Nuclear Data for Science and Technology (1988 MITO), 497- 500, Copyright © 1988 JAERI.

EXPERT SYSTEM FOR EVALUATION OF EXPERIMENTAL UNCERTAINTY FROM EXFOR FILE

Yuji Uenohara, Mitsuhaya Tsukamoto, Toshiya Mori, Mitsuru Kihara, and Yukinori Kanda

Department of Energy Conversion Engineering, Kyushu University
Kasuga, Fukuoka 816, Japan

Abstract: An expert system have been designed to estimate experimental uncertainties from comments and numerical data in EXFOR files. The expert system designed in the present work has knowledge bases and inference engines written in the computer programming language LISP. Scarce information in the report to be demanded in evaluation of errors is inferred and implemented from the other information given in the same report. The expert system is programmed to conduct reasonably these processes. Typical examples are presented.

(Expert System, Production Rule, LISP, Covariance, Nuclear Data)

Introduction

Covariances of evaluated nuclear data are important in the field of nuclear engineering. Error information on experimental cross section data are important sources for the evaluation of covariances. Systematic errors of experimental cross section data play a significant role in the evaluation of covariances. It, however, is impossible to evaluate accurately systematic errors. The information on systematic errors are given in nuclear reaction database. They, however, are not always described in detail. Ones hardly obtain the information on systematic errors in the nuclear reaction database. The systematic errors, however, show patterns with respect to experimental conditions e.g. facilities, detectors, and so on. Evaluators can roughly deduce systematic errors. The deduction is performed by using their abundant knowledge and experiences. This work becomes hard and troublesome with the number of experimental data. Many experts on the evaluation are required for the evaluation of systematic errors. A mechanical evaluation of systematic errors interests us. The knowledge cannot be almost represented by mathematical formulas and numerical data. The traditional computer programing techniques are unsuitable to the development of it.

A knowledge engineering has recently been developed. This technique can be applied in many fields. An expert system is the most successful technique in the knowledge engineering. We have tried to design the expert system for evaluation of the systematic errors. The knowledge base and inference engine of the expert system are written in the computer programing language LISP. The systematic errors of experiments of some activation cross section data for ^{55}Mn in EXFOR[1] file were evaluated for an example.

Conversion of EXFOR File

An EXFOR(EXchange FORmat) is the format for experimental nuclear reaction database. The numerical data of experimental cross section are almost stored and circulated to users in the EXFOR format. This database includes experimental conditions in detail. The EXFOR file is a set of entries, i.e. reports. An ENTRY is a set of some subentries, i.e. experiments. A SUBENTRY is an elemental unit identifying experiment. The example of EXFOR file is shown in Fig. 1. The bibliography and experimental conditions are stored in the BIB section. The keywords identify the experimental and bibliographic information in EXFOR file. Keywords are specified by codes and free documentations. The structure of component in BIB section is written by

```
KEYWORD, CODE-1, FREE_DOCUMENTATION-1
         CODE-2, FREE_DOCUMENTATION-2
           .
           .
         CODE-n, FREE_DOCUMENTATION-n
```

We represent this structure as

```
(KEYWORD  (CODE-1 CODE-2 ·· CODE-n)
          (FREE_DOCUMENTATION-1
           FREE_DOCUMENTATION-2
             .
             .
           FREE_DOCUMENTATION-n)
                               )
```

by the LISP. If code and free documentation are not given, they are replaced by a NIL, which means a null set. The changed database into LISP is shown in Fig. 2.

Uncertainty Evaluation System

The present system is designed by using the technique of the production system. A production system in a knowledge engineering is characterized by an inference engine and knowledge base[2]. The traditional computer program was composed of the program and data parts. The procedure of problem solving was only in the program part. In the case of production system, it is in the data part. It is represented by the form of knowledge. A set of knowledge is called knowledge base. It is composed of facts and rules. The program part is called inference engine. The inference engine deduces new facts by using knowledge base. The block diagram of this system is shown in Fig. 3. The implementation of the system by LISP is referred to the reference[3].

Inference Engine

The flow chart of the inference engine is shown in Fig. 4. The pre-processor, FREE

DOCUMENTATION ANALYSER processes the free documentations in the EXFOR file.

It is easy to implement the system processing the CODE in the EXFOR file. Many information are available in the free documentation. The key words, ERR-ANALYS and CORRECTION, are specified by few kinds of code. They are almost expressed by free documentations. Since the free documentations are hardly analysed by the traditional computer programming technique, the system accesses frequently the operator. In order to avoid the such redundant operator's works, we tried to implement the analyser of free documentation by applying a theory of CD, "Conceptual Dependency", representation[4] to the present system.

