

## Dialogue Computer System BIKEY as applied to diagnostics of Cnidaria (illustrated by an example of hydroids of the genus *Symplectoscyphus*)\*

A.L. LOBANOV, S.D. STEPANJANTS, and M.B. DIANOV

Zoological Institute, Russian Academy of Sciences, St.Petersburg, Russia.

**SUMMARY:** The paper is a short outline of the interactive computer system BIKEY intended for identification of biological objects. The basic system functions are discussed with special attention for its dialogue program-operated illustrated keys. The system operating procedure is illustrated by an example of key construction for Antarctic species of hydroids of the genus *Symplectoscyphus*.

**Key words:** identification, key, monoentry, multientry, dichotomous, polychotomous, characters, state of character, matrix, hydroids, hydrotheca, gonotheca.

### INTRODUCTION

The first efforts of computer-aided identification of biological objects date back to the end of 60ies (Kiskin *et al.*, 1965; Goodall, 1968; Morse, 1968 etc). As an applied branch of systematics involved in theory and practice of diagnostic key construction, biological diagnostics shaped up into an independent scientific branch in the 70ies resulting from wide spreading of personal computers (Lobanov, 1972; Sviridov, 1973; Pankhurst, 1975; Pankhurst *et al.*, 1984; etc). Later, truly convenient programs for dialogue identification and automatic construction of biological keys began appearing (Pankhurst, 1978, 1986, 1991; Payne, 1978; Payne and Preece, 1980; Watson, 1981; Boswell and Gibbs, 1986; Rubio,

1986; Dallwitz and Paine, 1986; Dedet *et al.*, 1990; Miller and Day, 1990; Schalk, 1992, etc.).

### TOOLS AND PROCEDURES

The present effort to optimize the procedures for identification of Cnidaria utilizes a unique computer system BIKEY (Biological Identification KEYS) which, as evidenced by Raitwyir and Kull (1988), proves to be the first Russian-made dialogue computer diagnostic system. The earlier versions of this system DIAGNOSTICS-1 and DIAGNOSTICS-2 were developed by Lobanov in 1974 (Lobanov, 1975a) for computers highly primitive to the present standards. Subsequently the system has been developed to accommodate more sophisticated computer design (Lobanov *et al.*, 1981; Lobanov, 1983).

\*Received November 18, 1994, Accepted December 10, 1995.

The latest version BIKEY-6, oriented for operation with both the text descriptions and images of taxa, has been worked out by Lobanov and Dianov (1994, 1995) to operate in MS-DOS environment with IBM PC/AT-compatible computers. The BIKEY-6 system is intended for automatically controlled manipulation of various-type diagnostic information on biological taxa and is capable of performing the following functions:

1. biologically convenient input of taxonomic diagnostic data (including images of organisms and their fragments) and its compact storage on magnetic carriers (hereinafter referred to as machine keys);
2. editing and enlargement of initial information in a convenient presentation for user;
3. comprehensive analysis, logical control and assessment of diagnostic adequacy;
4. conduct of optimized object identification by the system-constructed machine keys in "man-computer" dialogue mode. The identification pathway is formed dynamically in the process of dialogue implying selection of characters most convenient for operator-taxonomist and due consideration of recommendation of the program which tend to reduce the identification pathway;
5. automatic computer generation of optimized monentry textual identifiers (one machine key is adequate for receiving an infinite number of textual identifiers since an expert may attribute different importance to different characters);
6. printout in the form of standard diagnosis with respect to a single taxon or taxonomic group (thus an expert is aided in controlling the subject-matter of accumulated diagnostic information);

The BIKEY-6 package is functionally composed of three parts which are self-contained but interconnected through the following common data files:

1. a program package in FoxPro language for checking and analyzing the final textual monentry keys (including those generated outside the BIKEY system);
2. a unique program system in FoxPro for generation, analysis and subsequent refinement of machine multientry keys whose initial information is stored as database of DBF format compatible with dBase, FoxBase, FoxPro, Clipper systems;
3. a number of programs external with respect to FoxPro. They operate in the MS-DOS environment with the final machine keys and execute more complicated operations: image pretreatment, computer-controlled optimized dialogue identification, generation of optimized textual monentry keys for

publication and print-out of taxonomic standardized descriptions.

A block-diagram of connections and interaction between the system component parts is presented in Fig. 1. The central column describes 7 types of data storing files. The arrows pointing from the blocks denote the utilization of data inherent in the block by the relevant programs; the arrows pointing towards the blocks denote a possibility to change these data with the help of programs. Three upper blocks are databases related to the ASCII-format text-files (two central blocks) through a conversion program. The right-side column shows basic programs of the system. The lower blocks are programs operable in MS-DOS environment (all of them utilizing text-files with taxonomic data); the two upper blocks are programs for FoxPro environment. The left-side column represents support programs; the upper two are operable with database in FoxPro environment and the lower two with image files and text files in MS-DOS environment (two lower blocks in central column). The monentry key generation program is capable of transmitting information on the key to the structural database (upper block of central column) for subsequent analysis. The most complicated interactive identification program which makes use of drawings (lower block of right-side column) would require 4 types of data-files.

#### **Programs for checking and analyzing final textual monentry keys.**

They are responsible for input of data to the database, logical checking and analyzing the mathematics parameters of the key. The inputting is proceeded by checking the correctness in the numeration of theses and antitheses. The key analysis involves computation of the average length of identification pathway. Numerical estimate-index of efficiency of key structure (Lobanov, 1975b) varying between 10 and 0 (10 - for complete balance between the tree branches of the key; 0 - for most inadequate shape of the tree, so-called "comb-shape").

