No. 1, pp. 145–151, 1994

[3] Taylor, W.: "The Business of Innovation. An Interview with Paul Cook."*Harvard Business Review* March–April 1990, pp. 97–106

[4] Avthur, C.: "The future of work: It's all in the mind.*New Scientist*, April 1994, pp. 28–31

[5] Handy, C.: "Making Sense of The Future,*Harward Business School*, 1994

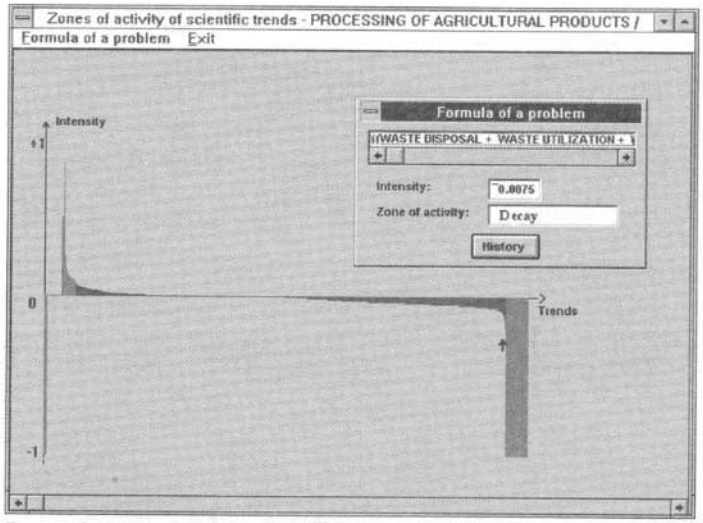

Figure 13: Example of the picture of tendency vectors for "common problems of waste treatment technology in fish industry"

[6] Makeev, B., and Zoueva, A.: "Elaboration of Forecast system for the separate scientific branches in atomic science and technology analysis."*TechnoRef*, 1, US, 1994

[7] Makeev B.A., Zoueva A.V.: "APL-Graphics application for Maps of Science Construction in Expert System 'Forecaster-E' used for scientific forecasting in Atomic Science and Technology (I)."*The International Conference Proceedings of FLINS'94*, 1994

[8] Kalinin, V., and Makeev, B., and Zoueva, A., and Pakhomov, A.: "APL as a Tool for Scientific Forecasting." *The International Conference on APL, Conference Proceedings*, pp. 175–182,1992

[9] Makeev, B.: "Some questions of the finding of functional dependencies among the scientific branches on the base of the analysis of information streams." Russian scientific conference "Safety of life people" materials. St. Petersburg, pp. 66–68, 1994

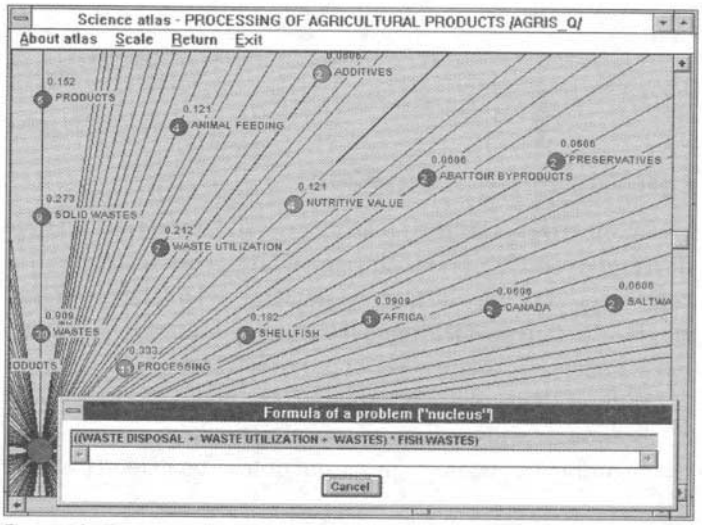

Figure 12: The map for "common problems of wastes' technology in fish industry"