This system is composed of sub systems, PARSER and MATCHING. The PARSER converts the natural language representation to the CD representation. The example for operations of PARSER is shown in Fig. 5. If the sentence being processed is

ERRORS ARE STATISTICAL,

the structure shown in Fig. 5 is built. The words ERRORS, ARE, and STATISTICAL have been stored in the dictionary by the form of CD representation , priorily. In the present system, the word ARE is equivalent to the word IS. The significant information is given by the nouns and the verbs are the key words to specify the structure of natural language representation. The structure of the CD formed word, "IS", is characterized by slots, ACTOR and STATE, and their slots are variables. Driving the PARSER, the ACTOR and STATE are specified as ERRORS and STATISTICAL, respectively. Consequently, the sentence is converted into the CD representation shown in the bottom of Fig. 5. Since the converted free documentation is written by the LISP, it can be easily processed by computers.

If the analyser fails to generate the CD representation, the operator reads the free documentation and takes the available information. The processes after the analyser is descirbed in the follwing chapters, Knowledge Base and Examples in detail.

Knowledge Base

The knowledge bases of present system are categorized into three types as shown in Fig. 3.

The first is the knowledge for searching keywords in EXFOR file. These knowledge are written by the form of production rule. The inference engine is driven by using these knowledge. If the keyword, ERR-ANALYS, is not found in a given subentry, the inference engine searches every keywords concerned with the systematic errors by the backward reasoning. A part of them are shown in Fig. 6. The second is a global knowledge with respect to cross section measurements. The code for keyword are specified by using these knowledges. A part of them is shown in Fig. 7. They are represented by the form of frames. The third is the table of systematic errors.

Example

In this example, the present system estimates a partial error derived from estimation of neutron flux, estimation of detector efficiency, and standard cross sections. The example experiment is the measurement of ^{55}Mn(n,p) cross section. In order to test the present system, the keywords of FACILITY, DETECTOR, and MONITOR are masked by adding the asterisk,* , to the head of these key words. The knowledges used in this example are shown in Fig. 6 and 7. The outputs of the operation are shown in Fig. 8. The system tries to match between ERR-ANALYS and keywords in this subentry. Since it was not found in this sub entry, the inference engine searches for the keywords relating with detector error. The first rule in Fig. 6 tells us that the key word DETECTOR is required for the deduction of detector errors. The keywords DETECTOR is not found in this subentry. In the next step, the inference engine searches the keywords for deduction of DETECTOR. The second rule tells us that the key word EXP-YEAR is required for the deduction of DETECTOR. This keyword is given in this subentry. The inference engine concludes that the estimation of detector error enable. This process is backward reasoning. The detector is specified by the global knowledge shown in Fig. 7. As the same manner, this system searches the experimental conditions required for the estimation of the error derived from estimation of neutron flux and standard cross section. Obtaining these information on experimental conditions, the values of systematic errors are decided by the error tables. In this example, the detector and facility are deduced accurately. The standard cross sections, however, disagree with that of original database. In the measurements of activation cross sections, the detector and facility can be specified easily. For example, there are few experiments performed with Ge(Li) detector before 1970. The almost experiments in which the one data or two are measured around 14 MeV neutron energy is performed with Cockroft-Walton's apparatus. On the contrary, the rules for use of standard cross sections cannot be found easily.

Concluding Remark

We have designed an expert system for evaluation of systematic uncertainty on experimental cross section data and tried to construct the prototype of the system.

The EXFOR format can be converted to LISP, easily. Using LISP interpreter, the inference engine and knowledge base can be implemented to the main frame FACOM780 and MS-DOS machines, easily. The LISP is very powerful tool to program this system.

The knowledge base is the most important to make this system fit for practical use. The knowledge of neutron cross section measurement needs to be collected internationally for the purpose of improvement of quality. Many discussions by experts on experiments are necessary to polish up the knowledge. It is expected that the discussions contribute to the improvement of experimental technique.

```
ENTRY            11010       820513
SUBENT         11010001      820106
BIB                  10          12
INSTITUTE   (1USAANL)
REFERENCE   (J,PR,82,69,5104)
AUTHOR      (V.HUMMEL, B.HAMERMESH)
TITLE       ACTIVATION CROSS SECTIONS MEASURED
             WITH ANTIMONY-BERYLIUM NEUTRONS

                          .
ENDSUBENT            15           0
SUBENT         11010004      820106
BIB                   1           1
REACTION    (25-MN-55(N,G)25-MN-56,,SIG,,SPA)
ENDBIB                1           0
                          .
                          .
ENDENTRY
```