#### **Program package for generation, analysis and refinement of machine multientry keys.**

The initial data for BIKEY-6 is prepared in the form of multientry polychotomous digital tables (Lobanov, 1972). A computer key represented by such a table shall consist of three components: a list of taxa; a list of characters and some character

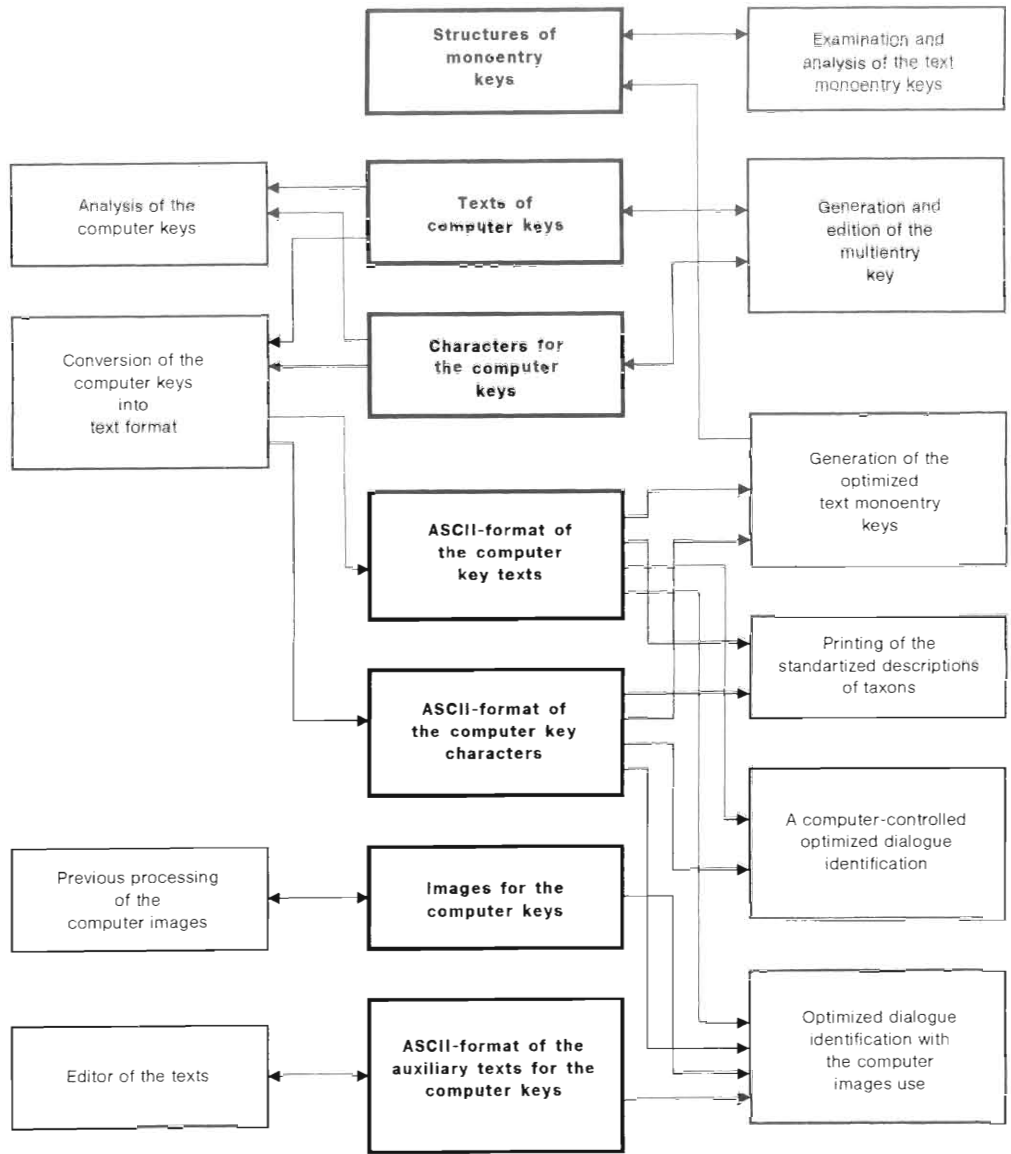


FIG. 1. – Block-Diagram of the connection of the separate parts of BIKEY-6 system. There are the data files into the thick frames, the programs into the thin frames; crossing lines do not join.

states; a digital matrix. In BIKEY-6 each character may be described by no more than 9 character states. In a matrix, rows correspond to taxa, and columns to characters. The row/column intersection is designated by a number which is reported for a particular state of a particular character. If, due to variability, one character of a taxon should be designed by several states, for space saving the state numbers are not divided by commas and are represented as a multiple-digit number. Absence of information on a particular character is denoted by zero. All the programs of the system interpret a zero as the whole set of such character states in the taxon.

The above three components of the computer key are prepared by an operator-taxonomist and introduced into two databases of DBF format using the input program. To assist the operator special operation modes are prescribed for introducing large-sized matrices. A set of characters for one taxon or states of one character for a set of taxa or any rectangular matrix fragment may be introduced at operators option.

