

IAEA-196

INIS: THESAURUS MAINTENANCE SYSTEM

BY

A. CHEPKASOV

(International Atomic Energy Agency, Vienna, Austria)



**A TECHNICAL DOCUMENT ISSUED BY THE
INTERNATIONAL ATOMIC ENERGY AGENCY, VIENNA, 1977**

INIS: THESAURUS MAINTENANCE SYSTEM, IAEA, VIENNA, 1977

**Printed by the IAEA in Austria
March 1977**

P R E F A C E

The origin of the INIS thesaurus dates to the late 1960's when, in its first form, it was an adaptation of the Euratom thesaurus modified according to INIS requirements. From the very beginning, the INIS thesaurus was implemented in computer-readable form which allowed for its computerized maintenance.

The first design of the computer internal form of representation of the thesaurus, however, did not allow for the full execution of numerous checks of validity of interrelationships between terms of the thesaurus on various hierarchical levels. This could lead to the existence of annoying anomalies in the structure of the thesaurus.

During the years 1974 - 1975 another computer internal form of representation of the thesaurus was designed by the author of this document. Accordingly, a new set of computer programs was developed to generate the thesaurus in this new form, to update it, and to transform the updated version into a form acceptable for use.

This set of programs is now in operation and forms the INIS Thesaurus Maintenance system. The main feature of the system is flexibility, which allows it to be used not only for INIS-like thesauri, but also for the development and maintenance of thesauri with a range of different structures.

This document describes the design and the operation of the system, which is independent of the subject area covered by the thesaurus terminology and ignores totally the intellectual aspects of terminology control. The document is intended for systems analysts and programmers who are involved in the design and development of computer systems for thesaurus maintenance and for whom the logic of the system described here, may be of interest.

The document contains full program listings as appendices, while the text is laid out as a logical development starting from consideration of abstract notions of relations on sets, through the development of an algorithm for checking the validity of chains of these relations, to the practical implementation of the algorithm in the programs.

The author is very grateful to the staff of the INIS section of the IAEA and particularly to Mr. D. Davis, former thesaurus specialist, and to Mr. V. Gadjokov, former subject specialist, for numerous suggestions and close collaboration at the time the system was designed and programmed.

The author is especially grateful to Mr. G. Timukhin, a former consultant at the IAEA, for invaluable assistance in compiling this document; without him it could not have been completed.

The author is also grateful to Mr. J. Crowell, the head of the IAEA Computer Section, who read the first draft, and to Mr. C. Todeschini, the present INIS thesaurus specialist, who undertook the final editing and made many useful comments.

TABLE OF CONTENTS

Page No.

1.	INTRODUCTION.....	1
1.1.	<u>Terms and their classification</u>	2
1.2.	<u>Relations</u>	3
1.3.	<u>Properties of the relations</u>	5
1.3.1.	Reciprocity of relations.....	5
1.3.2.	Symmetry of relations.....	5
1.3.3.	Transitiveness of relations.....	6
1.3.4.	Rules for defining the specification of relations.....	7
1.3.5.	Contradiction and excessiveness of relations.....	8
1.4.	<u>Chains of relations</u>	10
1.4.1.	The transparency of relations.....	11
1.4.2.	The valency of relations.....	13
2.	FUNCTIONS OF THE TMS.....	19
2.1.	<u>The U-form of the thesaurus</u>	21
2.2.	<u>The W-form of the thesaurus</u>	23
2.2.1.	The dictionary data set.....	23
2.2.2.	The cross-reference data set.....	26
2.2.3.	The Scope Note data set.....	27
2.2.4.	Linkage between files of the W-form.....	28
2.3.	<u>The T-form of the thesaurus</u>	31
3.	THE PROCEDURES TO ACCESS RECORDS ON THE XF-FILE.....	34
3.1.	<u>The READER procedure</u>	36
3.2.	<u>The WRITER procedure</u>	37

4.	THE PROGRAM THUPDTE.....	40
4.1.	<u>Action 1.</u> Add term, scope note or relation.....	46
4.1.1.	The procedure LCHECK.....	53
4.2.	<u>Action 2.</u> Delete terms, scope note or relation.....	65
4.2.1.	The procedure-function TCHECK.....	68
4.3.	<u>Action 3.</u> Print a block of information about a Term..	70
4.3.1.	The procedure Printer.....	70
4.4.	<u>Action 4.</u> Replace the spelling of a Term.....	73
4.5.	<u>The procedure IGNORE</u>	76
5.	THE THBKUP PROGRAM	77
5.1.	<u>UTABLE procedure</u>	84
5.2.	<u>THDEV procedure</u>	85
6.	<u>CREATION OF THE W-FORM (WORK FORM) OF THE THESAURUS</u>	87
6.1.	<u>The program THBASE1</u>	89
6.2.	<u>The program THBASE2</u>	93
6.3.	<u>The program THBASE3</u>	97
6.4.	<u>The program THBASE4</u>	98
APPENDIX 1 - Program Source Listings.....		100
	Program WRITER.....	101
"	READER.....	103
"	THBASE1.....	105
"	THBASE2.....	109
"	THBASE3.....	112
"	THBASE4.....	115
"	THUPDTE.....	118
"	THBKUP.....	142
"	THDEV.....	150