Fig. 1. The BIB section in EXFOR file. The strings in the first to 11th columns are keywords. The next strings enclosed by parenthesis are codes. The others are free documentations.

```
(ENTRY 11010 820513
 ((SUBENT 11010001 820106

    ((INSTITUTE ((1USAANL )) (NIL))
     (REFERENCE ((J,PR,82,69,5104 )) (NIL))
     (AUTHOR ((V'HUMMEL, B'HAMERMESH )) (NIL))
     (TITLE (NIL) ((ACTIVATION CROSS SECTIONS
       MEASURED WITH ANTIMONY-BERYLIUM NEUTRONS)))
                          .
                          .
   ))
  (SUBENT 11010004 820106
    ((REACTION ((25-MN-55{N,G}25-MN-56,,SIG,,SPA
     )) (NIL))
   ))
                          .
                          .
))
```

Fig. 2 The EXFOR files changed into LISP. Since the parentheses and peiod are reserved words for the LISP, they are changed into braces and back-quotation mark, respectively.

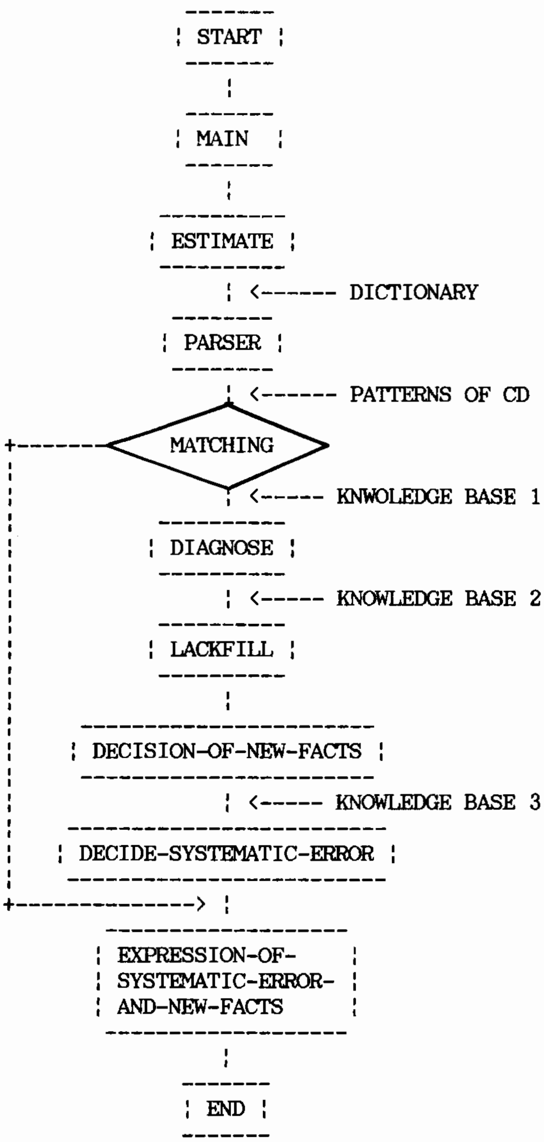

Fig. 4. The flow charts of the present inference engine. The sub routine ESTIMATE searches for the keywords, CORRECTION and ERR-ANALYS in a given EXFOR file. The PARSER changes the free documentations into CD representation. The MATCHING searches for the error information in the CD representation. The DIAGNOSE tests the hypothesis that the systematic errors are available. The LACKFILL deduces the experimental conditions by the Forword Reasoning. The DECIDE-SYSTEMATIC-ERROR gives the values of systematic errors.

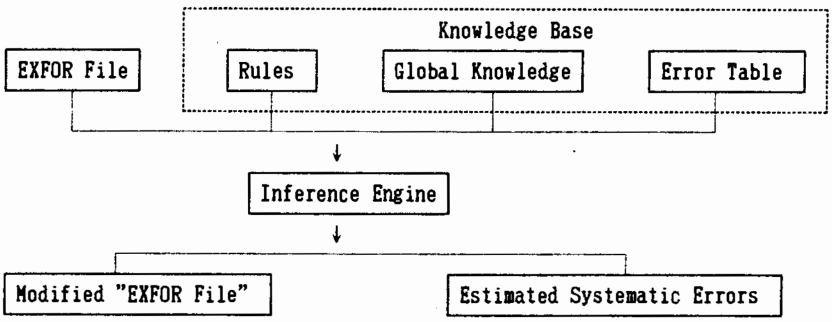

Fig. 3 The block diagram of the present system.

```
>(parser)
 (INPUT IS (ERRORS ARE STATISTICAL))
 PROCESS ERRORS
         *CDFORM* (ERROR (NAME (ERRORS)))
         *PARTOFSPEECH* NOUN
 (nextword test)
 success
 PROCESS ARE
         *PARTOFSPEECH* VERB
         *CDFORM* (IS (ACTOR (*VAR* ISVAR1))
                      (STATE (*VAR* ISVAR2)))
         ISVAR1 (ERROR (NAME (ERRORS)))
         ISVAR2 nil
 (nextword test)
 success
 PROCESS STATISTICAL
         *PARTOFSPEECH* ADJECTIVE
         *CDFORM* STATISTICAL
         ISVAR2 STATISTICAL
 (CD-FORM IS)
 (IS (ACTOR (ERROR (NAME (ERRORS))))
     (STATE STATISTICAL))
 (End of parser)
 nil
>
```