Special programs are intended for editing a key. Special care is given to detailed intercomparison between pairs of closely related taxa. A special program enables at one time to edit the whole set of characters of such a pair, or only such characters

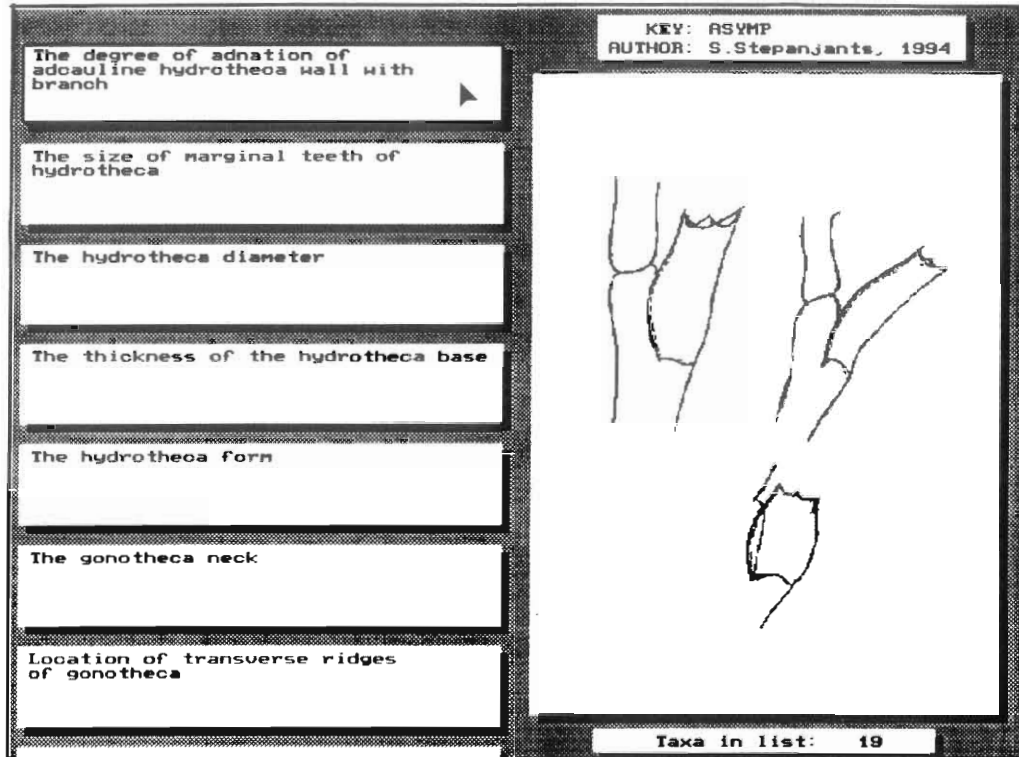


Fig. 2. – The first working screen with names of characters (left frames) and drawings (right frame). The current character is pointed by means of the cursor.

which are not different for the two taxa, or only those where they differ or only such characters where information is missing for at least one of the taxa (zero in the matrix). Several programs are specially developed for key checking. The most important of those programs is designed for checking differences between a particular taxon and the remaining ones by recalling all the possible pairs. As long as probability of indistinguishable pairs exists, a key would not be considered an identifier and would need editing. After elimination of indistinguishable pairs, the same program is utilized for key "safety margin" provisions, i.e. for providing two or more distinguishing characters for each taxonomic pair. The provision of the above safety margin is indispensable for a key to be considered a purely multientry one.

**Operation programs for final machine keys: optimized computer-controlled dialogue identification (DIAKEY and PICKEY programs).**

The most interesting programs in the BIKEY-6 package are dialogue polychotomous multientry keys DIAKEY (operable in the text mode) and PIC-

KEY (operable in graphic mode and actively utilizing pictures). PICKEY is a further modification of DIAKEY and is completely compatible with the latter with respect to utilizable text datafiles. The PICKEY functioning is described below.

The process of computer dialogue identification of a specimen with the aid of the above program comprises a number of steps. At each step the user has to specify the most convenient character for observation (e.g. number of marginal tentacles for medusae) and to select a state of the specified character corresponding to the properties of the examined specimen. The first program screen (Fig.2) comprises text frames with the names of characters, frames with the name of the key and the number of taxa included in the list. Major part of the screen is occupied by a frame containing the illustrations of the current character which are shifted by the cursor. The second operational screen (Fig.3) is separated into several equal areas (between two and nine) filled with illustrations of states of the selected character. The user carefully analyses the relevant element for a specimen identification and chooses the more adequate image for the specimen. The program offers a possibility to check for correct unders-

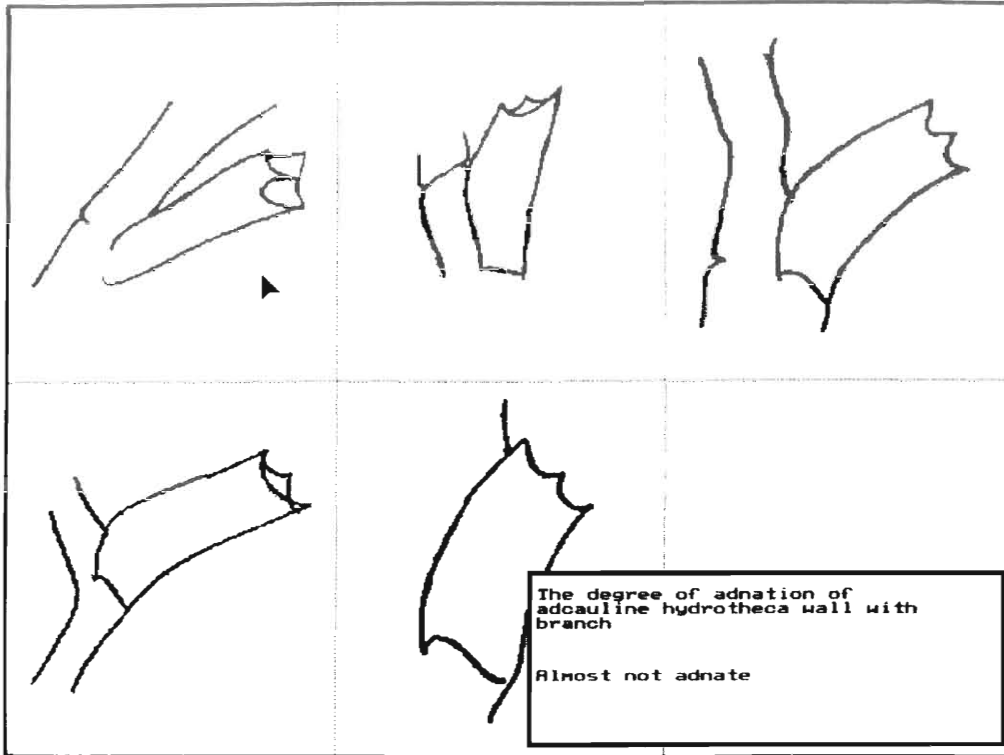


FIG. 3. – The second working screen with the illustrations for the states of the chosen character: in this case for the character 3 - "the degree of adnation of adcauline hydrotheca wall with branches". The cursor points the first state of this character: "almost not adnate".

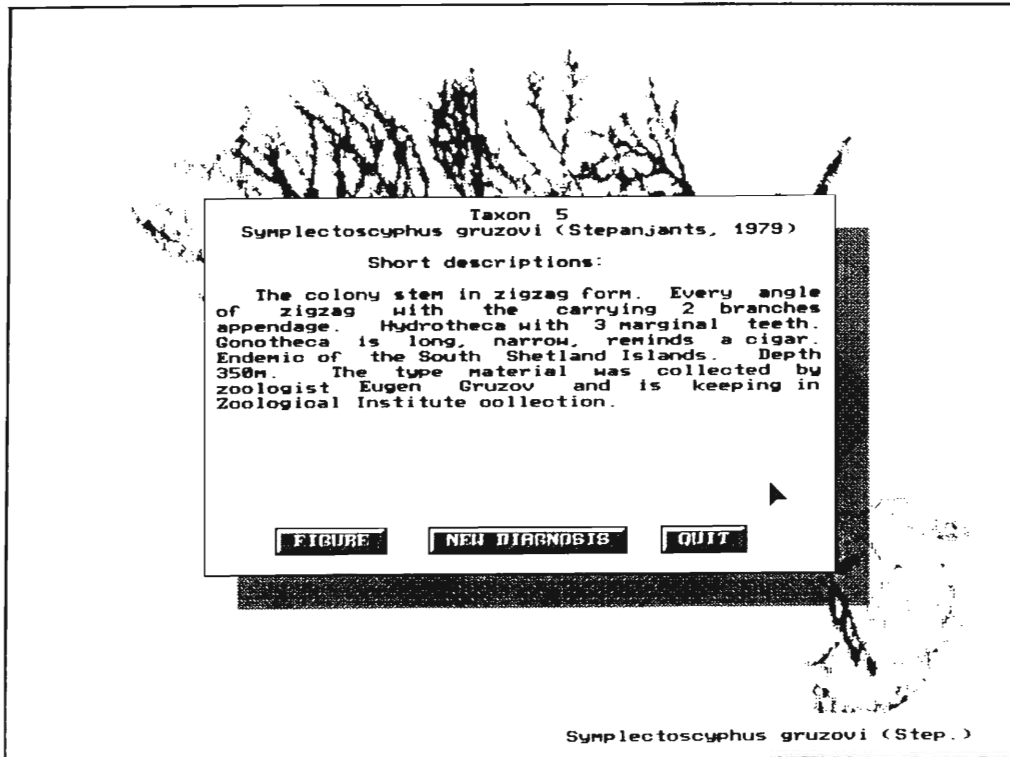


FIG. 4. – Frame with the textual information of taxon. 3 lower buttons permit to continue to work ("image" or "new diagnosis") or to stop to work ("quit").

TABLE 1. – Polychotomous key of the antarctic species of *Symplectoscyphus*

The list of characters	
Ch. 1.	The plane of hydrothecae arrangement 1. In one plane 2. In different planes
Ch. 2.	Hydrothecae arrangement on branch internodes 1. On the one side 2. On the two sides
Ch. 3.	The degree of adnation of adcauline hydrotheca wall with branch 1. Almost not adnate 2. About 1/2 3. About 1/3-1/4 4. About 1/5 5. About 2/3-3/4
Ch. 4.	The hydrotheca form 1. Oblong-cylindrical 2. Curved oblong-cylindrical 3. Curved short-cylindrical 4. Tube-form
Ch. 5.	The length of abcauline hydrotheca wall 1. About 0.20-0.60mm 2. About 0.61-0.90mm 3. About more than 0.90mm
Ch. 6.	The internal marginal teeth of hydrotheca 1. Absent 2. Present
Ch. 7.	The internal abcauline septum of hydrotheca 1. Absent 2. Present
Ch. 8.	The size of marginal teeth of hydrotheca 1. One adcauline and two abcaulines are equally small 2. One adcauline and two abcauline are equally large 3. One adcauline is larger than abcaulines 4. One adcauline and one abcauline are equally larger than middle one 5. Two abcauline are larger than one adcauline
Ch. 9.	The hydrotheca diameter 1. Equal through out the length 2. Narrower near the mouth 3. Wider near the mouth
Ch.10.	The thickness of the hydrotheca base 1. Present 2. Absent
Ch.11.	The gonotheca form 1. Elongate-cylindrical 2. Oval 3. Elongate-oval 4. Costate
Ch. 12.	The gonotheca neck 1. Absent 2. Elongate-cylindrical 3. Tube-form 4. Short cylindrical
Ch. 13.	Location of transverse ridges of gonotheca 1. Absent 2. Over entire length 3. Over part of the length 4. Implicit
Ch. 14.	Number of transverse ridges of gonotheca 1. Absent 2. Several smoothed ones 3. 3-6 4. 7-12 5. 13-27

tanding of the illustration from textual description of the selected character placed on a shifting textual frame. The program selects the taxa possessing the

specified state of the character from the list (at the first step the selection is made from the complete list of taxa included into the key).