Fig. 5 The example for the CD representation. The sentence "ERRORS ARE STATITICAL" was processed. The words ERRORS, ARE, and STATISTICAL have been stored in the dictionary by the form of CD representation , priorily. The structure of the CD formed word, "IS", is characterized by slots, ACTOR and STATE and their slots are variables. Driving the PARSER, the ACTOR and STATE are specified as ERRORS and STATISTICAL, respectively.

```
(rule identify1
  (if (DETECTOR)) (then (DETECTOR-ERROR)))
(rule identify2
  (if (EXP-YEAR)) (then (DETECTOR)))
(rule identify3
  (if (MONITOR)) (then (FLUX-ERROR1)))
(rule identify4
  (if (FACILITY)) (then (FLUX-ERROR2)))
(rule identify4
  (if (FACILITY) (INC-SOURCE)) (then (MONITOR)))
```

Fig. 6. Production rules for inference of unknown keywords. The 5th rule means if keywords, INSTITUTE and EXP-YEAR are given in SUBENTRY, the facility can be found.

```
(
  (EXP-YEAR ((70 100)) (NIL))
  (DETECTOR ((GELI)) (NIL))
)
(
  (EXP-YEAR  ((0 75)) (NIL))
  (DETECTOR  ((NAICR)) (NIL))
)
```

Fig. 7 Global knowledges for detector. The first knowledge means that few experiments were performed with Ge(Li) detector before 1970.

```
>(main)
(identify2 (DETECTOR))
(identify1 (DETECTOR-ERROR))
(HYPO (DETECTOR-ERROR) is-true)
Hypotheses-estimation-is-end
((DETECTOR-ERROR DETECTOR) (DETECTOR EXP-YEAR))
Fill-lack-of-information

(identify6 (FACILITY))
(identify5 (MONITOR))
(identify3 (FLUX-ERROR1))
(HYPO (FLUX-ERROR1) is-true)
Hypotheses-estimation-is-end
((FLUX-ERROR1 MONITOR FACILITY)
 (MONITOR FACILITY INSTITUTE))
Fill-lack-of-information

(identify4 (FLUX-ERROR2))
(HYPO (FLUX-ERROR2) is-true)
Hypotheses-estimation-is-end
((FLUX-ERROR2) (FACILITY))
Fill-lack-of-information

((8 %))
((8 %))
((2 %))
>fact
((FACILITY ((CCW)) (NIL))
 (DETECTOR ((NAICR)) (NIL))
 (MONITOR ((29 -CU-65{N,2N}29-CU-64,,SIG)) (NIL))
 (INSTITUTE ((3INDMUA )) (NIL))
 (EXP-YEAR ((69 )) (NIL))
 (*FACILITY ((CCW )) (NIL))
 (INC-SOURCE ((D-T )) (NIL))
 (*DETECTOR (( NAICR )) (NIL))
 (*MONITOR ((26-FE-56{N,P}25-MN-56,,SIG )) (NIL))
 (COMMENT (NIL) ((NOT GIVEN)))
 (REACTION ((25-MN-55{N,P}24-CR-55,,SIG )) (NIL))
 (RANGE (14 14) (NIL))
 (POINT ( 1) (NIL))
)
```

Fig. 8 Operation of the present system. The under scored lines are operator's input. The three groupes at the top of figure bounded by space lines are the traces of the inference of uncertainties on detector efficiency, on monitor, and on neutron source, respectively. The next three lists are the values of estimated systematic erros. The inferred facts are enclosed by a box. The keywords of FACILITY, DETECTOR, and MONITOR are masked by adding the asterisk in the original EXFOR file.

References

1. H.D. Lemmel, ed., NDS EXFOR Manual, International Atomic Energy Agency Nuclear Data Services, 1979.

2. E. Rich, "Artificial Intelligence", McGraw-Hill Book Company, 1983.

3. P.H. Winston and B.K.P. Horn, "LISP", Addision-Wesley Publishing Company, Inc., 1981.

4. R.C. Schank and C.K. Riesbeck, "Inside Computer Understanding: An inquiry into Human Knowledge Structure", Lawlence Erlbaum Associates, 1977