In case of two or more taxa included, the program proceeds with the next step. At each step the program proposes the list of characters in the order of decreasing diagnostic value, which is calculated for a particular step by a unique algorithm. It is more advisable to make use of the characters presented by computer in the first place; however, the final decision is taken by the user as a trade-off between safety, accessibility and value of the character. The algorithm for identification of diagnostic value of a character is based on simple formula (Lobanov, 1974):

$$d = \frac{T^2}{\sum_{i=1}^s t_i^2}$$

in which  $d$  = diagnostic value,  $T$  = total number of taxa in the key,  $s$  = number of character states, and  $t_i$  = taxa number which have state  $i$ .

Multiple practical tests have proved the larger efficacy of this formula than the conventional ones (Pimenov, 1987; Sviridov, 1994). The suitable choice of characters would complete the identification process by the demonstration of the total black-and-white and, if possible, colored image of organism as well as the textual description of the taxon (third operating screen, Fig.4).

Strictly speaking, both DIAKEY and PICKEY are not expert systems but they possess numerous features characteristics for the latter and may successfully compete with them in convenience, speed and reliability of identification. Simplicity of the dialogue between user and computer may represent an important advantage of these programs over the genuine expert systems.

#### Operation programs for final machine keys: generation of optimized textual monoentry keys for publication (GENKEY program).

The program algorithm is based on an earlier proposed procedure for constructing a tree of optimal diagnoses (Lobanov, 1983). A stepwise optimal diagnosis is modeled in the key for each taxon. If the properties of a taxon are polymodal with respect to some of the characters (containing more than one state of the character), diagnoses are provided for each of the possible combinations of characters. Pathways of all the diagnoses gradually form a tree in the computer

TABLE 2. – Polychotomous key of the antarctic *Symplectoscyphus* species. The matrix "taxon-character"

NT	Species	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	<i>S. elongatus</i> (Jaed.,1904)	1	2	2	3	1	1	1	4	2	1	2	1	1	1
2	<i>S. spiralis</i> (H.et Gr.,1907)	1	2	2	3	1	2	1	4	2	1	3	1	1	1
3	<i>S. bififormis</i> (Jaed.,1905)	2	1	1	2	2	1	1	2	1	1	3	2	1	1
4	<i>S. mawsoni</i> Briggs,1938	2	2	1	4	2	1	1	3	3	1	0	0	0	0
5	<i>S. gruzovi</i> (Step.,1979)	2	2	1	1	12	1	1	2	1	1	1	1	1	1
6	<i>S. pushi</i> (Step.,1979)	1	2	4	1	3	1	1	2	13	1	0	0	0	0
7	<i>S. naumovi</i> Blanco,1969	1	2	3	2	2	1	1	2	1	2	3	4	4	2
8	<i>S. curvatus</i> (Jaed.,1917-18)	1	2	2	2	2	1	1	2	1	2	4	12	2	3
9	<i>S. plectilis</i> (H.et Gr.,1907)	1	2	3	3	12	1	1	2	2	2	3	24	4	2
10	<i>S. glacialis</i> (Jaed.,1904)	1	2	2	3	1	1	1	1	2	2	2	3	4	2
11	<i>S. confusus</i> Totton,1930	1	2	2	3	1	1	1	2	1	2	4	3	3	4
12	<i>S. vanhoffeni</i> Totton,1930	1	2	2	3	1	1	1	1	1	2	4	3	3	4
13	<i>S. subdichotomus</i> (K.,1884)	1	2	3	3	1	1	12	1	1	2	4	3	3	34
14	<i>S. aggregatus</i> (J.,1916-17)	1	2	5	3	1	1	1	1	3	2	4	4	3	4
15	<i>S. liouvillei</i> (Billard,1914)	1	2	5	3	1	2	1	1	2	1	4	3	2	5
16	<i>S. millardi</i> (Step.,1979)	1	2	5	3	1	1	1	1	1	1	4	3	2	5
17	<i>S. cumberlandicus</i> (J.,1905)	1	2	5	3	12	1	1	2	3	1	4	1	3	4
18	<i>S. magellanicus</i> (M.-T.,1890)	1	2	2	3	12	1	2	1	1	2	0	0	0	0
19	<i>S. unilateralis</i> (Lmx.,1824)	2	2	3	3	12	2	1	5	2	2	2	4	14	12

storage (with a special matrix representation developed for space-saving registration). The tree formation process is terminated by calculating and printing out mathematical parameters of generated monoentry key and subsequently introducing the key itself into the text file in a form adequate for publication (Table 3, 4). The program-formed keys are of somewhat unconventional appearance. Essentially they are bracket-shaped, though the text, general for all the alternatives is removed therefrom and is represented only once before the first alternative.

The program provides a mechanism of swift reevaluation of particular characters by the user. These estimates (prior to final selection of one character from a number of alternatives) are multiplied by the diagnostic evaluations of the characters provided by the program at each step. By changing the weight of a character, the researcher may affect the process of key construction and use one data set for generating an infinite number of different monoentry keys.

## DISCUSSION

Automatic identification keys with application of BIKEY system have been originally constructed for 14 families of medusae from Filifera (with respect to 17 characters) (Lobanov and Stepanjants, 1993); for 7 families of scyphomedusae Coronata (with respect to 10 characters) (Lobanov *et al.*, 1995) and 70 families of medusae (in terms of 72 characters) (Stepanjants, in press). The dichotomous key for the families of medusae of the world served the virtual

base for the construction of the above keys (Stepanjants and Sheiko, 1989). Advantages of machine approach to key compiling are demonstrated in Lobanov and Stepanjants (1993).

In the present paper the philosophy of using BIKEY for automatic construction of diagnostic keys is demonstrated on a small group of species (19) of Antarctic hydroids of the genus *Symplectoscyphus*, practically based on the monograph by Stepanjants (1979) on hydroids of the Antarctic and Subantarctic. The characters were formed (14) with respect to the above 19 species *Symplectoscyphus*; each character comprising 2-5 states (Table 1). On their basis a matrix was formed (Table 2) which itself may serve a key. For instance, *S. pushi* is simultaneously identified by character 3 or 5; *S. gruzovi* by character 11; *S. mawsoni* by 4; state 4 of character 12 ("short cylindrical gonotheca neck") is typical for 3 species: *S. naumovi*, *S. aggregatus* and *S. unilateralis*, these species differing between themselves in characters 3 and 9.

As multientry key, the matrix is adequate for identification of the target by selecting any character mostly accessible for the user according to his professional level as well as level of distinguish ability of this or that character: in this particular case, in fertile colonies, the shape of gonotheca is the best character.

An important advantage of this key is a possibility of its application for identifying damaged specimens. Even separate fragments can be identified if they have retained at least one or several morphological parts which may be suggestive of some matrix characters: for example, it may be an edge of hydrottheca with dentation or a fragment of branch with

TABLE 3. -- The monoentry polychotomous key, composed by GENKEY program in automatic mode from the computer key "Polychotomous key of the antarctic *Symplectoscyphus* species"

1. The degree of adnation of adcauline hydrotheca wall with branch	
1. Almost not adnate	4
2. About 1/2	2
3. About 1/3-1/4	5
4. About 1/5	<i>S. pushii</i> (Step.,1979)
5. About 2/3-3/4	10
2. The size of marginal teeth of hydrotheca	
1. One adcauline and two abcauline are equal small	8
2. One adcauline and two abcauline are equal large	7
3. One adcauline and one abcauline are equal larger than middle one	3
3. The internal marginal teeth of hydrotheca	
1. Absent	<i>S. elongatus</i> (Jaed.,1904)
2. Present	<i>S. spiralis</i> (H.et Gr.,1907)
4. The hydrotheca form	
1. Oblong-cylindrical	<i>S. gruzovi</i> (Step.,1979)
2. Curved oblong-cylindrical	<i>S. bifurcatus</i> (Jaed.,1905)
3. Tube-form	<i>S. mawsoni</i> Briggs,1938
5. The size of marginal teeth of hydrotheca	
1. One adcauline and two abcauline are equal small	
<i>S. subdichotomus</i> (K.,1884)	
2. One adcauline and two abcauline are equal large	6
3. Two abcauline are more large than one adcauline	
<i>S. unilateralis</i> (Lmx.,1824)	
6. The hydrotheca form	
1. Curved oblong-cylindrical	<i>S. naumovi</i> Blanco,1969
2. Curved short-cylindrical	<i>S. plectilis</i> (H.et Gr.,1907)
7. The hydrotheca form	
1. Curved oblong-cylindrical	<i>S. curvatus</i> (Jaed.,1917-18)
2. Curved short-cylindrical	<i>S. confusus</i> Totton,1930
8. The internal abcauline septum of hydrotheca	
1. Absent	9
2. Present	<i>S. magellanicus</i> (M.-T.,1890)
9. The hydrotheca diameter	
1. Equal through the all length	<i>S. vanhoeffeni</i> Totton,1930
2. Narrow near the mouth	<i>S. glacialis</i> (Jaed.,1904)
10. The hydrotheca diameter	
1. Equal through the all length	<i>S. millardi</i> (Step.,1979)
2. Narrow near the mouth	<i>S. liouvillei</i> (Billard,1914)
3. Wider near the mouth	11
11. The size of marginal teeth of hydrotheca	
1. One adcauline and two abcauline are equal small	
<i>S. aggregatus</i> (J.,1916-17)	
2. One adcauline and two abcauline are equal large	
<i>S. cumberlandicus</i> (J.,1905)	

## hydrotheca.

The PICKEY Program is presenting much wider capabilities for dialogue identification of the species *Symplectoscyphus* especially if a taxonomist has inadequate special knowledge or finds it difficult to make use of basic diagnostic characters of this group: hydrotheca, gonotheca etc. By the use of PICKEY a taxonomist may see on the screen the drawings of these colony parts.

In the computer generated polychotomous monoentry key, for identification of 19 species *Symplectoscyphus* (Table 3), preference is given to the character: "degree of adnation of adcauline hydrotheca

Table 4. -- Monoentry polychotomous key, composed by GENKEY program in dialogue with taxonomist from the computer key "Polychotomous key of the antarctic *Symplectoscyphus* species" (the starting values of the 11 character is increased 3 times)

1. Gonotheca form	
1. Elongate-cylindrical	7
2. Oval	2
3. Elongate-oval	3
4. Costate	6
2. Size of marginal teeth of hydrotheca	
1. One adcauline and two abcauline are equal small	12
2. One adcauline and two abcauline are equal large	<i>S. pushii</i> (Step.,1979)
3. One adcauline is more large than two abcauline	<i>S. mawsoni</i> Briggs,1938
4. One adcauline and one abcauline are equal larger than middle one	<i>S. elongatus</i> (Jaed.,1904)
5. Two abcauline are more large than one adcauline	<i>S. unilateralis</i> (Lmx.,1824)
3. Degree of adnation of adcauline hydrotheca wall with branch	
1. Almost not adnate	5
2. About 1/2	4
3. About 1/3-1/4	9
4. About 1/5	<i>S. pushii</i> (Step.,1979)
4. Internal marginal teeth of hydrotheca	
1. Absent	<i>S. magellanicus</i> (M.-T.,1890)
2. Present	<i>S. spiralis</i> (H.et Gr.,1907)
5. Hydrothecae arrangement on branch internodes	
1. On the one side	<i>S. bifurcatus</i> (Jaed.,1905)
2. On the two sides	<i>S. mawsoni</i> Briggs,1938
6. Degree of adnation of adcauline hydrotheca wall with branch	
1. Almost not adnate	<i>S. mawsoni</i> Briggs,1938
2. About 1/2	10
3. About 1/3-1/4	<i>S. subdichotomus</i> (K.,1884)
4. About 1/5	<i>S. pushii</i> (Step.,1979)
5. About 2/3-3/4	14
7. Degree of adnation of adcauline hydrotheca wall with branch	
1. Almost not adnate	8
2. About 1/2	<i>S. magellanicus</i> (M.-T.,1890)
3. About 1/5	<i>S. pushii</i> (Step.,1979)
8. Hydrotheca form	
1. Oblong-cylindrical	<i>S. gruzovi</i> (Step.,1979)
2. Tube-form	<i>S. mawsoni</i> Briggs,1938
9. Hydrotheca form	
1. Curved oblong-cylindrical	<i>S. naumovi</i> Blanco,1969
2. Curved short-cylindrical	<i>S. plectilis</i> (H.et Gr.,1907)
10. Size of marginal teeth of hydrotheca	
1. One adcauline and two abcauline are equal small	13
2. One adcauline and two abcauline are equal large	11
11. Hydrotheca form	
1. Curved oblong-cylindrical	<i>S. curvatus</i> (Jaed.,1917-18)
2. Curved short-cylindrical	<i>S. confusus</i> Totton,1930
12. Internal abcauline septum of hydrotheca	
1. Absent	<i>S. glacialis</i> (Jaed.,1904)
2. Present	<i>S. magellanicus</i> (M.-T.,1890)
13. Internal abcauline septum of hydrotheca	
1. Absent	<i>S. vanhoeffeni</i> Totton,1930
2. Present	<i>S. magellanicus</i> (M.-T.,1890)
14. Hydrotheca diameter	
1. Equal through the all length	<i>S. millardi</i> (Step.,1979)
2. Narrow near the mouth	<i>S. liouvillei</i> (Billard,1914)
3. Wider near the mouth	15
15. Size of marginal teeth of hydrotheca	
1. One adcauline and two abcauline are equal small	<i>S. aggregatus</i> (J.,1916-17)
2. One adcauline and two abcauline are equal large	<i>S. cumberlandicus</i> (J.,1905)



wall with branch". In our opinion, it does not seem to be the most appropriate character but the computer has selected it according to the highest diagnostic value. In the result, this machine key has the average identification pathway of 2.632. Table 4 presents another key generated with the interference of taxonomist who arbitrarily enhanced the diagnostic value of the character: "the gonotheca form". In his opinion, it is a more appropriate character for starting the identification pathway. The resulting key (Table 4) has the pathway length of 2.857 which is somewhat inferior to the first alternative with respect to formal indicator but more convenient for the taxonomist.

## CONCLUSIONS

An optimization procedure for biological species identification using a unique computer system BIKEY is described through the case of 19 *Symplectoscyphus* species (Cnidaria, Hydrozoa).

The most taxon-informative characters specified by a taxonomist and integrated into a matrix are the base of the proposed computer system.

The matrix itself may serve as multientry key permitting to identify even damaged specimens or fragments, starting the identification process from any accessible character. At the same time, the matrix represents a space-saving form of taxonomic data and a base for compiling a description of a taxon, its diagnosis and the comparison with closely related taxa.

DIAKEY and PICKEY programs permit to perform identification in dialogue with the computer. It is important that DIAKEY and PICKEY are responsible for the identification pathway, selecting the most important diagnostic character, at the same time leaving the right of final decision to the taxonomist. These programs permit to identify one target from various sets of characters, thus promoting a much better identification reliability.

PICKEY seems especially promising due to the graphic facilities for the identification purposes. This program uses images as basic information while text description is of secondary, explanatory importance. The advantages of such approach were evident in the identification of the families of medusae Coronata and species of hydroids of the genus *Symplectoscyphus* where the important morphological details may be mostly unfamiliar to a non-expert user.

There is a possibility of making automatic compilation of monentry polychotomous keys, either completely computer-generated (the computer selects the most meaningful character) or with the interference of the user who makes a final decision on the most diagnostically valuable character.

## ACKNOWLEDGMENTS

The authors wish to express their cordial gratitude to Drs. J. Watson, S. Piraino, P.F.S. Cornelius and 3 anonymous consultants that reviewed this manuscript for their valuable advices. The authors are indebted to Mr. S. Solovjev for his help in translation of this manuscript.

## REFERENCES

- Boswell K. and A.J. Gibbs. – 1986. *The VIDE data bank for plant viruses. Development and Applications in Virus Testing*. Association of Applied Biologists, London.
- Dallwitz M.J., T.A. Paine. – 1986. *User's Guide to DELTA System. A general system for coding taxonomic description*. Division of Entomology, Report N 13 (Third Edition), CSIRO, Canberra.
- Dedet J.P., J. Lebbe, and R. Vignes – 1990. C.I.P.A.: A multinational program for computer identification of phlebotomine sandflies of America. *Bull. Soc. fr. parasitol.* 8: 1131.
- Goodall D.W. – 1968. Identification by computer. *BioScience*, 18: 485-488.
- Kiskin P.H., I.N. Pecherskaja, and Yu.N. Pecherskyi. – 1965. Computer-aided ("Minsk-I") diagnostic searching of grapes grades. *Wine making and viticulture in the USSR*, 1: 21-22 (in Russian).
- Lobanov A.L. – 1972. The logical analysis and classification of the existing diagnostic keys. *Entom. Obozr.*, 51: 668-681 (in Russian).
- Lobanov A.L. – 1974. Assessment of diagnostic value of sets of characters in multientry keys intended for computer usage. In: *Summary rep. VI Komi Rep. Young Scientists Conference*, pp. 125-126. U.S.S.R. Acad. Sci., Syktyvkar. (in Russian).
- Lobanov A.L. – 1975a. Results of experiments with biological diagnostic systems based on "Nairi-C" computer. In: T.V. Zabojeva. *Biological Research in the North East of the European Part of the USSR*, pp. 162-168. Syktyvkar. U.S.S.R. Acad. Sci. (in Russian).
- Lobanov A.L. – 1975b. Mathematical aids for evaluation, estimation and comparison of structural parameters of diagnostic keys. *Zool. J.*, 54: 485-497 (in Russian).
- Lobanov A.L. – 1983. Principles of computer-aided construction of insects' identification. PH.D. Thesis L., ZIN AN SSSR: 1-19 (in Russian).
- Lobanov A.L., M.B. Dianov. – 1994. Dialogue computer diagnostic system BIKEY and its potential application in entomology. *Entom. Obozr.*, 73: 465-478 (in Russian).
- Lobanov A.L. and M.B. Dianov. – 1995. Computerized diagnostic system BIKEY and its potentialities as plant identifiers. *IInd Conference "Computer databases in botanical investigations"*, St. Petersburg, 17-19 April 1995. Abstracts: 29-30 (in Russian).
- Lobanov A.L., W.F. Schilow, and L.M. Nikritin. – 1981 - Zur Anwendung von Computern für die Determination in der Entomologie. *Deutsch. Entomol. Z.*, 28: 29-43.
- Lobanov A.L., S.D. Stepanjants. – 1993. Optimization of key of families of suborder Filifera medusae (Hydrozoa, Athecata) with the computer-aided diagnostic system. In: S. Stepanjants (ed.), *Marine Plankton. Taxonomy, ecology, distribution, Issled. Fauni Morey*, 45: 38-50 (in Russian).

- Lobanov A.L., S.D. Stepanjants, and M.B. Dianov – 1995. In: S.D. Stepanjants (ed.), *Cnidaria. Modern and perspective investigation. Trudi Zool. Inst. Rus. Acad. Sci.*, 261: 20-45 (in Russian).
- Miller M.G., E.R. Day. – 1990. Interactive taxonomy: name that bug in three touches or less. *Amer. Entomol.*, 36: 219-224.
- Morse L.E. - 1968. Construction of identification keys by computer. *Amer. J. Botan.*, 55: 737.
- Pankhurst R.J. – 1975. Biological Identification with computers. Syst. Assoc. Special Volume. 7. Academic Press. London.
- Pankhurst R.J. – 1978. Biological Identification. The principles and practice of identification methods in biology. Edward Arnold. London.
- Pankhurst R.J. – 1986. A package of computer programs for handling taxonomic databases. *CABIOS*, 2: 33-39.
- Pankhurst R.J. – 1988. Taxonomic biology seen as a database. In: *Classification and Related Methods of Data Analysis*. pp. 709-716. Academic Press. London.
- Pankhurst R.J. – 1991. *Practical taxonomic computing*. Cambridge University Press. Cambridge.
- Pankhurst R.J., M.J. Dallwitz, and R.W. Payne. – 1984. Notice to users of computer programs for identification and description. *Taxon*, 33: 556.
- Payne R.W. – 1978. *GENKEY. A program for constructing and printing identification keys and diagnostic tables*. Rothamsted Experimental Station. Harpenden
- Payne R.W., and D.A. Preece. – 1980. Identification keys and diagnostic tables: a review. *J. Roy. Statist. Soc., Ser. A*, 143: 253-292.
- Pimenov M.G. – 1987. The mathematical methods and computerized technology in the higher plants systematics. In: Results of the Science and the Technology. *Botany*, 8: 96.
- Raitvyir A., and K. Kull. – 1988. Dialogue key for fungal identification, realized on a personal computer. *Mycol. Phytopat.*, 1: 111-117 (in Russian).
- Schalk P.H. – 1992. Systematic goes multimedia: computerizing biological knowledge. *ASC Newsl.*, 20: 117-119.
- Stepanjants S.D. – 1979. Hydroids of the Antarctic and Subantarctic waters. *Biol. Result. Soviet Antarct. Exped. 6. Explor. Fauna of the Sea N° 20* (in Russian).
- Stepanjants S.D., and O.V. Sheiko – 1989. The identification key of the classes, orders and families. In: M.G. Petrushevskaya and S.D. Stepanjants (eds.), *Marine plankton. Taxonomy, ecology, distribution. Issled. Fauni Morey.*, 41: 100-131 (in Russian).
- Stepanjants S.D. – 1996. Identification of medusae families and some hydroids species with the BIKEY system help. In: EDITOR. *Data bases and the computer graphics in zoological investigations. Trudi Zool. Inst.*, 30: in press (in Russian).
- Sviridov A.V. – 1973. Interrelationship between biological diagnostics and taxonomy. *J. General Biol.*, 34: 900 - 906. (in Russian).
- Sviridov A.V. – 1994. *Keys in biological taxonomy: theory and practice*. Moscow University. Moscow. (in Russian).
- Watson L. – 1981. An automated system of generic descriptions for Caesalpinioidea and its application to classification and key making. *Adv. Legume Syst.*, 2: 65-81.